

Table of Contents

Abbreviated Table of Contents	ii
Summary	iii
Acknowledgements	iv
Table of Contents	v
List of Tables	x
List of Figures.....	xv
Abbreviations used in the study	xvii
List of Appendices	xviii
Chapter 1: Overview and Orientation	1
1.1 Introduction	1
1.2 Purpose of the study.....	1
1.3 Defining core concepts and terminology	2
1.3.1 Cognitive load	2
1.3.2 An introduction to style	4
1.3.3 Cognitive style.....	5
1.3.4 Learning style.....	5
1.3.5 Multimedia learning.....	5
1.4 Background to the study.....	6
1.5 Rationale of the research	8
1.5.1 The empirical imperative	8
1.5.2 The theoretical imperative	9
1.5.3 The methodological imperative	9
1.5.4 The media imperative	9
1.5.5 The contextual imperative	10
1.5.6 Other considerations	13
1.5.7 Pulling it together	13
1.6 Research questions.....	14
1.7 Research design and methodology	14
1.8 Analysis of the data	17
1.9 Limitations and strengths of the research.....	17
1.10 Organisation of the thesis.....	18
1.11 Summary.....	19
Chapter 2: Literature Review and Theoretical Framework.....	20
2.1 Introduction	20
2.2 Overview of Chapter 2.....	20
2.3 Literature sources	20
2.4 Introduction to the theoretical frameworks of this study	21
2.5 The cognitive theory of multimedia learning	23
2.6 Cognitive load theory.....	27
2.6.1 The human cognitive architecture	27
2.6.2 The construct 'cognitive load'	28
2.6.3 Practical application of cognitive load theory.....	31

2.7	Riding's cognitive style model	31
2.7.1	Wholist-Analytic Style	34
2.7.2	Verbal-Imagery Style	34
2.8	Summing up Part 1 of the Literature Review	35
2.9	Introduction to Part 2	36
2.10	Cognitive styles and multimedia learning	38
2.10.1	Contributions to the theory	39
2.10.2	Measurement of cognitive style.....	42
2.10.3	Cognitive style and learner characteristics	50
2.10.4	Cognitive style and behaviour	52
2.10.5	Cognitive style and instructional design practice.....	55
2.10.6	Cognitive style and achievement	56
2.10.7	In summary	59
2.11	Cognitive load and multimedia learning.....	61
2.11.1	Instructional design issues	62
2.11.2	Theory development and directions of research	72
2.11.3	Measurement of cognitive load	82
2.11.4	Learning from multiple representations	89
2.11.5	Animations, including animated pedagogical agents	97
2.11.6	In summary	105
2.12	Multimedia in health sciences education	107
2.12.1	The media effect	108
2.12.2	The multimedia effect	109
2.12.3	The Contiguity effect.....	112
2.12.4	Interaction effect	113
2.12.5	Perceptions and attitudes.....	113
2.12.6	Meta-analyses.....	115
2.12.7	Multimedia in Physiology Education	115
2.12.8	In summary	120
2.13	Conclusion following the literature view.....	120
	Chapter 3: Research Methodology and Design	123
3.1	Introduction	123
3.2	Overview of this chapter	123
3.3	Purpose of the study.....	124
3.4	The research questions	126
3.5	The research approach	127
3.6	The research design.....	128
3.6.1	The experimental design	129
3.6.2	Managing threats to validity.....	131
3.7	The research sample.....	133
3.8	The research data	134
3.8.1	Cognitive style data	135
3.8.2	Cognitive load data.....	135
3.8.3	Data for the presentation format	137
3.8.4	Learning performance data	137

3.9	The research instruments.....	137
3.9.1	Measuring cognitive style	139
3.9.2	Measuring cognitive load.....	139
3.9.3	Measuring learner performance	141
3.9.4	Integrating the research instruments into the intervention	142
3.10	Ethical considerations in this study	143
3.10.1	Obtaining consent to participate in the study.....	143
3.10.2	Informing participants about their cognitive style.....	144
3.10.3	Anonymity	144
3.10.4	Making all the formats available to the participants.....	145
3.11	Summary for Part 1	145
3.12	Introduction	146
3.13	Source of the content	146
3.14	Design and development of the multimedia	147
3.14.1	Design of the program	147
3.15	Instructional strategies and media used	149
3.16	The hypotheses and expected findings	152
3.16.1	Distribution of the styles across the sample	152
3.16.2	Cognitive load of the presentation formats	154
3.16.3	Rating of cognitive load according to style	156
3.16.4	The relationship between cognitive style, cognitive load and learning.....	157
3.17	Introduction	158
3.18	Conducting the pilot study.....	159
3.18.1	Pilot study 1.....	160
3.18.2	Pilot study 2.....	163
3.19	Conducting the main study.....	167
3.20	Summary.....	170
	Chapter 4: Presentation and Analysis of Empirical Data	172
4.1	Introduction	172
4.2	Statistical analysis	173
4.3	The profile of the participants	174
4.3.1	Age, gender and cultural group.....	175
4.3.2	Home language of the sample	177
4.3.3	Prior knowledge of the participants	178
4.3.4	In summary: profile of the participants	180
4.4	Time spent on each version of the intervention.....	180
4.4.1	Time spent on full program.....	180
4.4.2	Time spent on individual screens.....	183
4.4.3	In summary: time spent on the program	186
4.5	Exploring the role cognitive style plays in an authentic multimedia learning environment	186
4.5.1	Cognitive style as measured using Riding's CSA	187
4.5.2	Cognitive style and time spent on the program	194
4.5.3	Cognitive style and use of the multimedia program	203
4.6	Exploring the role cognitive load plays in an authentic multimedia learning environment.....	206
4.6.1	Self-report of cognitive load.....	207

4.6.2	The relationship between cognitive load and cognitive style	224
4.6.3	Subjective rating of cognitive load and other variables	239
4.6.4	Cognitive load and time spent on the program.....	241
4.7	The correlation between self-report of cognitive load and the direct measure of cognitive load	247
4.7.1	Comparing correlation of the measurement techniques by version	247
4.7.2	Comparing correlation of the measurement techniques without considering version.....	249
4.7.3	In summary	249
4.8	The interaction between cognitive style, cognitive load and learning performance in an authentic multimedia learning environment	250
4.8.1	Results of the pre- and posttest, independent of cognitive style and cognitive load..	250
4.8.2	Cognitive style and learning performance	266
4.8.3	The subjective rating of cognitive load and learning performance	268
4.8.4	Cognitive style, cognitive load and learning performance.....	269
4.9	Summary of Chapter 4	277
Chapter 5: Discussion and Recommendations		278
5.1	Introduction	278
5.2	Summary of the research	278
5.2.1	Purpose of the study.....	278
5.2.2	Background to the study.....	279
5.2.3	The literature review	279
5.2.4	Finding the research question.....	280
5.2.5	The rationale of the study	280
5.2.6	The research methodology.....	281
5.2.7	Conducting the study	284
5.2.8	Analysis of the results.....	285
5.3	Methodological reflection.....	286
5.3.1	The research approach and design	286
5.3.2	The research sample	289
5.3.3	The research instruments and data	290
5.3.4	The research intervention.....	291
5.4	Substantive reflection	292
5.4.1	The role of cognitive style in an authentic multimedia learning environment	292
5.4.2	The role of cognitive load in an authentic multimedia learning environment.....	300
5.4.3	The correlation between self-report and direct measurement as techniques in measuring cognitive load.....	306
5.4.4	Presentation formats and their influence on cognitive load	308
5.4.5	Cognitive load, cognitive style and learning performance.....	316
5.5	Scientific reflection.....	322
5.5.1	Contribution of this study to the body of knowledge.....	322
5.5.2	Implications of this study for instructional design	324
5.5.3	Implications of this study for using multimedia in the learning environment.....	325
5.5.4	Recommendations for future research	325
5.6	Conclusion	328
References		330

Appendix A: Summary of COGNITIVE Styles research reviewed by Riding and Cheema (1991) ...	357
Appendix B: Letter requesting permission to conduct the study and letter providing this consent ...	358
Appendix C: permission from Smith (2007) to use data from her study	360
Appendix D: Pre test / Posttest: Computer-based test	361
Appendix E: Final two questions of the Post test.....	367
Appendix F: Title screen and practice session screen	369
Appendix G: Multimedia intervention: Screens common to both formats	371
Appendix H: Multimedia intervention: Different strategies to display SAME content.....	382
Appendix I: multimedia: Animation versus static images.....	384
Appendix J: Multimedia intervention: Whole view versus Parts view.....	389
Appendix K: A summary of the Multimedia and the integration of the research instruments.....	392
Appendix L: Electronic questionnaire to collect demographic data.....	394
Appendix M: Allocation of groups for the main study.....	395
Appendix N: Example of data for a single participant - written out to an .INI file	396
Appendix O: Student handout - Participation in the study	407
Appendix P: Riding's Cognitive styles Analysis	412

List of Tables

		Page
Chapter 1	Overview and Orientation	
Table 1.1	Cognitive load - types and sources of load	3
Table 1.2	Contexts of cognitive load research	11
Table 1.3	Contexts of cognitive style research, using Riding's CSA and other measures of cognitive style	12
Table 1.4	Summary of the research instruments and nature of the data	16
Table 1.5	Summary of the methods used to analyse the data	17
Chapter 2	Literature Review and Theoretical Framework	
Table 2.1	Detailed outline of the parts of Chapter 2	20
Table 2.2	List of special edition education journals devoted to cognitive load and/or multimedia research	21
Table 2.3	Approach to the literature review	38
Table 2.4	Assumptions underlying different techniques for measuring cognitive load	83
Table 2.5	The media effect in health sciences multimedia education	108
Table 2.6	The multimedia effect in health sciences multimedia education	110
Table 2.7	The contiguity effect in health sciences multimedia education	112
Table 2.8	The interaction effect in health sciences multimedia education	113
Table 2.9	Perceptions of and attitudes towards multimedia education in the health sciences	113
Table 2.10	Meta-analyses of research into multimedia education in the health sciences	115
Table 2.11	Review of recent research in Physiology education	115
Table 2.12	The literature review and the research questions	122
Chapter 3	Research Methodology and Design	
Table 3.1	Detailed outline of Chapter 3	124
Table 3.2	Managing threats to validity	131
Table 3.3	Operational definitions of the variables	135
Table 3.4	Summary of research instruments used in this study	138
Table 3.5	Number of times cognitive load was measured in each version	143
Table 3.6	A summary of the major design similarities and differences between the programs	148
Table 3.7	Scope of the content of the multimedia program	150
Table 3.8	Profile of the research sample	159
Table 3.9	Placement of the secondary task across four programs	165

	Page
Chapter 4	Presentation and Analysis of Empirical Data
Table 4.1	Distribution of sample across research intervention, by age and gender 176
Table 4.2	Number of participants who had studied topic previously 178
Table 4.3	Time groupings for each version of the program 181
Table 4.4	Mean time spent on program for each version 182
Table 4.5	Frequency of access for screen 12 in the animation version and screen 13-16 in the static images & text version 185
Table 4.6	Profile of the sample for the WA dimension of the CSA using three style groups 188
Table 4.7	WA style ratios according to gender 189
Table 4.8	WA style ratios according to culture 189
Table 4.9	Stepwise regression equation for WA style 190
Table 4.10	Profile of the sample for the Verbaliser-Imager Dimension of the CSA using three style groups 191
Table 4.11	VI style ratios according to gender 192
Table 4.12	VI style ratios according to culture 193
Table 4.13	Stepwise regression equation for VI style 193
Table 4.14	Frequency of WA and VI styles for each version of the program 195
Table 4.15	Comparison of time spent on program for WA style and version 195
Table 4.16	Comparison of time spent on selected screens for WA style 196
Table 4.17	Comparison of time spent on the animation version for the WA style dimension 197
Table 4.18	Comparison of time spent on the static images & text version for the WA style dimension 198
Table 4.19	Comparison of time spent on program for VI style and version 199
Table 4.20	Comparison of time spent on selected screens for VI style 200
Table 4.21	Comparison of time spent on the animation version for the VI style dimension 201
Table 4.22	Comparison of time spent on the static images & text version for the VI style dimension 201
Table 4.23	Cognitive style of participants who accessed the program once only 204
Table 4.24	Results of Chi-square analyses to determine relationship between style and access in the animation version 205
Table 4.25	Results of Chi-square analyses to determine relationship between style and access in the static images & text version 206
Table 4.26	Overview of content for the animation version on which self-report of cognitive load was based 207
Table 4.27	Frequencies of cognitive load reported as low, medium and high for the animation version 208

	Page	
Table 4.28	Mean cognitive load ratings for individual measurements of cognitive load in the animated version	209
Table 4.29	A comparison of the cognitive load for selected screen pairs in the animation version	210
Table 4.30	Effect sizes for comparisons between screen in the animation version	211
Table 4.31	Overview of content in the static image & text version on which self-report of cognitive load was based	212
Table 4.32	Frequencies of cognitive load reported as low, medium and high for the static images & text version	213
Table 4.33	Mean cognitive load ratings for individual measurements of cognitive load for the static images & text version	214
Table 4.34	A comparison of the cognitive load for selected screen pairs in the static text & images version	215
Table 4.35	Effect sizes for comparisons between screen in the static images & text version	216
Table 4.36	Cognitive load using direct measurement technique	217
Table 4.37	Results of GLM Repeated Measures Analysis for the cognitive load comparisons	217
Table 4.38	Effect sizes of comparison of mean cognitive load measured with the direct method	218
Table 4.39	Mean cognitive load per version using the subjective rating technique	220
Table 4.40	Screen-wise comparison of cognitive load for selected presentation formats	221
Table 4.41	Cognitive load of selected screens using the direct method of measurement	223
Table 4.42	Cognitive load for WA style dimension	225
Table 4.43	Cognitive load and the WA style dimension by version	225
Table 4.44	Cognitive load of the WA styles for each version	226
Table 4.45	Cognitive load levels and the WA style dimension	227
Table 4.46	Cognitive load for the Analytic learner grouped by time spent on program	228
Table 4.47	T-test for significance for the comparisons of cognitive load per time group for the Analytic learner	229
Table 4.48	Cognitive load and VI style	230
Table 4.49	Cognitive load and the VI style dimension by version	230
Table 4.50	Cognitive load of the VI styles for each version	231
Table 4.51	Cognitive load levels and the VI style dimension	232
Table 4.52	Cognitive load for the VI style grouped by time spent on the animation version	232
Table 4.53	T-test for significance for the comparisons of cognitive load per time group for the Verbaliser learner	234

	Page	
Table 4.54	Cognitive load for the VI style grouped by time spent on the static images & text version	235
Table 4.55	T-test for significance for the comparisons of cognitive load per time group for the Imager learner	236
Table 4.56	Stepwise regression equation for self-report rating of cognitive load	240
Table 4.57	GLM analysis results for cognitive load and rating of knowledge about the topic	241
Table 4.58	GLM analysis results for cognitive load and home language	241
Table 4.59	Cognitive load for time spent on program by version	242
Table 4.60	Cognitive of screen 12 where there were multiple entries to the screen	243
Table 4.61	Cognitive of screen 19 where there were multiple entries to the screen	243
Table 4.62	Cognitive load of screen 23 where there were multiple entries to the screen	244
Table 4.63	Cognitive load of screens 13, 14, 15 and 16 where there were multiple entries to each of these screens	245
Table 4.64	Mean cognitive load using direct measurement (Smith, 2007)	248
Table 4.65	Means for the cognitive load, using different measurement techniques	248
Table 4.66	Pretest questions and answers and the number of participants who answered correctly and incorrectly	253
Table 4.67	Learning performance and self-report of previous exposure to topic	255
Table 4.68	Learning performance and self-report of level of knowledge of topic	256
Table 4.69	Pretest results and self-rating of prior knowledge and understanding of topic	257
Table 4.70	Pretest by version	258
Table 4.71	Posttest questions and answers and the number of participants who answered correctly and incorrectly	261
Table 4.72	Posttest by version	263
Table 4.73	Mean score for Question 1 of the open-ended questions for each version of the program	264
Table 4.74	Mean score for Question 2 of the open-ended questions for each version of the program	264
Table 4.75	Mean scores for the pre- and posttests by version	266
Table 4.76	GLM analysis results for WA style and learning performance	267
Table 4.77	GLM analysis results for VI style and learning performance	267
Table 4.78	GLM analysis results for version and performance	268
Table 4.79	Posttest results for the Analytic learner, by time and load	271
Table 4.80	Posttest results for the animation version learner for the VI style by time and load	273
Table 4.81	Posttest results for the static images & text version learner for the VI style by time and load	275
Table 4.82	Stepwise regression equation for posttest results	276

		Page
Chapter 5	Discussion and Recommendations	
Table 5.1	Overview of Chapter 5	278
Table 5.2	A summary of the major design similarities and differences for the programs	284
Table 5.3	Profile of the research sample	285
Table 5.4	Reflecting on the research design	288
Table 5.5	Alignment of the instruments used in the study with criteria for good instruments	291

List of Figures

		Page
Chapter 2	Literature Review and Theoretical Framework	
Figure 2.1	A framework for a cognitive theory of multimedia learning (Mayer, 2003:129)	26
Figure 2.2	The construct cognitive load	28
Figure 2.3	Riding's Cognitive Style Model	33
Figure 2.4	Relationships between cognitive load, cognitive style and multimedia learning	37
Figure 2.5	New relationships between cognitive style, cognitive load and multimedia learning	37
Figure 2.6	Progression for the 'effects' research within the cognitive load research stream	63
Figure 2.7	Cognitive load research timeline	76
Figure 2.8	Assumptions of cognitive load (Gerjets & Scheiter, 2003, page 35)	78
Figure 2.9	Goals and strategies as moderators between instructional design and cognitive load (Gerjets & Scheiter, 2003, page 36)	79
Figure 2.10	Relationships of physiology with related Medical Science (Sefton, 1998, pg. 54)	115
Chapter 3	Research Methodology and Design	
Figure 3.1	A visual representation of the research methodology	123
Figure 3.2	Graphic representation of the issues that define this research design	129
Figure 3.3	Screen format of self-reporting a rating for mental effort	140
Figure 3.4	Menu for multimedia intervention used in the second pilot study	142
Figure 3.5	A visual representation of the structure of the multimedia program	148
Figure 3.6	Display of secondary task in the program used in pilot study 1	162
Figure 3.7	How remaining time was displayed on screen in the pretest	164
Figure 3.8	The message that was displayed in pre- and posttest once time had elapsed	164
Figure 3.9	Placement of the secondary task in Programs 1 and 3	165
Figure 3.10	Placement of the secondary task in Programs 2 and 4	165
Chapter 4	Presentation and Analysis of Empirical Data	
Figure 4.1	Gender profile of the sample	175
Figure 4.2	Profile of sample by gender and age	175
Figure 4.3	Profile of sample by cultural group	177
Figure 4.4	Percentage of participants where English is their home language	178

		Page
Figure 4.5	Rating of level of knowledge and understanding	179
Figure 4.6	Access to screen 16 in the static images & text version	184
Figure 4.7	Example of question in pretest	251
Figure 4.8	Distribution of pretest scores	251
Figure 4.9	Distribution of posttest scores	259
Figure 4.10	Distribution of pre- and posttest scores: Low, Average and High Performance	259
Figure 4.11	Comparison of pre- and posttest scores for each score in the range	265
Chapter 5	Discussion and Recommendations	
Figure 5.1	Screen 5 – First image	310
Figure 5.2	Screen 5 – Second and third image	311
Figure 5.3	Screen 5 – Fourth image	311
Figure 5.4	Section of screen 13 in Version 2	313
Figure 5.5	Section of screen 20 in Version 2	314
Figure 5.6	View from screen 18 in the static images & text version	315
Figure 5.7	Perspectives of the study	328

Abbreviations used in the study

ANS	Autonomic Nervous System
AV	Analytic–Verbaliser (unitary style)
AI	Analytic–Imager (complementary style)
CAI	Computer-Assisted Instruction
CAL	Computer-Assisted Learning
CBT	Computer-based instruction
CLT	Cognitive Load Theory
CSA	Cognitive Styles Analysis
CSI	Cognitive Styles Index
Exp	Experiment
Extended CSA-WA	Extended Cognitive Style Analysis–Wholistic–Analytic test
FD	Field-Dependence
FI	Field-Independence
GEFT	Group Embedded Figures Test of Witkin
GCSE	General Certificate of Secondary Education (UK)
IT	Information Technology
fMRI	Functional Magnetic Resonance Imaging
PBL	Problem-based learning
PNI	Psychoneuroimmunology
PNS	Parasympathetic Nervous System
SBLSQ	Santa Barbara Learning Style Questionnaire
SNS	Sympathetic Nervous System
SRCL	Self-report cognitive load
SRCLV	Total self-report cognitive load of the lesson for each participant
UP	University of Pretoria
Vers	Version
VI	Verbaliser-Imager (style dimension)
VICS	Verbal Imagery Cognitive Style
VVLSR	Verbal-Visual Learning Style Rating
WA	Wholistic-Analytic (style dimension)
WM	Working memory
WV	Wholistic–Verbaliser (complementary style)
WI	Wholistic–Imager (unitary style)
WA-IT	Wholistic-Analytic Inspection Time test

List of Appendices

	Page	
Appendix A	Summary of Cognitive Styles research reviewed by Riding and Cheema (1991)	357
Appendix B	Letter requesting permission to conduct the study and letter providing this consent	358
Appendix C	Permission from Smith (2007) to use data from her study	360
Appendix D	Pre test / Posttest: Computer-based test	361
Appendix E	Final two questions of the Post test	367
Appendix F	Title screen and practice session screen	369
Appendix G	Multimedia intervention: Screens common to both formats	371
Appendix H	Multimedia intervention: Different strategies to display content	382
Appendix I	Multimedia intervention: Animation versus static images	384
Appendix J	Multimedia intervention: Whole view versus Parts view	389
Appendix K	A summary of the multimedia and the integration of the research instruments	392
Appendix L	Electronic questionnaire to collect demographic data	394
Appendix M	Allocation of groups for the main study	395
Appendix N	Example of data for a single participant - written out to an .INI file	396
Appendix O	Student handout - Participation in the study	407
Appendix P	Riding's Cognitive styles ANALYSIS	412

Overview and Orientation

1

1.1 Introduction

This thesis documents an empirical study that used quantitative methods to investigate the relationship between cognitive load and cognitive style when using animation and static images in instructional material.

This chapter presents:

- the rationale and purpose of the study.
- an orientation to the design and implementation of the study.

1.2 Purpose of the study

The purpose of this study was to explore the role that cognitive load and cognitive style play in the successful achievement of learning outcomes when narrated animation and static images are used in multimedia learning formats in an authentic learning environment. The study also investigated the relationship between cognitive load, which is influenced by both the nature of the content and the specific design strategies used, and the cognitive style of the individual who uses different multimedia formats.

Schnotz and Lowe (2003) are of the opinion that new technologies in general, and multimedia in particular, play an increasingly important role in education. They describe multimedia as:

...the combination of multiple technical resources for the purpose of presenting information represented in multiple formats via multiple sensory modalities.

(Schnotz & Lowe, 2003, p 117).

Multimedia resources can be considered at three levels: a technical, semiotic and sensory level. The technical level focuses on the effect of the medium per se on learning, while the semiotic level refers to the effect of the representational format of the content, be this text, pictures or sound. Early research often focused solely on the media-effects (technical level) (Erwin & Ricardo, 1999; Mayer, 1997a; Quealy & Langan-Fox, 1998; Williams, Aubin, Harkin & Cottrell, 2001). The current view is that it is misguided and overly simplistic to compare different technical media with regard to their effects on learning without taking into consideration the semiotic and sensory effects (Mayer, 1997a; Mayer & Moreno, 2002; Schnotz & Lowe, 2003).

Cognitive science and educational psychology literature calls for the continued investigation of the conditions under which different forms of representation of specific media, such as texts and graphics (either static or animated), influence comprehension and learning (Goldman, 2003; Mayer & Moreno, 2002). The quest for finding evidence for optimal presentation formats and factors that need to be considered when designing multimedia instruction has not lost momentum in the new millennium, as evident from the number of special issues devoted to cognitive load and multimedia learning since 2002 (Kirschner, 2002; Paas, Renkl & Sweller, 2004; Paas & van Gog, 2006; Ploetzner & Lowe, 2004; Robinson, 2002; Schnotz & Lowe, 2003; van Merriënboer & Ayres, 2005).

There is considerable empirical evidence, to be discussed in Chapter 2 of this thesis, that cognitive load and cognitive style are two factors that impact on learning effectiveness and efficiency. Exploring cognitive load and cognitive style, and the **relationship between these two factors**, as an instructional designer becomes meaningful and worthwhile if the empirical evidence or outcome of such exploration can build on the available design principles and heuristics for designing effective instruction.

1.3 Defining core concepts and terminology

This section provides a brief explanation of the core concepts and terminology in order to facilitate the reading of this thesis. Chapter 2 will discuss these concepts and their position within the theoretical frameworks used in this study in more detail.

Four concepts are described:

- Cognitive load
- Cognitive style
- Learning style
- Multimedia learning

1.3.1 Cognitive load

Understanding the concept cognitive load requires taking a step back to consider the process of acquiring and understanding information. An understanding of cognitive load is based on several important assumptions:

- Working memory has limited capacity for information processing.
- In contrast, long-term memory has unlimited capacity for storing information.
- Information is also stored in long-term memory for an unlimited duration.
- Information is stored in knowledge structures called schemas.

- A schema categorises elements of information according to the manner in which they will be used. It is a hierarchically organised, domain-specific structure.
- There are many different schemas in long-term memory and they vary in complexity.
- The effectiveness of retrieving information from long-term memory depends on the quality of schema construction and automation.
- Controlled and conscious use of schemas also uses up working memory capacity (Kalyuga, 2006; Paas, Renkl & Sweller, 2003; Sweller, van Merriënboer & Paas, 1998).

With this as background, one perspective is to view information, which enters working memory via the senses for the first time or as a schema from long-term memory, as a burden. This is a negative perspective. Instead, cognitive scientists have conceptualised this as the cognitive load.

The literature is elusive when it comes to providing a formal definition of cognitive load. Explanations of cognitive load are provided within the framework of cognitive load theory (CLT) (Chandler & Sweller, 1991; Sweller et al., 1998; Sweller & Chandler, 1991). The only concise definition of the broader concept of cognitive load is that of Paas and van Merriënboer (1994a):

Cognitive load is generally considered a multidimensional construct that represents the load that performing a particular task imposes on the cognitive system of the learner

(Paas & van Merriënboer, 1994a, p353).

In-depth descriptions and explanations of cognitive load theory in the literature consider the concept of cognitive load in terms of three types of cognitive load. Each type of load has a specific source. This conceptualisation is summarised here in Table 1.1.

	Types of cognitive load	What determines this load
Cognitive load	Intrinsic load	The actual material/information to be learned.
	Extraneous load	The way in which the material/information is presented
	Germane load	Processes that contribute to the construction and automation of schemas.

Table 1.1: Cognitive load - types and sources of load (Kirschner, 2002; Paas, Renkl & Sweller, 2003).

This study measures the cognitive load of two presentation formats (animation and static images) used to teach Physiology to adult learners in an authentic learning environment.

1.3.2 An introduction to style

The literature on styles is vast. A search of the Academic Premier Database returned over 600 references for learning style and approximately 450 references for cognitive style.

It is a very controversial area of research that has been criticised in many areas, including the:

- plethora of style models found in the literature (Riding & Cheema, 1991);
- confusion between cognitive and learning style and the inter-changeable use of these two concepts (Ford & Chen, 2001);
- difficulty in measuring styles;
- low reliability and validity of the instruments used to assess styles (Curry, 1990; Peterson, Deary & Austin, 2003a, 2003b).

In spite of the criticism the research has continued. In fact there has been a renewed interest in both cognitive and learning styles research in recent years. Riding and Cheema (1991) note a resurgence of interest in this field after a decline in the prominent interest of the 1960s and 1970s. A review of the more recent literature does seem to provide ground for this opinion (Calcaterra, Antonietti & Underwood, 2005; Chen, Ghinea & Macredie, 2006; Ford & Chen, 2001; Guinea & Chen, 2003; Graff, 2005; Smith & Woody, 2000; Triantafillou, Pomportsis, Demetriadis & Georgiadou, 2004).

Cognitive and learning styles have emerged as a key dimension of individual differences (Ford & Chen, 2001). But what is the difference between these two constructs, if any at all? There are a great variety of so-called 'cognitive styles' and 'learning styles' (Riding & Cheema, 1991). Some authors use the terms interchangeably, while others consider them to be different concepts. Ford and Chen (2001) interpret learning styles as 'cognitive styles entailing information processing taking place specifically in a learning context.'

When can patterns of behaviour be called a style? What is the difference between a style and a strategy? When an individual displays a consistent tendency to behave in a certain manner, or, in an educational context, consistently uses a selected combination of strategies, this tendency is called a style. Witkin, Moore, Goodenough and Cox (1977) view style as the:

|| *'....characteristic approach the person brings with him to a wide range of situations....'*

They go on to state that when this approach involves both the person's perceptual and intellectual activities it is called their cognitive style (Witkin et al., 1977). Style probably has a physiological basis and is fairly fixed for an individual. Strategies however, are ways that may be learned and developed to cope with a variety of situations and tasks, including learning tasks (Riding & Rayner, 1998).

The position taken in this study is that cognitive style and learning style are two different but related concepts.

1.3.3 Cognitive style

Cognitive style is seen as an individual's preferred and habitual approach to both organising and representing information

(Riding & Rayner, 1998, p8).

This study uses a model that conceptualises style on two bi-polar dimensions: the Analytic-Wholistic and the Verbaliser-Imager dimensions.

An extreme analyst, who picks up a textbook to study, will look at the Index and then read systematically through the text. The extreme Wholist will look at the Table of Contents and the Executive Summary (if it exists). He or she might not look at anything else, but rather go and have a cup of coffee with their buddy. The extreme Verbaliser will read all the text and glance over the diagrams, while the extreme Imager will primarily look at the images.

This study measures the cognitive styles of the participants before they learn content that is presented using either animation or static images.

1.3.4 Learning style

Running parallel to the body of research on cognitive style is another research stream that seeks to better understand individual differences between students. The focus has been on an individual's active response to a learning task as opposed to the more internal processes inherent to cognitive style. Learning style therefore addresses the way a learner approaches a learning task. The results of learning style research are focused on ways to design better learning environments that would more effectively meet individual learning needs (Howard, Ellis & Rasmussen, 2004; Laight, 2004; Sabry & Baldwin, 2003; Sonnenwald & Li, 2003; Wieseman & Portis, 1990).

Riding and Rayner (1998) surveyed the literature on learning style theory and categorised the various models into four style groups. These style models are based on the learning process, orientation to study, instructional preferences and cognitive skills development. These will not be discussed further in this thesis.

1.3.5 Multimedia learning

The most simplistic understanding of the concept 'multimedia' would be that of many different forms (multi) of media – text, sound, video, static images and animation.

The Oxford English Dictionary Online (2006) defines multimedia as:

'The use of a variety of artistic or communicative media; (Computing) the incorporation of a number of media, such as text, audio, video, and animation, esp. interactively.'

Wikipedia, an online encyclopedia (2006), defines multimedia as:

'...the use of several media (e.g. text, audio, graphics, animation, video) to convey information. Multimedia also refers to the use of computer technology to create, store, and experience multimedia content.'

The definition of multimedia by Schnotz and Lowe (2003), quoted on page 1 of this thesis, emphasises the multiplicity of the field: multiple technical resources, multiple formats and multiple sensory modalities. The possibilities for including media in instruction today are vast. Design decisions address the question of which media or media combinations to use. Instructional designers must be critical of the practice that allows technology to generate the learning experience rather than using the growing knowledge of cognitive processes to guide the decisions regarding effective technology utilisation. Any approach to instruction that ignores cognitive processes is likely to be deficient (Chandler, 2004).

The field of multimedia learning therefore not only considers the media to be used, but also the impact of this media on the cognitive processes. There is a very close link between the theory of multimedia learning (Mayer, 2003) and cognitive load theory. The theory of multimedia learning will be considered in more detail in Chapter 2.

The multimedia formats used in this study are animation and static images. Both formats will be used to teach the same Physiology content to young adult learners.

1.4 Background to the study

As a multimedia instructional designer working in the context of tertiary health science education, I often reflected on the following questions:

1. Is multimedia development worth the effort in terms of the time and resources (financial and human) that need to be allocated to such development?
2. To what extent do multimedia learning resources contribute to the achievement of learning outcomes?
3. How do learners use and learn with these resources and learning materials?
4. How can the instructional design be improved in order to facilitate learning and meet the learning needs of diverse learners?

These questions, and the fact that as an instructional designer I have both overseen the design of the learning materials and frequently designed them myself, provided the direction for a review of the literature. The first area covered in the literature was that of multimedia learning. It was then that I started looking more closely at both the theoretical and empirical literature on cognitive load. The second area of review initially considered both learning and cognitive styles, but I eventually narrowed my focus to the cognitive style literature as I view these as two related but different concepts.

Questions of a similar nature are often asked by clients in corporate settings. These questions are being asked increasingly by management who must allocate resources from an often shrinking resource base.

Answering the question 'To what extent do multimedia learning resources contribute to the achievement of learning outcomes?' will provide part of the answer to the first question I asked: namely, 'Is multimedia development worth the effort in terms of time and resources (financial and human) that need to be allocated to such development?' The question implies some sort of measurement of these learning outcomes. There are however many factors that influence the successful achievement of learning outcomes, including the design of the learning materials, the learning environment, the learning strategies used by the learner, motivation and many more. It is also no great secret that we are all different, but to what extent do these differences apply to learning? Answers to this question and the question 'How can the instructional design be improved in order to facilitate learning and meet the learning needs of diverse learners?' seem to point in the direction of individual differences of learners.

The question 'How do learners use and learn with these [multimedia] resources and learning materials?' is one that instructional designers do not consider often enough. The common practice, particularly in the environment where there is a team approach to development, is for the designer to hand the final product over to the client or lecturer who is responsible for integrating it into the teaching and learning environment.

There is still much that needs to be discovered about how learners, in particular those studying at a distance, use and learn with electronic multimedia learning materials. This question indicated the need to explore the relationship between multimedia and learning behaviour in general. Exploring learning behaviours will assist in the identification of those behaviours and instructional conditions that enable and facilitate rather than hinder learning. When learning behaviour, or a particular instructional condition (animation or static images), increases the range of possible cognitive processes and therefore allows more cognitive processing to take place learning is enabled. When learning material has been designed to make processing and schema acquisition easier, or the interaction with the program is made easier (easier navigation for example) then learning is facilitated (Schnotz & Rasch, 2005). Designers must consider both the enabling and facilitative function of multimedia learning resources.

There are many studies in the literature describing the attitudes and perceptions of learners towards a wide array of technology-based learning materials and environments (Collaud, Gurtner & Coen, 2000; Ghinea & Chen, 2003; Jha, Widdowson & Duffy, 2002; Kerfoot, Masser & Hafler, 2003; Regnard, 2000). Attitudes do not always translate into action. I feel that it is important to consider what learners DO with the electronic learning materials, rather than looking only at how they FEEL about these materials? How does the design influence their use and the learning processes? Since these electronic learning materials must often stand alone in their ability to teach, does the design enable, facilitate or hinder the learning process? What learner support is needed to ensure that they use these learning materials optimally? How must learners be helped to make the transition from using paper-based to electronic learning materials? How should these electronic materials be designed to meet different learning needs and styles? Which designs will achieve the best learning outcome – for the majority of learners and for learners with specific learning styles?

1.5 Rationale of the research

The constructs cognitive load and cognitive style are both widely researched in their own right. The research spans close on twenty years (Paas & van Gog, 2006; Peterson & Deary, 2006; Sweller & Chandler, 1991; Thornell, 1976). New uses of educational technology in teaching and learning provide the rationale for re-visiting old research questions (Mayer, 1997a). The results from empirical studies serve to guide the practice of instructional designers, teachers and facilitators (Leahy, Chandler & Sweller, 2003; Mayer, 2003; Mayer & Moreno, 2003; Riding & Rayner, 1998). Chapter 2 will provide a critical review of this research.

This section builds the immediate case for the research question. The case is built by exploring the theoretical, empirical, methodological, media and contextual dimensions in cognitive load, cognitive style, learning style and multimedia learning literature.

1.5.1 The empirical imperative

Recent empirical research in the field of cognitive style and multimedia either fails to address the fact that the outcomes of the research might be due to cognitive overload (Ghinea & Chen, 2003) or only hints that there might be a relationship between cognitive style and cognitive load (Graff, 2003b). This is an avenue of research that does not seem to have been explored yet in great detail.

The results of Riding, Grimley, Dahraei and Banner's study (2003) seem to indicate that effective working memory capacity has a major influence on the performance of learners with specific styles. These researchers call for more investigation into this finding.

This study will look at working memory from a cognitive load perspective.

1.5.2 The theoretical imperative

As early as 1994 the literature on cognitive load theory stated that there are three factors that contribute to cognitive load: task characteristics, learner characteristics and the interactions between these two (Paas & van Merriënboer, 1994a). One of the learner characteristics listed included cognitive style. Other learner characteristics include cognitive capabilities, preferences and prior knowledge.

Cognitive load research has explored the influence of prior knowledge and learner experience on cognitive load in considerable depth (Kalyuga, 2006; Kalyuga, Ayres, Chandler & Sweller, 2003; Kalyuga, Chandler & Sweller, 2001), but the field has been strangely silent on the influence of cognitive style.

This study will investigate the theoretical link between cognitive style and cognitive load.

1.5.3 The methodological imperative

Methodological limitations in previous research include:

- studies that did not measure achievement directly (Riding, Grimley, Dahraei & Banner, 2003).
- small samples of under 100 participants (Ayres, 2006a; Dutke & Rinck, 2006; Mayer, Sobko & Mautone, 2003; Riding & Grimley, 1999; van der Meij, & de Jong, 2006).
- cautions that not all findings are easily generalisable to the classroom setting (Tabbers, Martens & van Merriënboer, 2004) since the context of the study was the laboratory setting.
- failure to measure the cognitive load of the intervention (Chandler & Sweller, 1991; Ghinea & Chen, 2003; Mayer, Moreno, Boire & Vagge, 1999; Moreno, 2006).
- giving the participants material to learn that is generally not relevant to their own coursework (Mayer, Fennell, Farmer & Campbell, 2004; Moreno, 2004).

This study will measure achievement, use a larger sample than many of the studies reviewed in the literature and will take place in an authentic learning environment.

1.5.4 The media imperative

Why consider animation and static images above other media options? The use of static images and text in instructional resources has received considerable attention in the research community since the early 1980s (Carney & Levin, 2002; Mayer, 2003; Mayer & Gallini, 1990; Mayer, Mautone, & Prothero, 2002; McKay, 1999; Moreno & Valdez, 2005; Verdi & Kulhavy, 2002). Graff (2003a) acknowledged in his research that incorporating other multimedia components in the instruction could lead to different findings. More recent research considered the use of text, images and sound in

varying combinations (Ginns, 2005; Jeung & Chandler, 1997; Kalyuga, Chandler & Sweller, 1999; Leahy, Chandler & Sweller, 2003; Mikk & Luik, 2003).

Improvement in technology has seen the increased use of sound, video, animation and 3D presentation formats in instructional materials. It is time to research the impact of these newer media formats on cognitive processes with the same rigour and vigour that have been applied to researching the use of text and images.

Chandler noted the following:

...despite this seemingly endless potential and unbridled enthusiasm for technology-based instruction, there is little empirical evidence to indicate that the widespread use of dynamic visualisations has resulted in any substantial benefit to learners.

(Chandler, 2004, p 353).

Of even more concern is the reflection by Kalyuga, Chandler and Sweller (1999, p369):

Many multimedia instructional presentations are still based on common sense rather than theory or extensive empirical research. Visual formats tend to be determined purely by aesthetic considerations while the use of sound and its interaction with vision seems not to be based on any discernible principles.

This study is an empirical investigation of the use of animation and static images in health science education.

1.5.5 The contextual imperative

The content domain of a large majority of the studies includes mathematics, science, technical subjects such as electrical circuits, computer applications or statistics. These are conducted primarily within primary, secondary and vocational education contexts. Table 1.2 and 1.3 summarise a selection of these studies.

There have been calls in the cognitive load and cognitive style research literature to consider replication studies using other contexts and subjects fields. The context of Graff's study (2003a) was psychological ethics. He indicated that it is possible that an instructional system with information on a different subject may yield a different finding. Riding et al. (2003) call for further research that includes designing a similar study for another context.

Content	Primary School	Secondary School	Vocational education	Higher Education
Statistics		Paas (1992)		Bodemer, Ploetzner, Feuerlein & Spada, (2004) - Exp 2
Mathematics and Science	Ginns, Chandler & Sweller (2003) – Exp 2 & 3 Jeung & Chandler (1997) Leahy, Chandler & Sweller (2003) Marcus, Cooper & Sweller (1996)	Ayres (2006a) Ayres (2006b) Clarke, Ayres & Chandler (2005) Kalyuga (2006) Kalyuga & Sweller (2005) Kalyuga & Sweller (2004) Kester, Kirschner & van Merriënboer (2004) Sinclair, Renshaw & Taylor (2004)		Bodemer, Ploetzner, Feuerlein & Spada, (2004) - Exp 1 Gerjets, Scheiter & Catrambone (2006) Große & Renkl (2006) Mayer et al., (1999) Reisslein, Atkinson, Seeling & Reisslein (2006) Schnotz & Rasch (2005)
Technical subjects			Kalyuga, Chandler & Sweller (2001) Kalyuga, Chandler & Sweller (2000) Kalyuga, Chandler & Sweller (1999) Pollock, Chandler & Sweller (2002) Van Gog, Paas & van Merriënboer (2006)	
Computer applications			Chandler & Sweller (1996)	Ginns, Chandler & Sweller (2003) – Exp 1 Catrambone & Yuasa (2006)
Instructional Design				Tabbers, Martens & van Merriënboer (2004)

Table 1.2: Contexts of cognitive load research

Content	Primary School	Secondary School	Higher Education
Natural Sciences	Riding & Grimley (1999)(Gravity & Motion, Reptiles)		
Mathematics / Accounting			Riding & Staley (1998) (Accounting)
Social Sciences	Griffin & Griffin (1996) (Map reading) Riding & Grimley (1999) (Geography)		
Psychology			Rush & Moore (1991)
Computer applications / Information Technology (IT)/ Computer literacy			Cunningham-Atkins, Powell, Moore, Hobbs & Sharpe (2004). Graff (2003b) Riding & Staley (1998) (IT) Triantafillou, Pomportsis, Demetriadis & Georgiadou (2004) Workman (2004) (Programming)
Various subjects - excluding health science subjects	Douglas & Riding (1993) (Prose)	Riding & Agrell (1997) Riding & Caine (1993) Riding, Grimley, Dahraei & Banner (2003)	Calcaterra, Antonietti & Underwood (2005) (History) Graff (2005) (History) Riding & Staley (1998) (Management)
Nursing			Luk (1998)

Table 1.3: Contexts of cognitive style research, using Riding's Cognitive Styles Analysis (CSA) and other measures of cognitive style

Cognitive and educational psychology research in the health science education domain seems to lag behind other science-related disciplines. A literature search using the search terms (health science education OR medical education OR paramedical sciences) AND (cognitive load OR working memory) AND (cognitive style) did not find any relevant studies. There have been studies that explored learning styles in health sciences education (Laight, 2004; Martin, Stark & Jolly, 2000; McNeal & Dwyer, 1999; Ross & Schulz, 1999) but a clear distinction has been made in this chapter between cognitive style and learning style.

The context of this study is health science education. What makes health science a suitable context, beyond the fact that I worked in this field?

Medical curricula across the globe are undergoing major changes (Rees, 2000). Health science education in recent years has moved from a subject-based approach to curriculum development to a more systemic, problem-oriented, integrated approach (Treadwell, de Witt & Grobler, 2002). For example, when considering cardio-vascular disease the students will study the relevant Anatomy, Physiology, Pathology, Pharmacology, diseases processes and the appropriate medical and other interventions in an integrated fashion. In order to make clinical diagnoses and prescribe the correct medication health practitioners must understand and apply their knowledge of physiology, pathophysiology and pharmacology. Element interactivity is high. The cognitive skills that must be acquired, such as decision-making and problem-solving, are complex and of the higher-order variety. Medical education is also using a variety of technologies to support learning. Cognitive load research has not extended to tertiary health science education in a meaningful way. It is therefore appropriate and necessary to encourage research in this field and in this context.

1.5.6 Other considerations

Technology has enabled choices that did not exist ten years ago. Learners like to be able to choose. Cognitive style influences this choice. Will the learner's choice result in cognitive overload? How will we design the correct combination of media so that it aligns with individual choice and need, especially if there is a relationship between cognitive load and cognitive style?

1.5.7 Pulling it together

Having considered these imperatives the case for this study can be summarised as follows:

- The empirical, theoretical and media imperatives suggest that there may be a relationship between cognitive load and cognitive style.
- The existence and nature of such a relationship has not been explored in any depth.
- Issues in both cognitive style and cognitive load are under-researched in health science education.

- There are limitations in current research designs that this study will try to overcome.
- There is a lack of research in authentic learning environments.
- Instructional design can benefit from more and specific guidelines for design, especially when new technologies are introduced.

1.6 Research questions

There is one major research question with five sub-questions in this study:

Question: What is the relationship between cognitive style and cognitive load as factors in the achievement of learning outcomes when someone learns the same content by means of different multimedia formats?

- i. What were the cognitive styles of the participants who took part in the study?
- ii. How did the participants rate the cognitive load of selected multimedia content?
- iii. What was the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?
- iv. To what extent did the presentation formats influence cognitive load?
- v. How was learning performance influenced when content with different cognitive load was studied by learners with different cognitive styles?

This study did not measure the cognitive load using the direct measurement technique. The results of the direct measurement of cognitive load, needed to answer sub-question iii, were obtained from Smith (2007) who was a research assistant for this study. She used the same instructional materials and sample for her study.

1.7 Research design and methodology

This quantitative study used an experimental and correlation design to determine the answers to the research questions. In keeping with typical approaches in experimental research a pilot study was undertaken before commencing with the main study.

Both the pilot and the main study were conducted in an authentic learning environment. The multimedia intervention used for the study is part of the prescribed course work in Physiology for health science learners at second year level. The authentic learning environment for participants in the study included using a computer laboratory on campus in order to learn and interact with the prescribed learning material.

It was therefore possible to control several of the variables that are normally difficult to control in authentic learning environments. Some of these extraneous factors include:

- time on task,
- lack of ability to focus on the task at hand due to many other distractions,
- collaboration with peers instead of working individually on the task,
- providing a uniform hardware platform on which the multimedia was delivered,
- access to other resources such as textbooks, journals and other notes, and
- control during the testing phases of the study.

Data collection was electronic in nature. The participants **received** the relevant questionnaires electronically and were required to **complete** these electronically. Data that tracked the respondent's use of the program and response to various stimuli was recorded automatically as the respondent worked through the multimedia intervention.

Table 1.4 lists the research sub-questions and indicates the instrument(s) used to collect the data. The measurement scales used in this study are also presented in this table. The research design and methodology are described in detail in Chapter 3 of this thesis.

Research question		Research instruments				Measurement scale and nature of data
		Questionnaire	Log of user interaction	Paper and pencil test	Computer-based test	
i.	What are the cognitive styles of the participants taking part in the study?				✓	Ratio scale
ii.	How do the participants rate the cognitive load of selected multimedia content?	✓				Category-ratio scale
iii.	What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?	Answered by means of statistical analysis				
iv.	To what extent do the presentation formats influence cognitive load?	✓	✓			Categorical data (formats) and continuous variables (cognitive load)
v.	How is learning performance influenced when content with different cognitive load is studied by learners with different cognitive styles?			✓	✓	Ordinal scale, variables (pre- and posttest score) are continuous

Table 1.4: Summary of the research instruments and nature of the data

1.8 Analysis of the data

Table 1.5 summarises the analyses that were done to address the hypotheses and answer the research questions. Demographic data was analysed by looking at frequency distributions, which were often presented as two-way frequency tables. Chi-square analyses, used to establish the relationship between frequencies for nominal and ordinal data (Cohen, Manion & Morrison, 2000) were conducted for these tables. Reliability of the various instruments was determined using Cronbach's Alpha.

Research question	Analyses
What are the cognitive styles of the participants taking part in the study?	Frequencies, Chi-square analyses, regression analyses followed up with confirmatory general linear modeling (GLM).
How do the participants rate the cognitive load of selected multimedia content?	Frequencies and Means procedures. GLM procedures and t tests of significance for independent samples to establish significance.
What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?	Pearson's Correlation procedure.
To what extent do the presentation formats influence cognitive load?	GLM procedures and t tests of significance for independent samples to establish significance.
How is learning performance influenced when content with different cognitive load is studied by learners with different cognitive styles?	Correlation procedures. Multiple regression analyses were used to extract the variables of interest. GLM procedures to analyse the variance and establish main and significant effects between the dependent variable (performance in the posttest) and several independent variables.

Table 1.5: Summary of the methods used to analyse the data

1.9 Limitations and strengths of the research

As much as I would have liked to have a study that only has strengths and no limitations, such a situation is rarely, if ever, possible.

The strengths of this study include the fact that the study:

- is one of the first studies to explore the relationship between cognitive load and cognitive style in detail.

-
- addressed an authentic learning environment as a response to the call for more research in such learning environments, and this authenticity was directed at the curriculum, the physical learning environment and the profile of the participants.
 - used a larger sample than many empirical studies in the cognitive style and cognitive load research streams.
 - explored a subject domain that has been neglected in both the cognitive style and cognitive load research streams.

The limitations of the study include the fact that:

- the very same authenticity, which was also a strength, made it very difficult to control the variables as rigorously as they are usually controlled in an experimental laboratory.
- seventy participants used the intervention simultaneously in contrast to a experimental laboratory where it is unusual to test more than 3 – 5 participants simultaneously, which made the control of the environment difficult.
- the limited time available in interacting with the participants excluded the possibility of any follow-up to explore further some of the qualitative aspects of the study.
- there was no time constraint for studying the content.

1.10 Organisation of the thesis

The outline of the chapters in the research report is presented below.

1 Overview and Orientation

This chapter introduces the study by defining the key constructs and discussing the purpose, background and rationale of the study. The research questions are presented. The research methodology and design is introduced. A summary of the limitations of the study provides additional context for the reader.

2 Literature Review

Chapter two presents the theoretical frameworks guiding this study and continues the argument for the rationale of this study. This critique of the literature is organised around the research questions and includes a review of the major studies related to the questions guiding this study.

3 Research Methodology and Design

3

Chapter three has three distinct sections. Section 1 describes the research approach and design, the sampling strategies used, the data and the design of the instruments. Section 2 describes the multimedia intervention used in the experimental design. Section 3 describes the implementation of the research design and methodology, for both the pilot and main studies.

4 Presentation and Analysis of Empirical Data

4

Chapter four presents the findings of the study. The data analysis includes both descriptive and inferential statistical techniques to confirm or reject the hypotheses. Each sub-question is addressed individually. The analysis then considers the interrelationship between the independent and dependent variables. Where appropriate, the findings are discussed relative to cognitive load and multimedia theory and research.

5 Discussion and Recommendations

5

Chapter five summarises the most important results and interprets these by considering the broader research field. This chapter also discusses the relevance, value and limitations of the research. Design guidelines for instructional designers are presented, with particular reference to designing for health science education. Recommendations are presented at the conclusion of the thesis.

1.11 Summary

This chapter has provided an introduction and orientation to the study. The constructs that will be used in this study were defined. A discussion of the background to and rationale for this study was presented. This culminated in the formulation of the research questions. An overview of the research design was presented, together with a brief description of the limitations of the study.

Chapter 2 will provide a critical review of the relevant research literature.



Literature Review and Theoretical Framework **2**

2.1 Introduction

The literature review serves two main purposes. Firstly it determines what has already been established in the empirical literature about cognitive load, cognitive style, multimedia learning and the relationships between these, and to identify the theoretical frameworks that guided the design of the study and the interpretation of the findings. Secondly it aims to identify any contradictions, ‘silences’, and gaps in these three areas of research. It is these contradictions, ‘silences’ and gaps that provide pointers for this particular study.

2.2 Overview of Chapter 2

This chapter is divided into two parts. Part 1 presents and discusses the three theoretical frameworks that guide and inform this study. These frameworks are discussed at the beginning of Chapter 2 since many of the studies reviewed in Part 2 use these theoretical frameworks and models. Part 2 presents a critical review of the literature about multimedia learning, cognitive load and cognitive style.

Table 2.1 (below) offers a more detailed outline of the sections into which Parts 1 and 2 are divided.

Part	Section	Focus of the review
Part 1 - Theoretical frameworks	2.1	Introduction
	2.2	Overview of Chapter 2
	2.3	Literature sources
	2.4	Introduction to the theoretical frameworks of this study
	2.5	The cognitive theory of multimedia learning
	2.6	Cognitive load theory
	2.7	Riding’s cognitive style model
	2.8	Summing up Part 1 of the Literature Review
Part 2 - Literature review	2.9	Introduction to Part 2
	2.10	Cognitive styles and multimedia learning
	2.11	Cognitive load and multimedia learning
	2.12	Multimedia in health sciences education
	2.13	Summary and conclusion

Table 2.1: Detailed outline of the parts of Chapter 2

2.3 Literature sources

A search in early 2004 that targeted only peer-reviewed journals in the Academic Premier Database returned over 600 references for learning style and approximately 450 references for cognitive style.

Since then, between 40 and 60 publications have been added to the cognitive styles literature list annually.

A similar search using the terms 'cognitive load theory' or 'cognitive load' in peer-reviewed journals returned close on 200 references. This number grew to 351 by July 2007. The extent of interest in cognitive load research is also reflected by the number of special editions that have been devoted to this topic in leading education journals since 2002.

The special editions mentioned above are in Table 2.2. It is a point worth making that the authors of the introductions and commentaries of these special editions read like a 'Who's Who' of Cognitive and Educational Psychology.

Year	Journal	Vol	No	No of articles	Editorial by	Commentary by
2002	Learning and Instruction	12	1	9	P. A Kirschner	M. Bannert / M. Valcke
2002	Educational Psychology Review	14	1	7	D.H. Robinson	W. Schnotz
2003	Educational Psychologist	38	1	8	F. Paas, A. Renkl & J. Sweller	S. Goldman
2003	Learning and Instruction	13	2	10	W. Schnotz & R. Lowe	S. Goldman / Riemann
2004	Instructional Science	32	1/2	10	F. Paas & J. Sweller	R. Rikers, P. van Gerven & H. Schmidt
2004	Learning and Instruction	14	3	9	R. Ploetzner & R. Lowe	M. Hegarty
2005	Educational Technology Research & Development	53	3	9	J.J.G. Van Merriënboer	G.R. Morrison / G. J. Anglin
2006	Learning and Instruction	16	2	8	F. Paas and T. van Gog	R. Moreno / J. Sweller
2006	Applied Cognitive Psychology	20	3	9	F. Paas and L. Kester	R.M.J.P. Rikers / J. Sweller

Table 2.2: List of special edition education journals devoted to cognitive load and/or multimedia research

Part 1 - Theoretical Frameworks

2.4 Introduction to the theoretical frameworks of this study

Decisions about the development of instructional material are not only content and context dependent, but depend increasingly on a skill set that is not part of the average lecturer or facilitator's experience (Inglis, 1999). Contemporary multimedia for face-to-face, blended and/or distance

learning environments are currently developed by teams, and it is a feature of these teams that each member of a team will bring different skills and perspectives to the design and development table. It is primarily the task of the instructional designer in any such team to decide on the variety and blend of instructional strategies and methods that will be used in any particular case. When making such decisions, an instructional designer needs to take the whole gamut of current theories about learning, available research-based evidence and empirically tested theory into account.

There are three mainstream theoretical perspectives that address the field of learning. They are behaviourism, cognitive science and constructivism. Instructional design is currently informed by all three of these theoretical perspectives.

Behaviourism as a learning theory focuses on changes in overt behaviour that can be observed and measured. While early behaviourist researchers used mainly animals as experimental subjects (Mergel, 1998), current behaviourist experiments with human beings achieve a desired behaviour by presenting content together with a stimulus (usually in the form of a question), to which the learner is required to respond. If the response to the question is correct, the learner is rewarded. This procedure is repeated until the learner responds automatically to the stimulus and the question. Learning of this kind takes place primarily as a result of association, and the repetition and reinforcement that are a feature of this method contribute to the success of the learning effort. Instructional strategies used in technology-supported learning that are based on behavioural theories include programmed learning, drill and practice and mastery learning. Teaching practices that focus on the correct use of feedback and reinforcement are also based on methods perfected by behaviourists. Because of the work of behaviourists, instructional designers routinely pay careful attention to the design of feedback mechanisms in learning contexts (Alessi & Trollip, 2001).

Constructivist theorists regard knowledge as something that is constructed by learners rather than something that is merely reproduced (Murphy, 2003). It is the view that knowledge is constructed when learners, often working together in a group, try to harmonise their existing perceptions, understanding and interpretations of a subject with the perceptions, understanding and interpretations of others. The use of such a method implies that the learners' final understanding of a particular knowledge domain will be different from their original conceptualisation and understanding of that knowledge domain. Another central tenet of this theoretical perspective is that there may be a number of individually constructed knowledge representations that are equally valid (Dalgarno, 2001).

Dalgarno describes three broad principles of constructivist theory:

- Each learner forms[his or her] own representation of knowledge.
- Learning occurs when learners actively explore their environment and find some inconsistency between their current knowledge representation and their experience.
- Learning occurs within a social context – this interaction is a necessary part of the learning process.

While there is general agreement about these principles among educationalists, their implications for teaching and learning are not quite so clear cut. Educationalists still differ about how to implement these broad principles in practice. Mayer (1997b) questions the value of focusing on doctrine-based research when educational practitioners and leaders are still asking for answers to some of the most fundamental educational issues such as 'How do students learn?' and 'How can learning be fostered?'

'Cognitive learning theory' is in fact a collective term for a number of theories that developed from the 1950s onwards. These theories were formulated in response to the perceived shortcomings of behaviourist theory (Mergel, 1998). Cognitive theorists regard learning as more than mere changes in behaviour, and contend that learning can in fact take place even when there are no external changes in a learner's behaviour. These theorists view learning as an internal process that involves memory, thinking, reflection, abstraction, motivation and meta-cognition (Ally, 2004). Cognitive theory focuses on the mental activity that takes place during the event that we call learning, and is based on assertions about how information is processed and how the brain develops and uses schemas to consolidate the acquisition and construction of knowledge (Mergel, 1998).

Several cognitive learning theories are concerned with the processing of information, memory, mental models and schema construction (Kearsley, 2004). The better-known of these theories include:

- ACT (Anderson, Bothell, Byrne, Douglass, Lebiere & Qin, 2004)
- Cognitive load theory (Sweller, 1998)
- Component Display Theory (Merrill, 1983)
- Dual Coding Theory (Paivio, 1986)
- Information Processing Theory (Miller, 1956)
- Cognitive Theory of Multimedia Learning (Mayer, 2003)

Two theories subsumed under the more general category descriptor of cognitive theory underpin this study: Mayer's Cognitive Theory of Multimedia Learning and Cognitive Load Theory (CLT). These theories guided the focus of the literature review in such a way that I was able to clarify my final research questions and obtain a research design appropriate for this study.

The third framework in this study is the model that I used to define and measure the cognitive styles of the participants in this study – a model proposed by Riding as early as 1991 (Riding & Cheema, 1991), and still in use today (Chen, Ghinea & Macredie, 2006).

2.5 The cognitive theory of multimedia learning

There are three important concepts that need to be clarified when one discusses multimedia learning: the delivery media, the presentation modes and the sensory modalities. 'Delivery medium' refers to

the system that an instructor uses to present instruction. Two widely used contemporary examples of such media are the paper medium and the computer medium. The medium of instruction has been the focus of a great deal of research and debate (Boling & Robinson, 1999; Chang, 2002; Erwin & Ricardo, 1999; Frith, Jaftha & Prince, 2004; Kekkonen–Moneta & Moneta, 2002; Vichitvejpaisal, Sitthikongsak, Parakkamodom, Manon & Petcharatana, 2001; Williams et al.2001). 'Presentation modes' refer to the format that an instructor uses to present the information. Examples of such presentation modes are words and illustrations (Carney & Levin, 2002; Dutke & Rinck, 2006; Mayer, Mautone & Prothero, 2002; O'Donnell, Dansereau & Hall, 2002; Rieber, Tzeng & Tribble, 2004). 'Sensory modality' refers to the information processing channel that a learner must use to process the information. These include the auditory, visual, olfactory and tactile modes. Research has sought to understand how learners integrate verbal and visual information when they engage in multimedia learning (Dempsey & van Eck, 2003; Mayer, 1997a; Moreno, Mayer, Spires & Lester, 2001).

Mayer (1997a) drew on the work of the generative theorists, Wittrock (1992) and Sternberg (1985), and the dual coding theory of Paivio (1986) when he proposed a generative theory of multimedia learning which he presented in the literature in 1997.

According to the generative theory (Wittrock, 1992) meaningful learning will occur when the learner creates relationships between the new concepts **and** between prior knowledge, experience and new information. This theory focuses on the generation of relations and not on the storage of information. Important processes which work together include:

- the processes of attention which are necessary for the active selection of stimuli.
- motivational processes which are stimulated by the plans and intentions of the learner, and which determine the degree of activity that the learner will devote to the stimuli and their meaning.
- knowledge creation processes which are driven by and depend on the various abilities that each individual possesses, for example verbal, spatial, analytic, holistic and propositional mechanisms.
- integration processes which enable new knowledge to be incorporated into existing constructs.

Dual coding theory (Paivio, 1986) presents the idea that cognitive processing takes place within the following two separate information processing systems:

- A visual system that processes visual knowledge from illustrations
- A verbal system that processes verbal knowledge mediated either by the written or the spoken word

In multimedia learning, that is to say, learning from words, pictures and sounds presented in any number of combinations, the learner needs to process information that arrives from more than one sensory channel. The central tenet of the cognitive theory of multimedia learning is that the design of

multimedia instruction affects the degree to which learners engages in the cognitive processes within these two systems.

The process of **selection** consists of paying attention to incoming information, sorting the relevant from the irrelevant parts, and adding the relevant parts to the visual and verbal short-term working memory. The process of **organising** information consists of making logical associations between the same type of information – in this case visual or verbal – and devising a mental model that is both coherent and comprehensible. Once the learner has created a visual and verbal mental model, she needs to **integrate** these two models on a one-to-one basis. This process of integration consists of connecting the newly created mental model with mental models that already exist in short-term memory. The best connections are made when the verbal and visual information are in short-term memory at the same time.

Mayer (1997a) goes on to explain how this theory can be used to make predictions about the following questions:

- Is multimedia instruction effective? (Dempsey & Van Eck, 2003; Hall, 2002)
- When is multimedia instruction effective? (Leahy et al. 2003)
- For whom is multimedia instruction effective? (Graff, 2005; Ross & Schulz, 1999)

Current research provides a set of guidelines for designing the multimedia presentation of information. This theory is illustrated in Figure 2.1 on the next page.

This theory, and the empirical work that has used this theory to explain the findings, will be applied in the following two ways in this study:

- The design guidelines that arise from the research will guide the development of the intervention used for this study. This will be discussed in more detail in Chapter 3.
- The theory will also be used to assess the meaning of the findings that arise from the empirical component of this study.

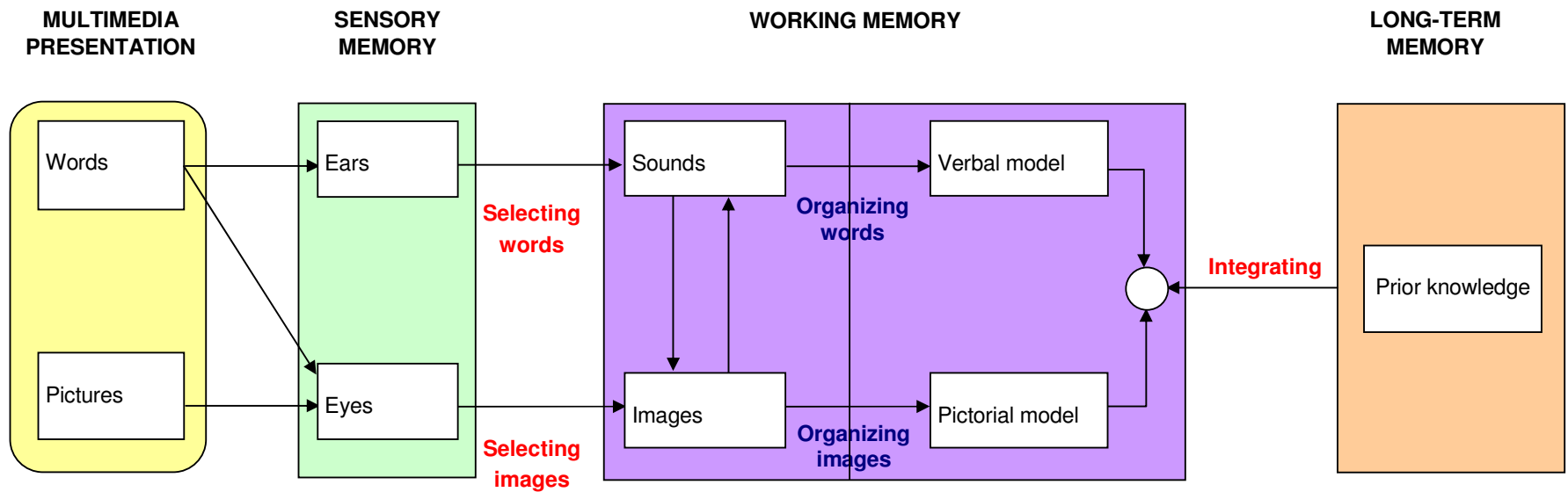


Figure 2.1: A framework for a cognitive theory of multimedia learning (Mayer, 2003:129)

It has already been established that the short-term memory has a limited capacity (Miller, 1956). The information that is held in short-term memory may be regarded as a cognitive load. A potential problem in all multimedia learning is that the processing demands of the learning task may exceed the capacity of the cognitive system (and the short-term memory in particular) to process large amounts of information in such a way that it can be transferred to long-term memory. When demand exceeds capacity, cognitive scientists talk about 'cognitive overload' (Mayer & Moreno, 2003).

This construct and the theory relevant to it will be discussed in more detail in the next section.

2.6 Cognitive load theory

The 'founding father' of cognitive load theory (CLT), which originated in the 1980s, is John Sweller (Sweller, 1988). The conventional classification of cognitive theories places this theory into the larger class of limited capacity theories (Goldman, 1991). This class of theories is based on the assumption that human beings have a finite cognitive capacity that includes, among other capacities, limited attentional resources and limited working memory. The consequence of this is that when a person reaches the limits of attention and working memory capacity, performance declines exponentially.

The development of this theory has been a work in progress that has now spanned almost twenty years (Paas, Renkl & Sweller, 2003; Sweller, 2006b; Sweller, 1988; Sweller & Chandler, 1991; Sweller & Chandler, 1994; Sweller, van Merriënboer & Paas, 1998). For the past twenty years, researchers from around the globe have continued to develop and refine this theory in practice and in carefully controlled empirical conditions. Ways of measuring cognitive load has also been a work in progress, with efforts to measure cognitive load described as early as 1994 (Paas, van Merriënboer & Adam, 1994b) and as recently as 2004 (Brünken, Plass & Leutner, 2004; Paas, Tuovinen, Tabbers & van Gerven, 2003). Sweller concluded a commentary for a special issue of the journal, *Applied Cognitive Psychology*, with the following statement:

Research using cognitive load theory is a vibrant, active enterprise around the globe. The papers of this issue provide both theory and data attesting to the strength of this continuing work.

Sweller, 2006b, pg 367.

2.6.1 The human cognitive architecture

The foundation for understanding and using the CLT is an understanding of the human cognitive architecture. The assumptions about the human cognitive architecture that underpin the development of cognitive load theory are based on the well known and widely accepted limitations of working memory. Pioneering work in this field was being undertaken by Miller as early as 1956 (Miller, 1956). The architecture referred to is a cognitive system that consists of three memory structures: a very

powerful long-term memory, a limited working memory and a limited sensory memory. Working memory can only handle a few new interacting elements at one time. Long-term memory expands processing ability due to the creation of schemas. Schemas are cognitive constructs that incorporate multiple elements of information into the ambit of one coherent cognitive construct (Paas, Renkl & Sweller, 2003). We are not consciously aware at any given moment of the information that exists in our long-term memory. We are only aware of whatever information appears in our working memory. Information is transferred into working memory from both the long-term memory and the sensory memory. Paradoxically, even though the working memory has a limited capacity, its limitations seem only to apply to information arriving from the sensory memory. A vast amount of information can be transferred from the long-term memory to the working memory without overloading the working memory in any way (Sweller, 2004).

2.6.2 The construct 'cognitive load'

Cognitive load is a multi-dimensional construct. It refers to the extent to which the cognitive system is loaded during the process of performing a task. The construct is conceptualised as having two factors: *causal factors* which affect cognitive load, and *assessment factors* which are effected by cognitive load. The construct 'cognitive load' is illustrated here in Figure 2.2 by means of an adaptation of an original representation by Paas and van Merriënboer (1994a). Figure 2.2 illustrates the interrelationships between the various elements in this complex construct.

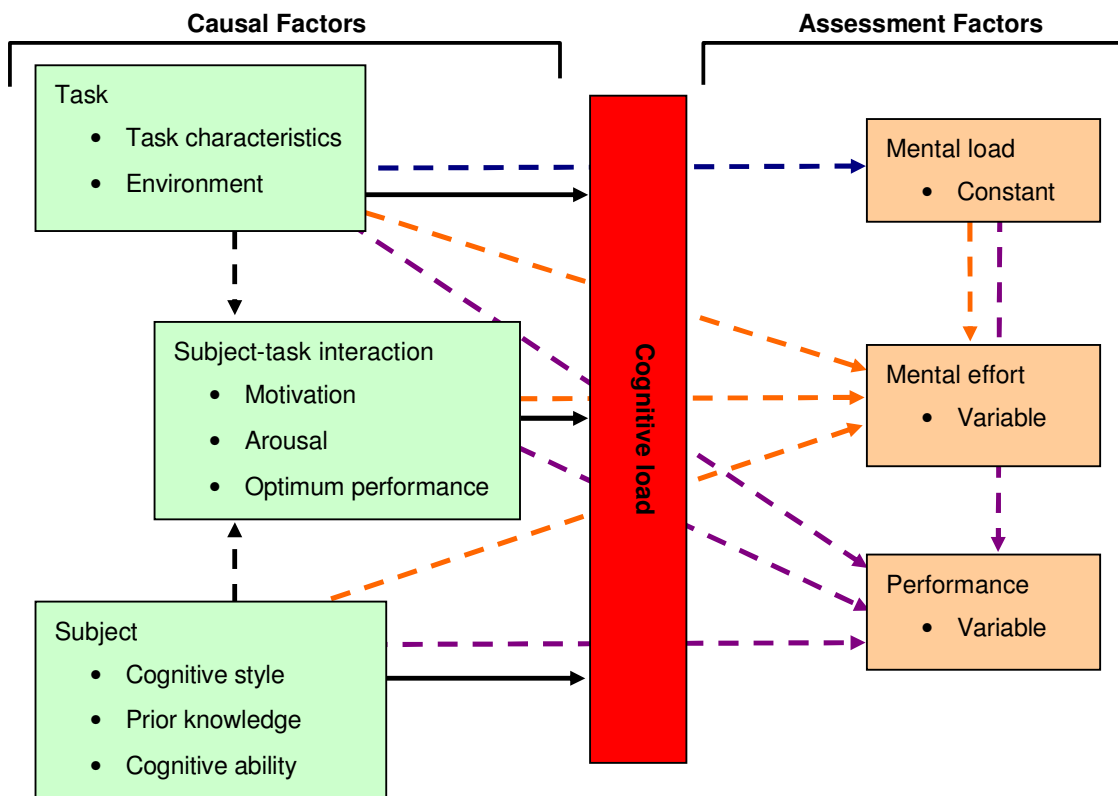


Figure 2.2: The construct cognitive load

Some of these factors are unstable because they change from one learning context to another or because they can be changed by someone such as an instructional designer. The various factors are often closely interrelated (the interrelationships are illustrated by the dotted lines in Figure 2.2). Factors that influence mental load are indicated by the blue dotted line; those that influence mental effort are indicated by the orange dotted line, and those that influence performance are indicated by the purple dotted line.

Task characteristics, which are generally unstable, and which can be manipulated, include:

- Structure of the task
- Novelty
- Type of reward system
- Time pressures to complete the task
- Environmental factors – which can include things like noise, temperature

Subject characteristics which, on the other hand, are relatively stable, include:

- Cognitive ability
- Cognitive style
- Prior knowledge
- Previous experience

This theory suggests that there is a relationship between cognitive style and mental load and effort. This study will explore this relationship in more detail.

Whenever task-subject interactions occur, we find the following set of relatively unstable factors:

- Motivation
- Arousal
- Optimum performance

The nature and dynamics of the assessment factors are illustrated in Figure 2.2. Mental load, which tends to be constant, is imposed by the task. Mental effort, by contrast, is determined primarily by the subject even though it is influenced by both the task characteristics and subject-task interactions. Mental effort is variable. Learners who are highly motivated usually invest more effort in the learning of a task (Paas, Tuovinen, van Merriënboer & Darabi, 2005). And because performance is the visible outcome of mental load and effort, it will also tend to vary.

Early research in this field tried to find the reasons why some things are more difficult to learn than others (Sweller & Chandler, 1994). By closely analysing what learners were learning, these early

researchers were able to reduce what was being learned to the most basic elements that made up the learning material and how these elements were interacting with each other. While some elements (such as vocabulary) do not interact at all and can be learned in isolation, other elements (such as the grammar of a language) need to be learned together. But learning to write a sentence, for example, will require both a knowledge of vocabulary and grammar, and such a task is therefore more complex than merely learning vocabulary in isolation. Interaction between the different elements of learning can vary from simple to fairly complex. According to the cognitive load theorists, the cognitive load of a task is primarily influenced by the number of elements that need to be processed simultaneously and the level of element interactivity inherent in whatever material is being learned. The higher the level of interaction, the greater the cognitive load (Sweller & Chandler, 1994).

During the course of testing and refining, the researchers conceptualised and described the following three categories of cognitive load: intrinsic, extrinsic and germane cognitive load.

Intrinsic cognitive load is determined by the level of element interactivity inherent in the information that must be processed. It is an intrinsic part of the material to be learned and cannot be altered by instructional manipulation. For example, there may be 10 steps in a particular process and these must follow each other exactly. These 10 steps must be learned in a specific sequence in order to understand the process. Reducing the process to 7 steps in order to make the instruction shorter or simpler is not an option, since the person executing the process must still go through all 10 steps. Because different materials differ in their levels of element interactivity, they also have different intrinsic loads. Intrinsic load has also been described in terms of the complexity of the learning content relative to the learner's prior knowledge (Renkl & Atkinson, 2003).

Extrinsic cognitive load is determined by the manner in which information is presented to the learner and the activities in which he/she is required to engage. Using this sample example, the designer can select to present these 10 steps as text across two or more pages. This design makes the process of learning these 10 steps unnecessarily more difficult. The learning becomes burdensome. If this load is unnecessary and interferes with schema acquisition and automation, it becomes extraneous and is referred to as extraneous or ineffective cognitive load.

Germane cognitive load is a concept that has only recently been addressed in the empirical literature (Paas & van Gog, 2006; Renkl & Atkinson, 2003). 'Germane load' refers to the effort that the learner applies when trying to understand a task or solve a problem (Renkl & Atkinson, 2003). Although deeper processing imposes a load on the cognitive system, it is a load that promotes schema formation. If, however, the capacity of a working memory has already reached its limit because of intrinsic and extrinsic loads, no spare capacity will be available to support deeper processing. 'Germane load' refers to the manner in which information is presented to the learner **and** the types of learning activities required of the learner.

Extrinsic cognitive load has two dimensions, and each of these is under the deliberate control of the designer who must both limit the factors which impose extraneous load and introduce deliberate strategies that increase germane load (Paas, Renkl & Sweller, 2003). When considering germane cognitive load, an instructional designer will deliberately look for ways to enhance schema acquisition and automation. Since extraneous load can be closely bound to germane load in some learning environments, designing with the purpose of accommodating both conditions may be very difficult. Some recent research has been devoted to understanding how intrinsic load can be reduced in circumstances where it is difficult to manipulate both extraneous and germane loads (Ayres, 2006a, Paas et al., 2004).

These different forms of cognitive load are all additive.

2.6.3 Practical application of cognitive load theory

The primary purpose and relevance of this theory has been to provide a framework for instructional design (Sweller, van Merriënboer & Paas, 1998). Cognitive load theory emphasises the constraints of the working memory, and these constraints exert a decisive influence on the effectiveness of the instructional design of learning materials and environments.

Early research also focused on the constraints of the working memory. Sweller (2004) is of the opinion that the if the major purpose of instruction is to store information in the long-term memory, then cognitive load theory must concern itself with how instruction will facilitate the acquisition of appropriately structured information in long-term memory. It is therefore necessary for instructional design to take cognisance of how long-term memory and short-term memory function in human cognitive architecture. When they are considering the function of long-term memory, instructional designers will try to find ways of facilitating the acquisition of appropriately structured information. When they are considering the function of working memory, instructional designers will try to find ways of compensating for the limited capacity of the working memory.

2.7 Riding's cognitive style model

In the section I will discuss the third model that I have used to guide the study, namely the cognitive style model of Riding.

Witkin et al. (1977) describe four general but essential characteristics of cognitive styles. These cognitive styles are:

- concerned with the form rather than the content of the cognitive activity in those cases where the focus is on the processes underlying the differences between the styles.
- pervasive because they cut across social, personality and intellectual aspects of life.
- stable over time because they are not susceptible to change.

- bi-polar. This is important in distinguishing style from intelligence and other ability dimensions.
- completely value-free. This offers both advantages and disadvantages to individuals on either end of the bi-polar dimension.

The term 'cognitive style' is used throughout the literature to describe an individual's preferred approach to organizing and representing information (Riding & Rayner, 1998). There are three elements which form the core of an individual's persona: affect or feeling, behaviour or doing, and cognition or knowing. The cognition or 'knowing' element of a persona is structured and organised in terms of an individual's cognitive ability and cognitive style.

The internal processes that accompany cognitive style are reflected in the way a person approaches learning. These are life-long processes. An individual's repertoire of learning strategies, in combination with that person's cognitive style, establish an individual's idiosyncratic and personal learning style. There are many constituent parts of cognitive style. They include the manner in which a learner imposes structure on learning materials, the ability of a person to perceive spatial location or orientation in space (which is important when one is navigating non-linear learning programs and resources), the value that an individual places on the bigger picture as opposed to the parts that make up the whole, and the extent of an individual's preferences for using textually rich or graphically rich resources to construct knowledge and meaning (Graff, 2003b).

Riding and Cheema (1991) undertook an extensive review of the literature on cognitive style theory. They identified over 30 labels that researchers have used for cognitive styles. A detailed analysis of these labels and concepts suggest that they all fall into two principal cognitive style groups – the Wholistic-Analytic cognitive style family and the Verbaliser-Imager style family. Appendix A contains a very brief summary of the Wholistic-Analytic and the Verbaliser-Imager cognitive style labels that are most often described in the literature prior to the early 1990s. Riding and Cheema (1991) point out that while many other labels exist not all of them have attracted equal attention and empirical evidence for these styles remains sketchy.

After they had analysed cognitive style, Riding and Rayner (1998) proposed a new cognitive style model (illustrated here in Figure 2.3). I used this model when designing this study.

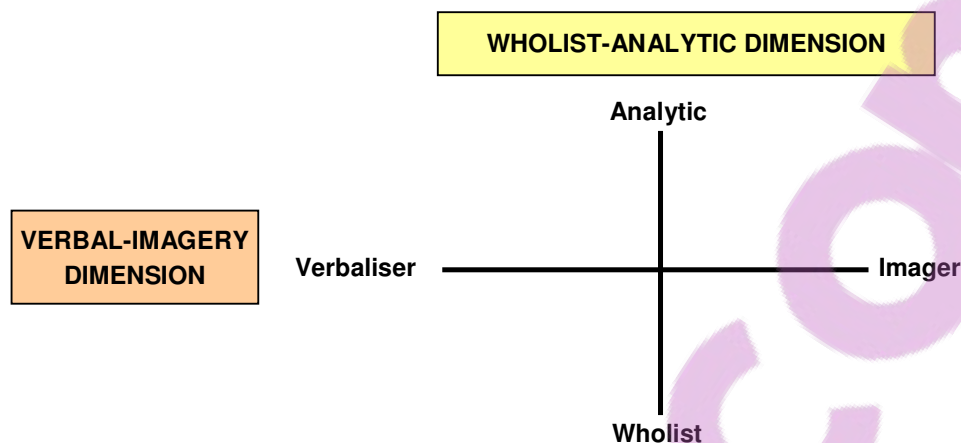


Figure 2.3: Riding's Cognitive Style Model

The two basic dimensions of cognitive style may be summarised as follows.

- The Wholist-Analytic Style determines whether an individual tends to process information as a whole or in parts.
- The Verbal-Imagery Style determines whether an individual is inclined to represent information during thinking verbally or thinking by means of mental images (Riding & Rayner, 1998).

These dimensions are measured using a test developed by Riding (Riding, 2005a). The Cognitive Styles Analysis (CSA) provides a score for each dimension in the cognitive style model. On the Wholistic-Analytic dimension, a low ratio corresponds to a Wholist and a high ratio to an Analytic style. Ratios in the middle are considered to be Intermediate. On the Verbaliser-Imager dimension, a low ratio corresponds to a Verbaliser and a high ratio to an Imager, with the intermediate position being described as Bimodal.

Riding proposed that each dimension is a continuum, and that labels are only attached to the ranges of a continuum for descriptive purposes (Riding et al., 2003). Individuals are distributed along the continua.

The two dimensions are furthermore independent of one another. This means that the position of an individual on one dimension of cognitive style does not affect his/her position on the other. This has been confirmed in the empirical literature (Douglas & Riding, 1993; Rezaei & Katz, 2004; Riding & Grimley, 1999; Riding & Mathias, 1991; Riding & Rayner, 1998; Riding & Sadler-Smith, 1992; Riding & Staley, 1998).

2.7.1 Wholist-Analytic Style

Wholists, as the term suggests, habitually look at the totality of any context, condition, situation or phenomenon. This ability to be able to see the whole 'picture' gives rise to a balanced view and is one of the strengths of this style. The negative attribute of this style is that such people find it difficult to separate information into its constituent logical parts. Not only do they not separate the parts, but they also tend to blur the distinction between them. The Wholist finds it difficult to distinguish the component parts of a whole piece of information.

Analytics, by contrast, see a situation as a collection of parts. Because they can analyse information into its constituent parts, they are able quickly to arrive at the heart of a problem. While their particular mode of perception makes them good at identifying similarities and detecting differences, they often focus on only one or two of these parts to the exclusion of the others. When a person focuses on just one aspect of the whole at a time, it creates the potential for distortion and exaggeration and for making one particular part of a whole more important than it should be. When this happens, such a part may be accorded an importance that is out of all proportion to the total situation.

When instruction is being designed, it is important to help Wholists to see detailed structure and sections of the learning material, and to help Analytics to see a unifying overview so that they can integrate sections into a unified view.

2.7.2 Verbal-Imagery Style

When it comes to the mode of presentation of information and learning performance, Imagery learners generally learn best from pictorial presentation while Verbalisers learn better from text (Riding & Douglas, 1993; Riding & Rayner, 1998).

When it comes to type of content, Imagery learners find it easier to cope with concrete and readily visualised information rather than semantically and acoustically complex details (Riding & Calvey, 1981). In another study, Riding and Read (1996) found that Imagery learners prefer picture materials, especially if the subject lends itself to pictorial content. Riding and Mathias (1991) explored the influence of Wholistic-Analytic style, Verbaliser-Imager style and gender on the mode of instruction. This research demonstrated a significant effect for the two-way interaction of Wholistic-Analytic style and Verbaliser-Imager style and their effect on mode. Since the preference of Wholists depended on their Verbaliser-Imager style, they preferred word formats if they had a Verbaliser style, while Imagery learners favoured the pictures. For the Analytics, the preference lay between the modes and slightly inclined to the pictorial, and verbal-imagery style had relatively little effect on their preferences. Riding and Mathias concluded that this lack of interaction with the verbal-imagery dimension could be explained by the fact that Analytics are able to adopt strategies that utilise the strengths of their style.

Riding and Read (1996) argue on the basis of some very early cognitive style research that looked at two levels of control of verbal and imagery performance (namely, voluntary and involuntary) that Verbalisers do not use images to any great extent during involuntary information processing, although they can generate them successfully by using conscious effort. Imagers, on the other hand, habitually use involuntary imagery as a means of representing information.

In summing up cognitive style, it is likely to affect learning performance on two levels: perceptual and conceptual. From a perceptual point of view, a Wholistic learner finds it more difficult to sort out the detail out while the Analytic learner finds it more difficult to integrate learning material into a whole. The conceptual level addresses the ability to analyse content. All learners need to be able to analyse information conceptually, and the structure of the learning material influences this process. Various instructional strategies such as overviews or advance organisers may help with this process. While Wholists, for example, will benefit from help in structuring material, Analytics are generally able to impose their own structure upon it.

Since most tasks require both structure and representation, Riding and his colleagues have proposed that individuals will develop strategies to use alternative methods of structuring tasks when their own style is inappropriate. In this way, combinations of methods serve to compensate for individual weaknesses (Riding & Staley, 1992; Riding & Rayner, 1998).

2.8 Summing up Part 1 of the Literature Review

This section concludes Part 1 of the Literature Review. Two theories and one model, both of which provide the theoretical framework for this study, were reviewed, and key issues relating to each were discussed.

The first theory, the cognitive theory of multimedia learning, addresses the way in which information is conceptualised as processes in the sensory, working and long-term memory. The cornerstones of this theory are the generative theories of Wittrock and the dual coding theory of Paivio. According to this theory there are two sensory modalities involved in the receiving and decoding of incoming sensor stimuli: the eye and the ear. Two distinct processes operate within each channel (selecting and organising) before the information is finally integrated (the third process). There is no reference to cognitive style in this theory.

The second theory that is used in this study is the cognitive load theory of Sweller. The foundation of this theory is a description and understanding of the function of the human cognitive architecture. Sweller takes the position that human beings possess a finite cognitive capacity which is characterised by, among other things, limited attentional resources and a limited working memory. When the limits of attention and working memory are reached, performance declines. The information that must be processed by this cognitive architecture is called cognitive load, and when cognitive load is too great, the resulting situation is called cognitive overload. Three following types of cognitive load

have been identified: intrinsic load (which relates to difficulty and the interrelatedness of the content), extraneous load (which relates to design and which is a hindrance to cognitive processing), and germane load (which is concerned with motivational issues and strategies that promote deeper processing). There are several different contributors to cognitive load, and these include cognitive style. The effects of cognitive load are manifested in the performance of the learner. A learner is also able to report the amount of mental effort invested in a particular learning task, and mental effort is used as an indicator of cognitive load. This theory is applicable to all learning, irrespective of the delivery medium.

The final model that I used in this study is Riding's Cognitive Style model, which has been described as a orthogonal, two-dimensional model. These two dimensions are independent of one another. The one dimension addresses the way in which information is processed by the elements of the human cognitive architecture, and has been called the Wholistic-Analytic style. The other dimension addresses the way in which information is represented in the human cognitive architecture, and has been called the Visualiser-Imager dimension. A person's individual styles can be plotted the points of these two continua. This model is applicable to all learning, irrespective of the delivery medium.

With these theories as background, I now turn to the second part of Chapter 2. This second part offers a critical review of the empirical research that makes use of these theories and the cognitive style model.

Part 2 - Literature Review

2.9 Introduction to Part 2

Part 1 of this chapter described and reviewed the theoretical frameworks that are used in this study. Part 2 expands upon the rationale of this study and furthers the arguments (presented in Chapter 1) that justify the research question by means of a critical examination of the theoretical, empirical, methodological, media and contextual dimensions of cognitive load, cognitive style, learning style and multimedia learning literature.

My review of the theoretical frameworks enabled me to make sense of the data implied by the research questions. The rationale of the study (Chapter 1) and my subsequent consideration of the theoretical frameworks provided the necessary background for the central theme of this study, which is multimedia learning. The context of the study is health science education. This study makes use of this theme and context to explore two of the many factors that impact on multimedia learning, namely cognitive load and cognitive style. I noted above that the constructs 'cognitive load' and 'cognitive style' have both already been extensively researched.

Figure 2.4 illustrates how a superficial and largely non-critical review of the literature about multimedia learning would define and illustrate the relationships between these three constructs.

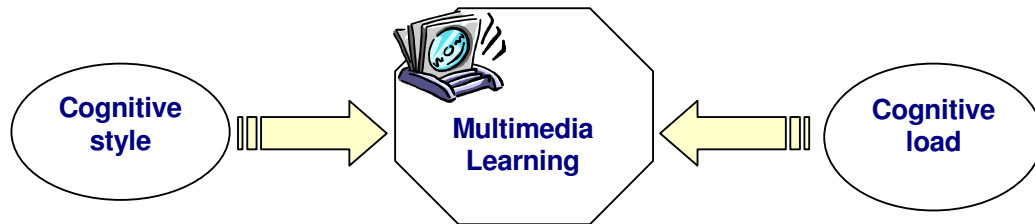


Figure 2.4: Relationships between cognitive load, cognitive style and multimedia learning

Although it might therefore seem simplistic to consider these three streams (cognitive style, cognitive load and multimedia learning) in parallel, two factors became evident in the early stages of the literature review: firstly, researchers in each of the streams often cross-reference the studies of researchers in at least one of the other two streams, and, secondly, researchers from different streams collaborate with one another on research projects.

The reality that informs this situation is that few of the factors that impact on the effectiveness of multimedia learning operate in isolation.

The relationship between the three constructs is illustrated in Figure 2.5 below.

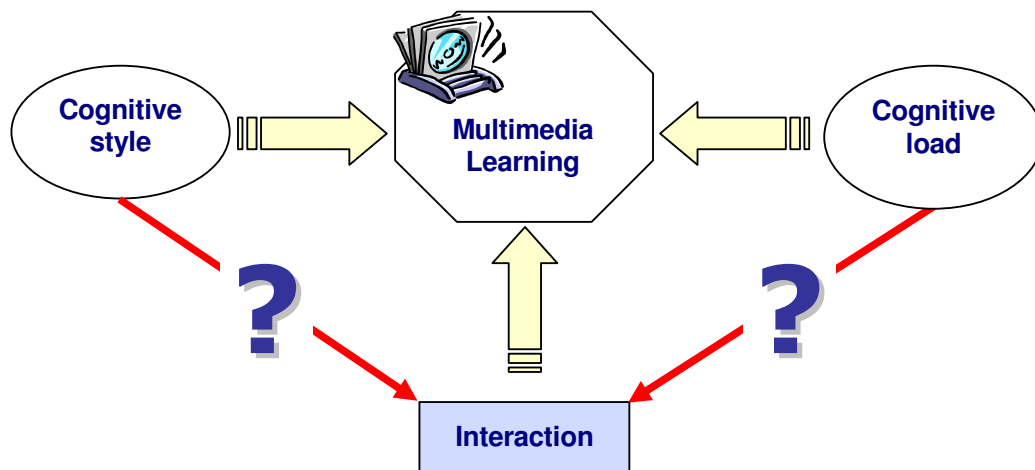


Figure 2.5: New relationships between cognitive style, cognitive load and multimedia learning

This conceptualisation supports a review and critique of the literature that is organised in terms of the research questions which are designed to explore the relationships between cognitive style, cognitive load and multimedia learning. It requires an in-depth review that looks for answers to questions such as: 'What does the research into cognitive style and multimedia learning have to say about cognitive load?' and 'What does the research into cognitive load and multimedia learning have to say about cognitive style?'

Since the volume of accumulated literature on these topics is vast, any attempt to offer a broad overview of the research might end up with providing only a superficial commentary on this literature. Even though reference will be made where necessary to more general areas of research, the focus of this review has been narrowed so that it considers only issues that are pertinent to this study. I have been careful to make reference to relevant research in health science education because that is the context for this study. This approach to the literature review is illustrated in Table 2.3.

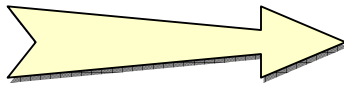
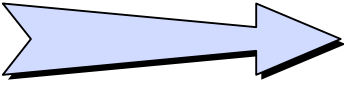
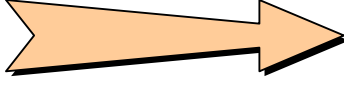
The streamwill focus on the	...specific issues
Multimedia learning		Use of animation, sound and static images
Cognitive load		Cognitive load of multiple representations.
Styles research		Cognitive style in tertiary / adult education

Table 2.3: Approach to the literature review

2.10 Cognitive styles and multimedia learning

The research question relevant to this review is 'What are the cognitive styles of the participants taking part in the study?' I have divided the discussion of the literature, which included a review of close on 80 articles, into six major areas:

- Contributions to the theory
- Measurement of cognitive style
- Cognitive style and learner characteristics – gender, age, personality, intelligence
- Cognitive style and behaviour – attitudes, work orientation
- Cognitive style and instructional design practice
- Cognitive style and academic achievement

The critical review examines the design guidelines that flow out of cognitive style research and reflects on what the cognitive load perspective would have to say about these findings and subsequent design guidelines.

The renewed interest in cognitive styles and its application in the field of education gained momentum in the late 1980s and early 1990s (Riding & Cheema, 1991). Initial efforts were aimed at reducing the confusion surrounding style groups and labels. Riding and his colleagues, working at the University of Birmingham, were some of the first researchers to propose a model that appears to consolidate the work of other researchers (Rayner & Riding, 1997; Riding & Cheema, 1991; Riding & Rayner, 1998).

I described this model in Part 1 of this chapter. Cognitive style models and dimensions that have been more widely used than others as theoretical frameworks for empirical research since 1990 include the following:

- Riding's cognitive style model. This model has been widely used by researchers **other** than Riding and his colleagues. There have been at least 38 of these empirical studies (Riding, 2005b) since 1990 (Chen et al., 2006; Cunningham-Atkins et al., 2004; Evans, 2004; Graff, 2005; Graff, 2003a; Graff, 2003b; John & Boucouvalas, 2002b; McKay, 1999).
- Field Dependence/Independence. This model is based on the work of Witkin (Ford & Chen, 2001; Ghinea & Chen, 2003; Griffin & Griffin, 1996; Macnab, Hansell, & Johnstone, 1991; Palmquist, 2001; Triantafillou et al., 2004; Wooten, Barner & Silver, 1994; Zhang, 2004).
- Reflection-Impulsivity. This model is based on the work of Kagan (Nietfeld & Bosma, 2003; van Merriënboer, 1990).
- Rational/Analytical-Intuitive/Global dimension. This model is based on the work of Allinson and Hayes. The CSI is used to measure these dimensions (Allinson & Hayes, 1996; Armstrong, 2004; Armstrong & Priola, 2001).

2.10.1 Contributions to the theory

Using only five cognitive style dimensions; field dependence-field independence, Wholist-Analytic, Visualiser-Imager, Reflective-Impulsive and Analytic-Impulsive, I went back to the literature to look for evidence of continued theory development within these five style dimensions. I only found evidence of theory development for the Wholist-Analytic and Visualiser-Imager dimensions. Erhman and Leaver (2003) have also proposed a new cognitive style model for use in language learning, but this will not be discussed in detail in this literature review. The reason for this is that they propose to apply the model to language learning and I feel such a focus would be too narrow to serve as a model in this study, even though it does consolidate the ten different style dimensions.

John and Boucouvalas (2002c) note that Riding's classification of cognitive style has not been as successful as was expected in accurately predicting the performance of individuals engaged in complex tasks that use multimedia. These researchers in fact use Mayer's cognitive theory of multimedia learning to explain this lack of success, even though they do not actually refer to this theory by name. They ascribe their findings to the differences in the ways that users perceive different computerised media. This happens because information from different media is received by means of different sensory organs and is processed by different parts of the brain.

John and Boucouvalas (2002a) propose a new definition of cognitive style that takes an auditory component into account and is based on the ways in which individuals perceive different types of media. In order to test this theory, they developed a different but equivalent set of questions that can be used with the questions in Riding's CSA. In this new test the Verbal-imagery questions were spoken instead of being displayed as text, and the Wholistic-Analytic questions were transmitted by

means of sound effects rather than graphics. The performance of subjects in the visual test was then compared to their performance in the audio version. The hypothesis of the researchers was that if cognitive style influences an individual's perceptions of both visual and auditory information, there should be a strong correlation between the ratios achieved in both versions of the test. The researchers found there were strong correlations between both versions of the text in the Analytic-style questions, as indeed there were with the questions in the Imager-style questions. In spite of this, the overall style dimensions did not correlate well across the two versions of the test. The differences in performance between the visual and audio experiments seems to indicate that Riding's classification of cognitive style does not explain the performance of users when information is presented using a combination of visual and audio components. Is this model of style therefore able to explain multimedia learning which – in multimedia research and the context of this study – includes both an audio and visual component? John and Boucouvalas's proposal to add a visual-audio dimension to the current Verbaliser-Imager classifications does seem to provide new direction for cognitive style theory.

The fact that current technology is able to deliver richer multimedia experiences than the technology of the late 1980s and early 1990s (the period during which Riding and colleagues published their work on styles) was able to do, might well provide a motivation for a timely re-consideration of cognitive style theory. There are, however, two other possible explanations for these findings, which warrant further exploration and empirical research. Firstly, do the results obtained by John and Boucouvalas not simply provide additional evidence for the dual-coding theory of Paivio, rather than being indicative of a third style dimension? Secondly, to what extent did the cognitive load of these different formats play a role?

Kozhevnikov, Hegarty and Mayer (2002) have worked on revising the Visualizer-Verbalizer dimension of style. Their review of the research into this style highlighted the fact that there has been (and still is) a debate about the actual validity of this style dimension. They are critical of previous research for focusing on results that confirm the classification of Visualizers as people who have a developed imagery ability and Verbalisers as those who have a less-developed imagery ability, rather than for attempting to establish a clear relation between the preference to process information visually and performance on imagery tasks. Kozhevnikov et al. conducted research with a view to extending the visual construct in this style dimension. Their hypothesis was that there are two qualitatively different types of Visualizers who process visual spatial information, generate visual images and solve problems presented visually in different ways, and they devised three experiments to investigate this style dimension. The participants completed spatial ability tests, a verbal ability test, and a Visualizer-Verbalizer cognitive style questionnaire. Experiment 1 established that the spatial ability tests did not correlate with the cognitive style test. When they compared the performance of the Verbaliser and Visualiser groups on the spatial ability tests, the findings revealed that the majority of the Verbalisers were of average spatial ability, while the Visualizers were not a homogenous group. They identified two groups of Visualizers: a group with high spatial ability (the spatial type) and one with low spatial ability (the iconic type). They followed this up with two experiments that investigated

how these two types of Visualizers process visual-spatial information, generate mental images and solve problems presented visually.

They found that:

- the two types of Visualizers interpret motion graphs differently. Low-spatial Visualizers interpreted the graphs as pictures and relied mostly on visual (iconic) imagery, while high-spatial Visualizers constructed more schematic images and manipulated them spatially.
- there were differences in the flexibility with which the two groups used images. Iconic types tended to generate images by activating their visual memories and looking for patterns with the closest match, while the spatial types analysed the problem in smaller chunks and worked it out that way.
- Verbalisers and Visualizers also differ with regard to problem-solving strategies. While Verbalizers of low and high spatial ability did not have any clearly observable preference for using visual or spatial imagery, Visualizers differed in how they solved problems: low-spatial Visualizers used visual-pictorial imagery and high-spatial Visualizers used spatial-schematic imagery.

Their findings lead them to conclude that imagery is not general and undifferentiated, but is composed of different, relatively independent visual and spatial components. Using the outcomes of their experiments they proposed two qualitatively different types of Visualizers – low and high spatial Visualizers. Another very interesting statement that the literature review for their study produced was that spatial imagery is not limited to the visual modality, but might include auditory or tactile imagery. Although they did not follow up on this in their experiments, this statement suggests that the work of John and Boucouvalas (discussed above) and their proposal to add an audio component to both the Wholistic-Analytic and Verbaliser-Imager classifications seems to be a valid direction for the development of this cognitive style model.

Educational practice in several disciplines could benefit from a clearer understanding of the relevance and importance of spatial ability for cognitive style theory. In the medical field, for example, technological advances have made it possible for surgeons to use laparoscopes to carry out surgical interventions in situations where the only visible image of the surgical site is a 2D radiographic image. If the Verbaliser-Visualizer dimension of cognitive style is expanded to include a more multidimensional Visualiser construct, instructional strategies using multimedia will need to be reconsidered and adjusted to facilitate learning. Strategies such as virtual reality and 3D modelling could be used to teach both the visual and spatial aspects of subjects like anatomy and surgical skills acquisition. There is evidence that such strategies are already being used in health sciences education today (Garg, Norman, Spero & Taylor, 1999; Levinson, Weaver, Garside, McGinn & Norman, 2007; Mantovani, Castelnuovo, Gaggioli & Riva, 2003; Vernon & Peckham, 2002).

2.10.2 Measurement of cognitive style

2.10.2.1 Riding's CSA and related research

The validity of research findings cannot be divorced from the reliability and validity of the instruments used to measure the variable under study, which, in this case, is cognitive style. The psychometric testing of both the Cognitive Styles Analysis (CSA), which assesses the style dimensions in Riding's model, and the Cognitive Styles Index (CSI), which assesses the rational/analytical-intuitive/global dimensions, has been the focus of debate between several authors over a period of time. I considered both debates when looking for an instrument to use in this study.

Riding's Cognitive Styles Analysis (CSA) is a computer-based test that takes approximately 20 minutes to complete. It was developed to overcome two problems: firstly those associated with the instruments developed by Witkin, which were used to test for Field Dependence/Independence, and secondly those associated with the traditional self-reporting scales used to assess imagery performance (Riding & Rayner, 1998).

The CSA, which is well described in the literature (Graff, 2003b; Riding & Cheema, 1991; Riding, 1997; Riding & Grimley, 1999; Riding & Rayner, 1998), directly assess both ends of the Wholist-Analytic and the Verbaliser-Imager dimension. There are three sub-tests. The first sub-test assesses the Verbal-Imagery dimension and the other two sub-tests assess the Wholistic-Analytic dimension. The computer records the response times to the various statements in the test and calculates the corresponding ratios. More detail about measuring cognitive style, using Riding's CSA are provided in Appendix P.

It has been established that these dimensions are independent of one another. This means that the position on one dimension does not influence the position on the other (Douglas & Riding, 1993, Rezaei & Katz, 2004; Riding & Grimley, 1999; Riding & Rayner, 1998; Riding & Sadler-Smith, 1992; Riding & Staley, 1998). Since these two dimensions are independent of one another, they are discussed separately in this study.

Riding's CSA, which is available commercially, is accompanied by a manual that explains how to administer the test and interpret the results (Riding, 2005a).

Is the CSA a valid and reliable test? Riding writes:

In considering psychological assessments the most important feature of a test is its construct validity - if there is no evidence that it assesses what it purports to measure then it is of no use

(Riding, 2005b, pg 6).

With research into the CSA the primary emphasis has therefore been to demonstrate its validity.

Riding and Rayner (1998) summarise the evidence for validity as follows:

- The independence from other variables: gender, intelligence, common personality measures.
- The relationships between a wide range of behaviours: learning, subject preferences, social behaviours, occupational suitability.
- Evidence of a physiological basis.
- Conformity to the requirements of style.

Riding's CSA, and especially the test to determine the Verbaliser-Imager dimension, has been criticised in the literature for its lack of reliability, stability and internal consistency (Ong & Milech, 2004; Peterson, Deary & Austin, 2003a; Rezaei & Katz, 2004). The validity of the Verbaliser-Imager subtest was also recently called into question (Massa & Mayer, 2005). Mayer's critique is based on his assertion that the measures were not derived from an authentic multimedia learning situation.

Peterson, Dreary and Austin (2003a) investigated the reliability of Riding's CSA by using parallel form, test-retest and split-half design techniques. The retest interval was only 7 days. The parallel form and test-retest reliabilities were low. Riding (2003) responded to this work by citing eight limitations in the study of Peterson et al (2003a). He did not dispute the need for more investigation to determine the reliability of the CSA. In the final response by Peterson et al. (2003b) to Riding's critique of their research, they asserted that the criticisms of Riding were incorrect. The thrust of their argument was that Riding's critique merely distracted from the issue on hand, which was the need for a more rigorous assessment of cognitive style. These researchers explained that they were compelled to develop a different form of the CSA because of the limitations of the commercial version. They concluded that their study represented an attempt to address a valid concern in the domain of cognitive style research, namely, that not enough research has been undertaken to address the reliability of the tools used to assess cognitive style.

There have been subsequent attempts to address the reliability of the CSA. Parkinson, Mullally, and Redmond (2004) used only the test-retest technique to determine reliability in order to avoid the same rebuttal by Riding if they used the split-half technique or tried to develop a parallel form of the CSA. They improved on the study by Peterson et al. (2003a) by making the time interval between the two test sessions was longer: 14 days in the first study and 23 months for the second study. The Wholistic-Analytic test-retest correlation was stable over time but still low at about 0.3 for both studies. The Verbaliser-Imager retest correlation was not only low, it was also not stable at -0.19 and 0.36 in study 1 and 2 respectively. Parkinson et al. are of the opinion that low reliability continues to raise concerns about the validity of the CSA test despite Riding's assurance about the validity of the CSA.

Rezaei and Katz (2004) also investigated the reliability of the CSA in three studies that used a test-retest interval of one week for the first study and one month for the second and third studies. The

range of reliability for the Wholistic-Analytic dimension across the three studies was between 0.42 and 0.55, and for the Verbaliser-Imager dimension it was between 0.30 and 0.45. This, like the findings in the other studies, is still low. Rezaei and Katz went one step further and made suggestions for improvements to the CSA by addressing the limitations of the commercial version. Their suggestions include the following:

- Including the ability to analyse the response status and reaction time for each item for each individual by getting access to the results used to calculate the ratios.
- The items in the Verbaliser-Imager subtext use differences in colour to determine the difference between Imagery and Verbalisers. The rationale is that the Imager will respond faster to the colour items. They propose that the use of colour alone as an indicator for Imagery is too simplistic because there are other dimensions of an image that are also important such as size, shape and dimensions.
- Increasing the number of questions for each category.
- Revisiting the use of red to indicate correct and blue to indicate incorrect responses because these colours contradict the convention in force in most Western cultures.
- Finding less controversial pairs because some of the pairs used are very controversial and culturally dependent.

The work of Peterson, Deary and Austin (2003a, 2003b, 2005a, 2005b) and Peterson and Deary (2006), focused on developing new versions of cognitive styles tests that could replace Riding's CSA, was first published in the literature in 2003. These authors describe three new tests that are used together to explore the Wholistic-Analytic and Verbal-Visual dimensions of cognitive style:

- Verbal Imagery Cognitive Style (VICS)
- Extended Cognitive Style Analysis – Wholistic–Analytic test (Extended CSA-WA)
- Wholistic-Analytic Inspection Time test (WA-IT)

In the VICS test, the questions that assess the verbal dimension ask whether two items are human-made or natural, and these are presented in both word and image form. Questions assessing the imagery dimension ask which of the two objects is the larger in size, and these are presented in both word and image form. There are a total of 232 stimuli in the test and it takes approximately 20 minutes to complete.

The Extended CSA-WA, which consists of 80 items (40 for each dimension), takes 15 minutes to complete. The two categories of items ask participants to compare how similar two objects are (Wholists are expected to respond faster than Analyticians), and indicate whether or not one object is part of the second object (Analyticians are expected to respond faster than Wholists).

A criticism of current W-A style tests is that they tell us little about the actual information processing that occurs. One approach to studying lower order information processing is to use inspection time.

This is the basis of the Wholistic-Analytic Inspection Time test (WA-IT), and the hypothesis is that individuals with a particular style preference will process stimuli that match this preference more quickly than those who do not have this match between style preference and stimulus. Stimuli were categorised as either global or local and were distributed across 320 trials, presented in four equal blocks of 80 trials each. It was predicted that Wholists would be more accurate in identifying the global stimuli and Analyticians would be more accurate with the local stimuli. The test takes approximately 45 minutes to complete. The style preference is determined by calculating the global-local accuracy ratio. Peterson and Deary's study (2006) found no correlation between the WA-IT and Extended CSA-WA tests. They suggest that the reason for this could be found in the fact that the two tests are measuring different kinds of task.

The VICS and Extended CSA-WA show promising results, and appear to be more reliable than Riding's CSA ($r = 0.2$) (Peterson & Deary, 2006; Peterson et al., 2005a; Peterson et al., 2003a). These new tests are not yet commercially available, and Peterson et al. (2005a) have documented the work that still needs to be performed on the instrument, including tests of validity. I communicated several times with Peterson in late 2005 and early 2006 about the possibility of using the VICS and Extended CSA-WA for my study, but I failed to convince her in time to let me see the actual tests or to give me permission to use these tests.

2.10.2.2 Allinson and Hayes' CSI

Another instrument for the assessment of cognitive style is the Cognitive Styles Index (CSI) developed by Allinson and Hayes (Allinson & Hayes, 1996). The aim of their study was to produce a psychometrically sound instrument that would be suitable for investigating the intuition-analysis dimension of cognitive style in large-scale organisational studies (and one that could specifically be used to assess managers and professionals). Intuition is a characteristic of right brain orientation and refers to the kind of immediate judgment that is based on feeling and the adoption of a global perspective. Analysis is a property of left brain thinking and refers to judgments that are based on mental reasoning and a focus on detail. The final instrument had 38 items. It is a self-report instrument that uses a trichotomous scale: true, false and uncertain. Tests for internal consistency yielded Cronbach's alphas ranging from 0.94 to 0.91 across the seven samples used. The test-retest coefficient for the one sample ($n = 30$, with a period of one month separating the two tests) was 0.90 ($p < .001$). The authors called for the study to be replicated and extended and requested test-retest studies to be conducted over a period of time.

This call was taken up by Sadler-Smith, Spicer and Tsang (2000). Their study ($n=1050$) sampled employees from a large range of occupational and professional contexts and included some graduate students. Sadler-Smith et al. found that the internal reliability was satisfactory across all sub-samples (0.84 - 0.90). They also found that it compared well with the values cited by the developers. When they performed factor analysis, they obtained results that were similar to those of Allinson and Hayes. They also examined these styles in relation to other constructs, and used Riding's CSA because it

had been demonstrated to possess both construct and concurrent validity. Their findings, however, indicated that there was no statistically significant correlation between the CSI and the Verbaliser-Imager or Wholistic-Analytic dimension of Riding's CSA. They suggested that the styles measured by the CSI and the CSA are orthogonal, and their suggestions for future research include a recommendation to consider the construct validity of the CSI. Both Sadler-Smith et al. and Riding feel that it is a matter of the greatest importance to establish the validity of the instrument. These researchers feel that this is crucial to the empirical and theoretical elaboration of the cognitive style construct, otherwise the cognitive style field will suffer the same fate as the learning style field, that is, it will become so diluted and confused by the use of terms that at face value have the same label, but have different meanings to different groups of researchers and may even describe different constructs.

Hodgkinson and Sadler-Smith (2003a) continued to research the construct intuition-analysis which underlies the CSI. By making use of a different methodology, they produced evidence that intuition and analysis should be treated as two separate dimensions. Hayes, Allinson, Hudson and Keasey (2003) responded to this critique with a counterargument that the theoretical argument forwarded by Hodgkinson and Sadler-Smith was insufficiently robust, and they proceeded to defend the methodology they had used for factor analysis. In a final response to this critique, Hodgkinson and Sadler-Smith (2003b) continued to criticise the methodology used by Hayes et al. and defend the methodology that they had used. While such debate is valuable and necessary, I could find no further evidence in the literature that this debate about the proper methodologies that should be used to establish validity had been furthered or decisively confirmed or rejected by any other group of researchers. The literature also produces no evidence that the nature of this construct has been clarified.

When it came to comparing the two instruments, I elected to use Riding's CSA because the literature indicated that the validity of the style construct appears to be more robust than the constructs measured by the CSI.

2.10.2.3 Determining the Verbaliser –Imager dimension of style

The Verbaliser-Visualiser or Verbaliser-Imager dimension of style has remained problematic, and there is still debate about the validity of this construct and how to create reliable instruments to test this style dimension. Several researchers have devoted their attention solely to this dimension.

Ong and Milech (2001) developed the Style of Processing Scale (SOP) – an adaptation of the work of Richardson – to assess verbal-visual cognitive style. Their findings were that the scale had good internal reliability ($r=.78$) and that it yielded good test-retest reliability ($r = .81$) after six months. Although they did not report on the validity of the scale in this study, they compared this scale with Riding's CSA in subsequent work in which they used the SOP (Ong & Milech, 2004), and they found low correlations between the verbal-visual sub-scales of the two tests. These researchers reported

that while the SOP has demonstrated reliability, there is little support for its validity. They are of the opinion that the only recent work in verbal-visual cognitive style that could be considered reasonably comprehensive is the work of Riding.

Other work that utilised self-report measures was carried out by Mayer and Massa (2003) at the University of California in Santa Barbara. They presented a battery of 14 cognitive measures related to the Verbaliser-Visualiser dimension to a sample of 95 students. These measures all crossed the domains of cognitive ability, cognitive style and learning preference. While some of the measures used were already in existence, others were adapted from existing tests. Mayer and Massa specifically created seven measures for the study. This new instrument included two cognitive style measures, namely, the Santa Barbara Learning Style Questionnaire (6 items) and the Verbal-Visual Learning Style Rating (1 item). The aim of the work of these researchers was twofold: firstly, to determine theoretically whether the Visualiser-Verbaliser dimension is unitary or multifaceted, and, secondly, to produce valid and economical measures of style and ability as well as valid and behavioural measures of learning preference. Correlational analysis of the cognitive style measures demonstrated that while three of the four cognitive style measures (Verbalizer-Visualiser Questionnaire, Santa Barbara Learning Style Questionnaire and Verbal-Visual Learning Style Rating) correlated highly with each other, the CSA did not correlate significantly with the other measures. Factor analysis revealed four factors with eigenvalues greater than 1. These four factors were labelled cognitive style, general achievement, learning preference and spatial ability respectively. The Santa Barbara Learning Style Questionnaire and the Verbal-Visual Learning Style Rating, both new measures, loaded strongly onto the cognitive style factor, as did the Verbalizer-Visualiser Questionnaire (VVQ) of Richardson and the Learning Scenario Questionnaire (from the Learning preference group of measures). The CSA did not load strongly onto any of the factors, although it had its strongest loading (.175) on the cognitive style factor. The Cronbach's alpha coefficient for the Santa Barbara Learning Style Questionnaire was .76. It was not reported for the Verbal-Visual Learning Style Rating. Mayer and Massa concluded that the Visualizer-Verbalizer dimension is multifaceted and that it covers the areas of cognitive style, cognitive ability and learning preference. The findings in this study that are related to the Verbal-Imager dimension of Riding's cognitive style model cast doubt on the validity of this measure in the CSA, despite the claims by Riding that the CSA is a valid test. Mayer and Massa observed that, on a practical level, this study provides good evidence for the fact that one or two self-ratings can yield effective and economical measures of spatial ability and learning style.

Massa and Mayer (2005) raise several concerns about the validity of the Verbal-Imager subtest of the Cognitive Styles Analysis. In their summary of the evidence, they state that this subtest of the CSA lacks face, construct and predictive validity. Their critique of face validity is that the test does not appear to measure the test-takers primary mode of processing or representing information in thought. No information is requested about how the information is processed. Massa and Mayer do, however, concede that the ability of any test to ascertain what thought modality a person is using is must be interpreted very cautiously because participants might not know whether they are using their verbal or

visual information processing. The work by Mayer and Massa (2003) that was discussed in the previous paragraph, provides evidence that the Verbaliser-Imager dimension of the CSA does not have construct validity. Massa and Mayer also report on a study in which they compared learning behaviour with scores on a test of style in order to determine the predictive validity of the instrument. The process measure of the Verbaliser-Visualiser cognitive style was the number of times a learner selected a pictorial help screen rather than a verbal help screen. They used Riding's CSA and their own Verbal-Visual Learning Style Rating to determine the Verbaliser-Imager-style of the participant. Riding's CSA did not correlate significantly with any of the process measures. By contrast, the Verbal-Visual Learning Style Rating had strong, significant correlations with the process measures. The implication for my study, which will use Riding's CSA is that results must still be interpreted cautiously due to the limitations of the instrument.

2.10.2.4 Other measures of cognitive style

Another question that has interested researchers is whether or not there is a cerebral basis for differences in cognitive style for the Field Dependence-Field Independence style (Tinajero & Paramo, 1993) and Riding's style model (Riding, Glass & Douglas, 1993; Riding & Glass, 1997). I will discuss the work of Riding, and the study reported in 1997 (Riding & Glass, 1997) in particular, because a review of the methodology and results of this study suggest that there might be a link between the tasks which had different processing loads, and cognitive style. The researchers did not examine this question any further because it was beyond the scope of their study at that time.

On the basis of existing knowledge that the right hemisphere is associated with the location of visuospatial function and the left with the verbal function, Riding, Glass and Douglas (1993) proposed that Verbalisers, who translate pictorial information into words or semantic representations, will present predominantly with left hemisphere activity, and that Imagers, who represent semantic information in mental pictures whenever possible, will present with predominantly right hemisphere activity. Since suppression of the rhythm is indicative of cortex activity in the interpretation of EEG alpha rhythm, they predicted that EEG alpha rhythms would be suppressed in the left hemisphere when Verbalisers were given selected cognitive tasks, and that right hemisphere EEG alpha suppression would be evident when Imagers were given the same cognitive tasks. The location for Wholistic-Analytic processing in the brain is less clear.

Riding and Glass (1997) explore this model in their study. They gave tasks of increasing information processing load to their subjects and then measured the alpha power for each task at different locations. Analytics demonstrated lower alpha power than Wholists over all the tasks and in all locations. This difference was more pronounced for the posterior locations. Riding and Glass also detected a style-hemisphere effect with the Verbal-Imagery dimension. This effect was that Verbalisers showed relatively more suppression on the left posterior temporal location in comparison to right, while Imagers showed relatively more suppression on the right posterior temporal location in comparison to the left (Riding & Glass, 1997).

The suggestion of a link between cognitive style and load is seen in the finding that, for the Analytics and Imagers, processing was fairly constant irrespective of task load, while for the Wholists and Verbalisers the alpha power decreased as the task processing load increased. In fact, the general pattern for these two styles was that the alpha power decreased with increasing information – processing load, up to a point, where it then increased slightly, suggesting a resting state. Could this indicate a working memory overload with which they could not cope – with the result that they stopped processing the information altogether? Does this suggest that the Analytic style can cope with higher cognitive loads than can the Wholistic style, and that Wholists will therefore perform more poorly as the load increases while, for the Analytics, there is little or no difference? Riding et. al.(1997) were of the opinion that these findings required further research, as they had implications for the use of EEG power output as an index of information-processing load.

Genovese (2005) tested three instruments that were designed to measure hemispheric cognitive style. The popular belief is that the verbal-analytic cognitive style is related to the functions of the left-brain hemisphere and that the visual-holistic cognitive style is related to the right. The validity of instruments to measure hemispheric cognitive style is determined by using one of the following two methods:

- The researcher looks for correlations between the scores on the instrument and biophysical measures such as EEG activity.
- The researcher looks for the ability of the instrument to predict certain individual choices which represent hemispheric differences.

An examination of the correlations between the scores of three instruments and the ability of the instruments to predict teaching licensure area, revealed that only two of the instruments correlated with each other and with teaching licensure area. The factor analysis also provided evidence for the existence of the two separate Wholistic-Analytic and Verbal-Visual dimensions proposed by Riding and Rayner (1998).

Although one can detect some progress with regard to the assessment of style preferences, the observation that Hodgkinson and Sadler-Smith made as far back as 2003 still rings true to a large extent today:

Basic lack of agreement over nomenclature and few reliable and valid instruments suitable for assessing cognitive style in applied settings threatens the viability of this construct. But if cognitive style was a unitary construct then it would be necessary that the majority of instruments developed to measure cognitive style are inter-correlated with one another. This evidence has not been forthcoming.

(Hodgkinson & Sadler-Smith, 2003).

2.10.3 Cognitive style and learner characteristics

I turn now to a discussion of cognitive style and learner characteristics. The literature addresses the characteristics of gender, age and intelligence. Since I have included the characteristics of gender and age in my study, a brief review of the literature is therefore relevant.

In 1995 Riding, Burton, Rees and Sharratt concluded that there appear to be no overall gender differences with respect to cognitive style. Any differences are usually small and non-significant on both dimensions ($p < 0.05$). While this was confirmed in three studies that I reviewed (Abouserie & Moss, 1992; Evans, 2004; Riding & Read, 1996), I also found several studies that indicated gender differences. For the Wholistic-Analytic style, Riding and Staley (1998) found that for the Wholistic-Analytic dimension there was a significant gender difference, with females having a slightly lower ratio than the males. Riding and Agrell (1997) reported similar results. Riding and Grimley (1999) also found good evidence for a gender effect and style. Their literature review reports on studies from the 1980s in which gender effects in information processing were observed. The materials used in these early studies were paper-based. The conclusion they reached in their literature review was that while males processed more quickly but at a more superficial level, females processed more thoroughly.

Riding and Grimley (1999) used primary school children to explore the relationships between cognitive style and performance and gender. Their subject material included multimedia learning materials, and they assessed performance by using a multiple choice type test and assessment at the level of recall. They found an interesting gender effect when they compared performance between the three formats of the multimedia lesson: pictures plus sound (PS), pictures plus text (PT), and pictures plus sound plus text (PTS). For the PS and PT condition, the gender effect was determined by whether the style groups were unitary (AV and WI) or complimentary (AI and WV). In the complementary groups, males perform better with the PS version than with the PT version and females performed better with the PT version. The results were reversed for the unitary style groups. In these, the males performed better in the PT version and females performed better in the PS version. Performance was the best for all style and gender groups when the multimedia was presented with picture, sound and text. This is an interesting finding if one looks at the results in the light of the both the cognitive load theory and the cognitive theory of multimedia. There is a substantial body of research into what has been called 'the modality principle' (Mayer, 2003; Mayer, Dow & Mayer, 2003; Moreno et al., 2001; Moreno & Mayer, 1999) or 'the modality effect' (Ginns, 2005; Kalyuga et al. 2000; Tabbers et al., 2004). The modality principle states that when one uses multimedia instruction that includes text and images, the words should be presented in the form of narration (that makes use of the auditory channel of processing) rather than in the form of on-screen text. Leahy et al. (2003) explored the conditions under which auditory presentation may be effective or ineffective. Their experiment led them to the conclusion that when the intrinsic load of material is high, a dual mode of presentation (audio and text in that experiment) produces better results than a text-only presentation. While Riding and Grimley (1999) did not describe the exact instructional

design of the multimedia used in their investigation, it is possible that the modality effect was responsible for the fact that the participants performed better with the picture, sound and text versions than with the picture and text or text-only versions. Their finding suggests that cognitive load might have played a role in their findings. It could have been that the picture-text version, which only used one channel for processing information, had a higher load than the picture-sound and picture-text-sound versions which used two channels for processing – thus dividing the cognitive load.

Abouserie and Moss (1992) investigated the relationships between cognitive style, gender, attitude toward computer-assisted learning (CAL), and academic achievement among university students who were taking a course in Physiology. This is one of the **few** studies that made use of the same context (namely, medical students taking a Physiology course) as the one I have used in my study. Cognitive style was measured by means of the Field Dependent – Field Independent (FD-FI) classification. While the participants had a positive attitude toward CAL, they indicated that they were not prepared to rely solely on CAL. Gender and style were not significant factors in the relationships explored in this study.

By making use of a sample of 119 twelve- to thirteen-year-old pupils from an urban school, Riding and Pearson (1995) examined the relationship between intelligence and cognitive style. Intelligence was determined by means of the British Abilities Scales Short-Form and the test gave an IQ score for each participant, and cognitive style was measured by means of Riding's CSA. The researchers detected no significant relationships between cognitive style and intelligence, either in terms of overall IQ or the individual sub-tests. Since all the coefficients were low, the inference was that intelligence and style are independent. Both style and intelligence will effect performance on a given task. The difference is that as intelligence increases, so does performance, while style exerts either a positive or negative effect depending on the nature of the task. Riding and Agrell (1997) also came to the conclusion that cognitive style and cognitive ability are independent of one another.

Peterson, Deary and Austin (2005b) also looked into the question of whether intelligence is related to cognitive style. By making use of a sample of 100 university students with a mean age of 20 years, they conducted a study using 2 performance-based cognitive style tests, 8 tests of mental ability and 3 personality tests. Out of 24 possible significant correlations between style and intelligence, only three tests of ability correlated with any of the style measures. These correlations were not higher than $r = .27$, which is lower than the correlation determined for the power of the study. Similar findings were obtained for the personality measures. Peterson and Deary concluded that the correlations that occurred could most probably be attributed to Type 1 errors because the findings were not consistent across the different test sessions, and because it appeared as though style and intelligence were independent constructs.

2.10.4 Cognitive style and behaviour

In this section of Chapter 2, I will review research that investigated the relationship between cognitive style and behaviour, where 'behaviour' included attitudes, perceptions, anxiety, work preferences or orientation, teaching style and navigation in multimedia (including the web). The focus of the research into learner behaviour has served the purpose of describing the different styles in more detail and it has become possible to create profiles of particular styles.

The study of Abouserie and Moss (1992) also looked at attitudes towards computer assisted learning. In response to the question of whether or not they would be prepared to rely entirely on computer tutorials, field dependent (FD) individuals were slightly more favourable towards this possibility than were field independent (FI) students. The researchers suggested that this preference might have been motivated by the structure and step-by-step approach of the tutorial with many examples and exercises. Field dependent learners, who prefer not to impose their own structure on learning material, may have found this option more to their liking than field independent students, who are able to, and prefer to, impose their own structure on their learning materials (Witkin, 1977). But since this finding was only relevant to one of the six items in the instrument, it is difficult to interpret because there were no differences between the field dependent and field independent students when the total scores for the instrument were compared. Cognitive style also had no effect on the achievement of the students in the Physiology course. Attitudes can influence motivation to learn and the willingness to use new technologies such as multimedia instruction. Paas et al. (2005) propose that meaningful learning can only commence if the learner is also motivated to learn and is willing to invest mental effort in processing the instruction. Motivation speaks to germane cognitive load and is one of the newer directions for cognitive load research.

I now turn to an examination of what the research says about what learners with different styles actually do, or what they prefer, when they use multimedia learning materials (which includes using web-based and hypertext applications).

Graff (2005) researched the different web browsing strategies used by older and younger participants, on one hand, and individuals who displayed Verbaliser and Imager cognitive styles, on the other hand. He developed different web architectures: a simpler hierarchical architecture that allowed users to browse up and down the hierarchy, and a more complex architecture that allowed users to browse hierarchically and laterally across the different topics. Graff hypothesised that Imagers would visit a greater proportion of pages in a complex hierarchy as they strove to obtain a big picture view, and that they would also tend to browse high in the hierarchy rather than dig deep in their efforts to locate specific information. His findings were that Verbalisers and Bimodals visited more pages and a higher proportion of pages in the simple hierarchy, while Imagers visited more pages and a higher proportion of pages in the relational condition. This is consistent with a Verbaliser's need to obtain detail by drilling down into the hierarchy for each topic, and with the need

of Imagers to get the big picture by browsing more laterally in the structure. My critique of this study is that participants were given 10 minutes to browse through 64 pages of content. Although they were told that they would be answering questions about the content at the end of the session, this was not clarified in the published article. Since ten minutes seems to be a very short time for such a task, it is valid to ask whether this would not impose a high cognitive load on the learner. Since no performance results were reported in the assessment, it is not possible to decide whether or not cognitive load might have played a role in this instance. What *is* relevant here is whether a similar pattern of navigation will be observed in my study in which one might expect Imagers to move quickly through the program to get the big picture before coming back to each screen to study the content in more detail, and in which Verbalisers might be expected to work systematically through the lesson only once and to spend the amount of time on each screen that would allow them to do this. Screen logs of this behaviour were recorded so that it would be possible to determine this information.

Calcaterra, Antonietti and Underwood (2005) used a different style dimension and reported different effects of navigational style in hypermedia environments. They examined the influence of cognitive style, spatial orientation and computer expertise on hypertext navigational patterns and learning outcomes and used a Wholistic–Sequential style dimension rather than a Verbaliser-Visualiser dimension. They found that hypermedia navigational behaviour was linked to computer skills rather than to cognitive style and that learning outcomes were unaffected by cognitive style or by computer skills. Performance on the learning outcomes was positively affected by specific search patterns: participants who re-visited the hypermedia sections and who visited the overview sections in the early stages of hypermedia browsing obtained higher scores. The total amount of time spent on the content did not affect performance.

Self-perception in general has many dimensions. These are efficacy, ability, concept and esteem. While the influence of self-perception on performance will sometimes be negligible, it will at other times exert a major influence. Perception about oneself as a learner is influenced by motivation, interest in a subject and performance. Riding and Staley (1998) explored the relationship between self-perception and cognitive style with a sample of first-year Business Studies university students. The main focus of the study was on style and the differences between perception and performance. The role of self-perception for a learner is in the regulation of learning. Riding and Staley propose that the level of self-regulation that learners adopt, and hence their motivation, will be influenced both by the learning experience itself and by the influence of performance on outcomes. Learning experience includes whether or not learners find the subject matter easy, the attractiveness or otherwise of the presentation, and the extent to which the material makes sense to them. When one looks at their proposal from a cognitive load frame of reference, the extent to which learning material is easy and extent to which it makes sense to the learners speaks to the issue of intrinsic cognitive load, while the attractiveness of any particular presentation is addressed by extraneous cognitive load. If both the learning experience and performance are likely to be influenced by the cognitive style characteristics of the individual learner, then the question 'And what role or influence does cognitive load have?' is one that may also require investigation. The results of this study indicated that performance in all the

subjects included in the study was related to the nature of the tasks within each subject area: thus, in those cases where the acquisition of the big picture was necessary for a proper grasp of the subject (as is the case in Management), Wholistic learners performed better than Analytic learners. One might therefore hypothesise that if there is a mismatch between subject and style, learners might report that they needed to invest more mental effort in order to understand the work, so indicating a higher cognitive load. What type of load is being influenced here? It is most likely to be germane load, which has to date not yet been directly measured.

Chen, Ghinea and Macredie (2006) examined the relationship between cognitive style and the learner's quality of perception of multimedia. Their definition of the quality of perception included an enjoyment and understanding of content in selected video clips which consisted of varying combinations of audio, video and text. They have published at least two articles on this topic. In their 2003 publication, they used the field independent- Field Dependent style to explore the relationship (Ghinea & Chen, 2003). In their 2006 publication, they made use of the Verbaliser-Imager dimension of style as measured with Riding's CSA (Chen, Ghinea & Macredie, 2006). Both these studies included the nature of the content and the informational load as parameters in the study. In both studies it was observed that participants experienced difficulty in concentrating on video clips that were very dynamic, that is they included use of video, audio and text, irrespective of their style. The 2006 study looked at whether the cognitive style of the participant (Verbaliser or Imager) influenced the level of information being assimilated from the two sources, video or audio. They found that cognitive style was not a significant factor. But they did find that in those cases where the information source was text, Verbalisers obtained a statistically significant higher score than did other styles. The explanation of the design used and the subsequent findings would seem to indicate that cognitive load did in fact play a role in the design of this experiment, although this was neither addressed nor investigated by the researchers.

Evans (2004) investigated the nature of the relationship between a teacher's cognitive style, as measured by Riding's CSA, and his/her teaching style. This study used a strategy that had already been used in early cognitive style research (Riding & Agrell, 1997; Riding & Grimley, 1999; Riding, Grimley, Dahraei & Banner, 2003), namely, the combination of the two style dimensions to give four style groups. Riding and Rayner (1998) have suggested that the Analytic-Verbaliser and Wholistic-Imager combinations are unitary, and that the Analytic-Imager and Wholistic-Verbaliser style combinations are complementary.

When remarking on the Analytic-Verbaliser style, Evans makes the following very interesting statement:

The unitary aspect of this cognitive style also adds another requirement: this style can deal with large amounts of information.

(Evans, 2004, pg 512).

Evans, however, neither clarifies, describes, defines nor further discusses the constructs 'deal with' and 'large amounts'. Does 'large amounts of information' also imply a possible cognitive load effect? What assumptions can be made about this information? My study will present the same content with a different cognitive load to learners whose cognitive style will be measured. This will be the first systematic attempt to explore this statement by Evans.

2.10.5 Cognitive style and instructional design practice

Cognitive style research also informs instructional design practice. Several studies have considered the impact of different instructional strategies on cognitive style and learning performance.

Riding and Sadler-Smith (1992) compared performance on two versions of computer-presented instructional material about central heating systems with students between the ages of 14 and 19. Their study manipulated the instruction along four dimensions – structure (large versus small step), advance organiser (absent or present), verbal emphasis (high versus low), and diagram type (abstract versus pictorial) – in order to determine which form best suited which cognitive style combination. The organiser version included an introduction to the entire lesson and an overview and summary at the beginning and at the end of each of the five sections of the lesson respectively. There was also an overall summary at the end of the five topics. The best overall learning performance was for the version that included an advance organiser. The recall performance for Analytic-Verbalisers and Wholist-Imagers was slightly better when compared to that of Wholist-Verbalisers and Analytic-Imagers. While Analytic-Verbalisers battled to get a big picture view of the material and Wholist-Imagers found it more difficult to analyse the content, the inclusion of an overview and summaries in the instructional design of the material provided them with the necessary support and so improved their performance. The form of the organiser did not appear to help the Wholist-Verbaliser and Analytic-Imager learners. The additional material seemed to reduce performance for the other two styles. Riding and Sadler explained that because the Wholist-Verbaliser and Analytic-Imager learners were able to get a big picture view and analyse information on their own, the additional material was redundant and probably depressed performance. This explanation fits the redundancy effect described by cognitive load theory exactly (Chandler & Sweller, 1991; Chandler & Sweller, 1996; Kalyuga, Ayres, Chandler & Sweller, 2004; Kester & Paas, 2005).

Graff (2003b) made use of a web-based lesson to investigate the influence of segmentation of information and the extent to which an overview facilitated learning. His findings suggested that cognitive style and segmentation have an effect on performance, and that the Wholist-Analytic dimension of cognitive style is the style that determines the degree of success that individuals have when they attempt to learn from web-based systems. Segmentation may exacerbate the Analytics tendency to see information in parts because it encourages them to focus more strongly on the parts at the expense of seeing the whole picture, and this – in the long run – may be detrimental to learning. Analytics should therefore be able to learn more efficiently from material that is not heavily segmented. For Wholists, segmentation would make very little difference because of their inherent

ability to see the whole. But the nature of the content is also important. Which is more important: the big picture or are the parts? Because of its inherent ability to link information that is conceptually related, the web has the potential to present instructional information more effectively than traditional linear methods. But one has to take into account the disadvantages of fragmentation and the lack of discourse cues (indicators of how information is presented). The findings also suggested that the provision of an overview had little effect on learning performance. This leads me to return to the question: 'Is the big picture more important or are the parts more important?' If the learning of the parts is more important, the provision of an overview would tend to benefit Wholists because they strive to see the big picture. It would make little difference to Analytics. This angle, which considers the nature of the content, was not part of Graff's investigation. In addition to this, users of web-based systems have to multitask (navigate, read and understand). All these activities increase the cognitive load. Graff's study contains a very brief reference to the fact that there might be a relationship between cognitive style and cognitive load, where the load is affected by the amount of extraneous load, and which includes tasks such as navigation, rather than information processing at a cognitive level. Graff did indicate that it is possible for an instructional system with different subject information and also (possibly) a system incorporating a multimedia component, to yield a different finding. Multimedia brings with it a greater potential for cognitive overload. This suggests another reason why it might be helpful to explore the relationship between cognitive load and style in the context of web-based and other multimedia learning.

2.10.6 Cognitive style and achievement

Although most of the studies reviewed in the sub-sections above do include achievement, the inclusion was more concerned in each case with learning performance in a posttest designed for the particular study. The achievement that I will address in this section looks at achievement as it is reflected in grade scores and final exit level examinations across several subjects in the curriculum.

Riding and Caine (1993) looked at how the habitual ways of representing and structuring information affect General Certificate of Secondary Education (GCSE) performance in Mathematics, English Language and French in a sample of 16-year-old pupils. When they looked at overall performance they found that learners who returned an Intermediate (on the Wholistic-Analytic dimension) or Bimodal (on the Verbaliser-Imager dimension) style performed the best. The inference they made from this was that because these style groups avoid the limitations imposed by an extreme style, they are free to use the most appropriate processing across the style dimensions as and when a specific task requires them to do so. When they analysed performance in individual subjects, the authors note that there were two factors that influenced performance: the extent to which the subject required a whole or a part view and the degree to which this requirement is matched by an individual's style.

Although the study revealed interactions between the two style dimensions for the different groups, the researchers did not discuss the following points:

- their analysis of the task requirement for each subject included in the study
- their expectations for the different style groups in the light of a task requirement per subject.

This lack of information makes it difficult to interpret and explain the findings objectively without making one's own assumptions about the context of the study. The essence of their findings seems to be limited to the fact that the nature of content seems to play a role in style differences, as proposed by Riding and Smith (1992). This point also needs to be taken into consideration in any interpretation of the results of my study.

Riding and Pearson (1995), using a sample of 119 12 – 13 year old pupils from an urban school, examined the relationship between intelligence and cognitive style. Performance in the following subjects was analysed: Mathematics, French, Science, History, Geography, English. They found a significant interaction between the Wholistic-Analytic style dimension and the subject. Wholists achieved more highly in Geography and French, Intermediates in English, History and Science and Analytics did poorly on Science, Geography and French. In this study performance was once again largely influenced by the extent to which the content requires a whole or parts view.

Riding and Agrell (1997) investigated the relationship between cognitive style, cognitive skills and school achievement in an English-speaking Canadian school environment. Like Riding and Pearson (1995), they examined a sample of students taking French, English, Mathematics, Geography and Science. The researchers point out that any investigation into educational achievement and style poses problems because of the interplay between the nature of the subject, the ways in which it has been taught, and the methods that are used to assess performance. There are usually variations within subjects themselves with regard to the type of content and the range of processing required at any given time. The researchers used Mathematics as an example of a subject in which sequential operations are required for arithmetic, abstractions are required for algebra, and spatial representations are required for geometry. Because variables such as these are not easy to control in educational research, they did not expect to find (and did not, in fact, find) any noticeable differences between performances on subjects with respect to style. They stated their conclusions as follows:

- Analytic-Verbalisers appear naturally suited to all the subjects considered in this study, probably because they all require an element of verbalisation together with an ability to analyse.
- By contrast, Analytic-Imagers appeared to be least suited to these academic subjects, perhaps because they lack fluent verbalisation ability.

The implication of this is that instructional strategies need to be put in place to assist those learners whose style does not naturally 'fit' the nature of the content. Riding and Agrell's study leads us to

believe in this regard that while the more intelligent and adaptable learners will probably be able to solve this problem for themselves, the less and adaptable intelligent ones will need and should be trained in the art of developing strategies. In those cases where, by contrast, the styles are naturally appropriate, there will be little incentive to develop strategies because learners will usually be able to cope reasonably well.

The conclusions provided by the research discussed thus far emphasise how extremely important it is for instructional designers to undertake thorough content and task analyses during the development of learning material. In cases where the cognitive load of the material is known and the style of the learner is also known, I must ask the following question: Is it possible to predict how learners with different styles will deal with the cognitive load of the particular content?

In the process of picking up on the suggestion that it is desirable to put strategies in place to assist learners whose style is not naturally suited to their learning context, I looked briefly at the work done by Triantafyllou, Pomportsis, Demetriadis and Georgiadou (2004, 2003) in the area of using adaptive hypermedia systems. In 2004 these researchers published the results of a study that looked at whether adaptive hypermedia systems that accommodate cognitive styles (the Field Independent/Field Dependent style group) could be beneficial for observed learning outcomes. Since they found no significant interaction between cognitive style and learner achievement, they concluded that cognitive style alone was not the only factor to impact on learner performance. The study did, however, demonstrate that the difference in the mean scores between the field independent and field dependent group was proportionally smaller for the posttest than for the pretest. It was the adaptive system that had made it possible for the field dependent learner to close the performance gap and start achieving at a level that was almost identical to the field independent learner group.

The work of Riding, Grimley, Dahraei and Banner (2003), which considered the relationship between working memory, cognitive style and gender on overall learning behaviour and performance in 10 different school subjects, addressed the hypothesis that cognitive load might be a consideration when looking at methods of improving performance. They assessed working memory efficiency by using an instrument that was developed by Riding – the Information Processing Index. They used the tutors to assess this learning behaviour on a 6-point rating scale. The authors' findings were that there was no relationship between the independent variables Wholistic-Analytic style, Verbaliser-Imager style, gender and working memory capacity. Research into the cognitive style stream has already established the independence of the two style dimensions (see Section 2.7). A study by Colom, Flores-Mendoza and Rebello (2003) has confirmed that working memory is one general cognitive resource and that it is strongly related to intelligence. If cognitive style and intelligence are not related, it follows that there should also be no relationship between the two style dimensions and working memory capacity. It therefore comes as no surprise that there was no relationship between the independent variables.

An analysis of variance of gender, Wholistic-Analytic style and memory with the dependent variable showed a significant effect for memory and an interaction between Wholistic-Analytic style and memory. In this analysis, working memory capacity produced little effect for Wholists but a large effect for Analyticians. While the Analytic learner performed well in those cases where sufficient working memory capacity was available, it seemed to be the case that working memory capacity made little difference for the Wholistic learner. A similar analysis using the Verbaliser-Imager style produced similar results: while working memory capacity produced little effect for Imagers, it produced a large effect for Verbalisers who, if there was high working memory capacity, performed well. My problem with this finding is that learning behaviour as defined for this study did not adequately reflect the profile of either the Wholistic-Analytic or Verbaliser-Imager style dimension, which has been described well in the literature (Riding & Rayner, 1998). Only five items were used for learning behaviour and five for conduct behaviour, and items regarded as learning behaviour included 'is attentive and has interest in school work', 'good learning organisation', 'is an effective communicator', 'works efficiently in a group' and 'seeks help when necessary'. I think this results in a very narrow view of learning behaviour. By using subjective ratings of performance across 10 subjects as the dependent variable, similar analyses of variance were carried out for the interactions between Wholistic-Analytic style and working memory. Working memory capacity in general exerts the biggest influence on the performance of learners with Analytic and Verbaliser styles. The researchers suggested that in those cases where the working memory capacity of the Analytic and Verbaliser is low, their performance could be improved by reducing the load on working memory or enhancing their effective working memory capacity. Where instructional designers are developing a program that uses style as a parameter in an adaptive learning environment, the analytic learner should preferably be routed to the version/strategies with the lowest cognitive load. Their study called for further investigation into the interaction between cognitive style and working memory capacity, and, by implication, cognitive load.

2.10.7 In summary

This section has examined cognitive style and multimedia learning in terms of the following six main themes:

- Contributions to the theory
- Measurement of cognitive style
- Cognitive style and learner characteristics – gender, age, personality, intelligence
- Cognitive style and behaviour – attitudes, work orientation
- Cognitive style and instructional design practice
- Cognitive style and academic achievement

For each of the above themes, I discussed and critiqued the most relevant research. This included a determination of whether the findings of the study could be explained from a cognitive load frame of reference.

The research has looked at adding an audio component to the assessment of cognitive style, with the proposal that Riding's cognitive style model be redefined to take an auditory component into account. Researchers have also looked at revising the Visualizer-Verbaliser dimension of style. This work has explored the possibility that there are two groups of Visualizers, namely those who had high and low spatial ability respectively.

Two instruments used to measure cognitive style were considered in detail: Riding's CSA and the CSI of Allinson and Hayes. Riding's CSA is criticised for its poor reliability. There has also been criticism of the validity of the Verbaliser-Imager dimension. In the absence of other reliable and valid instruments to measure cognitive style, the CSA still continues to be widely used. Peterson and her colleagues (2005a) are busy developing and testing alternative instruments that address some of the weaknesses of Riding's CSA. These instruments are not yet available commercially. The CSI has also been reviewed, critiqued and used in several replication studies. This instrument demonstrates good reliability, but there are still problems in establishing it as a valid measure of cognitive style. Finally the review looked at the work of Massa and Mayer, which has focused on developing a series of instruments to determine Verbaliser-Imager style. These instruments appear to be promising alternatives, but need to be tested in other cultural contexts, for example South Africa.

While there are contradictory findings with regard to the relationship between cognitive style and gender, the evidence appears to favour the position that cognitive style is not related to gender. There does not appear to be any relationship between cognitive style and intelligence.

The review of the literature under the section 'Cognitive style and behaviour' was aimed at establishing the extent to which current research validates or refutes the profiles for the different styles, as explained by the style models. The styles that seem to be well researched are the Field Dependent /Field Independent style, the Verbaliser-Imager dimension of Riding's model and studies that combine the Wholistic-Analytic and Verbaliser-Imager dimensions of Riding's model. Several of the studies looked at navigational behaviour in hypermedia and web-based systems.

Instructional design practice has also been informed by cognitive styles research. The studies reviewed covered the use of instructional strategies such as advance organisers, overviews and summaries and chunking of content. The focus of this research was often directed towards establishing whether a particular strategy assisted the learner whose style was weak in the area the strategy addressed. Overview and summaries, for example, assisted Analytic-Verbalisers who find it more difficult to get a big picture view of the material and Wholist-Imagers who find it more difficult to analyse content. Segmentation, for example, may exacerbate the Analytics tendency to see information in parts because it encourages them to focus more strongly on the parts at the expense

of seeing the whole picture, while for Wholists, segmentation would make very little difference because of their inherent ability to see the whole.

The achievement of the different style groups with respect to certain subjects has also been considered in several studies. Learners who were found to be Intermediate or Bimodal in style (using Riding's model) often performed the best. This is because these style groups avoid the limitations imposed by an extreme style and are free to use the most appropriate processing across the style dimensions as and when a specific task requires them to do so. Investigations into educational achievement and style are complex because of the interplay between the nature of the subject, the ways in which it has been taught, and the methods that are used to assess performance. There are also variations within subjects themselves with regard to the type of content and the range of processing required at any given time. These are all difficult to control in experimental conditions.

Several of the authors under review did in fact hint at the possibility that cognitive load might have played a role in the findings, and they called for further investigation into this dimension. Many of the early studies were conducted with samples of school-going participants. The studies conducted in the higher education sector were carried out mostly in the subject disciplines of information technology. Since research into the health sciences is under-represented in these efforts, this study will contribute to addressing such an imbalance.

2.11 Cognitive load and multimedia learning

The following research questions are relevant to this section of the review:

- How do the participants rate the cognitive load of selected multimedia content?
- What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?
- To what extent do the presentation formats influence cognitive load?

This critical, in-depth review will also be informed and directed by the question, 'What does the research on cognitive load in multimedia learning have to say about cognitive style?' Cognitive load is regarded as an important influence on the ability of human beings to process information. The use of many different media formats – sound, animation, text, images, animated pedagogical agents, virtual reality and various combinations of these – has become widespread in education. Learners who are faced with choices are influenced in their decision-making by many factors that include their particular cognitive style. It is often tempting for designers to use more when less might, in fact, be better.

Not all multimedia research literature addresses the question of cognitive load (that research will be discussed in Section 2.11.4 and 2.11.5 of this chapter). Multimedia learning environments were also not a topic of investigation in early cognitive load research (Chandler & Sweller, 1996; Sweller & Chandler, 1994). While research into instructional strategies that use paper-based material has been

ongoing (Robinson, 2002; Verdi & Kulhavy, 2002), the proliferation of technology in education has increased multimedia learning environments to such an extent that research on this topic is now included in the cognitive load research stream.

Because there has been some merging of these streams, it is no longer easy to categorise research streams into definite focus areas such as multimedia learning or cognitive load. While I will consider available multimedia research primarily from a cognitive load perspective in this section of the literature review, I will include studies that did not investigate multimedia learning from a cognitive load perspective.

I have divided my discussion of the literature, which included a review of close on 180 articles from the cognitive load and multimedia literature, into the following five major areas:

- Instructional design issues
- Theory development and directions of research
- Measurement of cognitive load
- Learning from multiple representations
- Animations, including animated pedagogical agents

2.11.1 Instructional design issues

Cognitive load theory attempts to explain the ways in which instructional design can successfully reduce extraneous cognitive load, increase germane cognitive load and manipulate intrinsic load. Research studies using this theory have investigated many different instructional design manipulations that highlight the effects of cognitive load. I will discuss this research in terms of the effects described in the research and the principles generated for design. The following effects are described in the cognitive load literature (Sweller, van Merriënboer & Paas, 1998):

- Worked example effect
- Completion problem effect
- Goal-free effect
- Variability effect
- Guidance fading effect
- Split-attention effect
- Redundancy effect
- Modality effect

Extensive research has examined the worked-example effect, often in combination with an investigation of the completion problem effect (Gerjets, Scheiter & Catrambone, 2004, 2006; Kalyuga,

Chandler, Tuovinen & Sweller, 2001; Tuovinen & Sweller, 1999; van Gog, Paas & van Merriënboer, 2006), even to the extent that a special edition of *Learning and Instruction* was devoted to worked-example research (Paas & van Gog, 2006). Worked examples appear to be a more effective instructional technique for teaching problem-solving to novice learners, a finding that has been empirically demonstrated in the domains of algebra (Ayres, 2006a, 2006b), probability (Gerjets, Scheiter & Catrambone, 2004, 2006), programming and relay circuits (Kalyuga et al., 2001; Tuovinen & Sweller, 1999) and electrical circuits (van Gog, Paas & van Merriënboer, 2006).

Only the 'effects research' that is relevant to my study will be described in detail in this study. The research in question investigated split-attention, redundancy, modality and expertise reversal effects.

Contributions to this work emanated from both the multimedia and cognitive load research streams because these research streams were often engaged in researching the same issue simultaneously. It is interesting to note that there is considerable convergence between both the findings and the design principles that flow from the two research streams. But this research is also characterised by the contradiction and divergence of its findings, in conjunction with the convergence mentioned above.

The cognitive load research stream reveals a definite progression in the effects described, as well as interactions between these effects. What happens is that the findings with regard to one effect trigger new investigations into the interactions between that effect and another effect.

This is illustrated in Figure 2.6 below. The blue arrows indicate the direction and time line, and the red dotted arrows indicating the triggers that motivate researchers to explore the synergies between these effects..

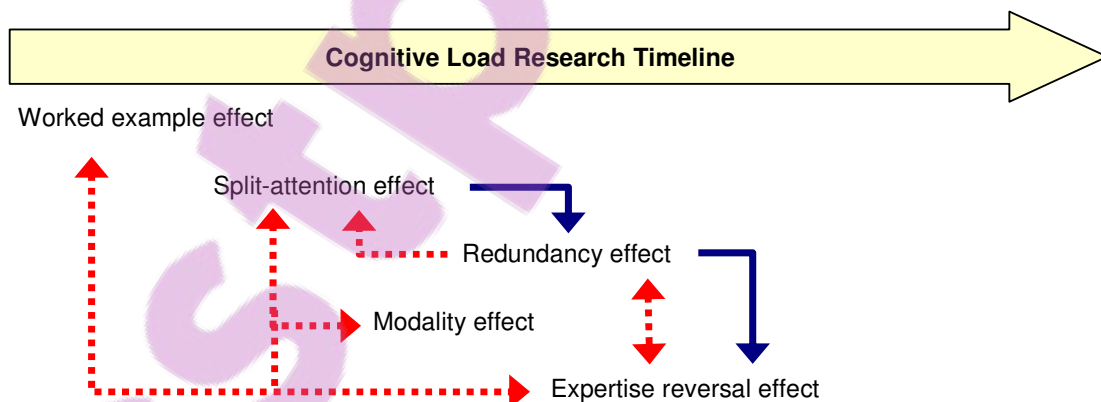


Figure 2.6: Progression for the 'effects' research within the cognitive load research stream

My review of the literature reveals that once these various effects had been described and tested in practice, it became increasingly difficult for subsequent researchers to consider each of the effects in isolation. The split-attention effect for one group and context, for example, later became a redundant

effect for another group and context (Yeung, Jin & Sweller, 1997). Another example is that researchers investigated the modality effect as a possible solution for solving the problems generated by the split-attention effect (Kalyuga, Chandler & Sweller, 1999). For the purpose of this review, however, I will discuss each effect separately and make appropriate references to the other effects where necessary.

2.11.1.1 Split attention effects

The split-attention effect became apparent during the course of investigations into the most effective ways for integrating different sources of information. If a learner needs to integrate several sources of information simultaneously in order to understand a concept or problem, and each source is unintelligible in isolation, then the most effective way of facilitating this mental integration is to present the different sources of information contiguously. If a learner must split his/her attention between sources of information that are not integrated but that need to be integrated, the need to split or divide attention may well place an unnecessary strain on the learner's limited working memory resources. Since splitting of this kind induces an extraneous cognitive load, it causes a deterioration in learning performance. Chandler and Sweller (1991) conducted six experiments that explored the integration of information, split-attention and redundancy effects from different perspectives. In their first experiment they compared two formats of instructional material: the split-source and the integrated format. The participants in their experiment used paper-based materials, together with a set of material for practical work, over a period of three months, and were tested three times during that period. While both groups improved in performance in the course of the three months, the group that was given the integrated material (which avoided the split-attention effect) achieved significantly better results. In a second experiment the groups were given material that used similar formats (split-source as opposed to an integrated format). In this case it was not necessary for the participants to integrate information from different sources so that they could understand the concepts that were being taught. In this experiment the integrated material did not enable the group that were using this material to produce better results. Chandler and Sweller (1991) therefore concluded that integrated material is not necessarily beneficial. If a learner is able to understand each source in the split-source format in isolation, then the use of a split-source format will not necessarily interfere with learning. They noted that because the learners in such cases would observe the redundancy in the sources, they would, in fact, focus on the source that was most meaningful to them.

Chandler and Sweller (1991) did not investigate style at all. Their non-integrated format used a text description and an image, and their integrated format combined text and image. It would appear that Chandler and Sweller assumed learners would be able to decide by themselves which of the two sources was best for that context. From a cognitive style perspective, however, it would have been interesting to determine where participants could be placed on the Verbaliser-Imager dimension, and then to observe their selection and use of material in order to determine whether or not cognitive style had played a role in these findings.

Chandler and Sweller then moved their research into a laboratory setting where they were able to control the variables more rigorously. The results of their 3rd and 4th experiments provided further support for the split-attention and redundant effects. While experiments 5 and 6 used material from a different domain altogether, they produced evidence that requiring the learner mentally to integrate sources of information that did not need to be integrated had a negative effect on learning performance. Experiment 6 was similar to Experiment 1 because Experiment 6 also required sources of information to be integrated before they could be understood. The evidence once again showed that learning from integrated materials produced superior learning performance. Chandler and Sweller (1991) concluded from this that instructional designers should avoid the split-attention effect wherever possible.

Mayer and Moreno (1998) demonstrated the split-attention effect while testing the predictions of the dual-processing theory. This study extended previous research in three ways: it used a computer-based multimedia environment rather than paper-based materials; it used multiple dependent measures, and it looked at cause and effect explanations rather than problem-solving scenarios. Mayer and Moreno constructed two experiments, In the first, they required the participants to explain how lightning was formed, and in the second, they required an explanation of how the braking system of a car worked. The first intervention made use of animation and narration (i.e. it used two processing systems, namely, visual and auditory), and the second used animation and text (it made use of the visual processing system alone). The evidence indicated that the animation and narration format was the better design. The split-attention effect was produced in the animation and text version. The learner had to select, organise and then integrate two sources of verbal information. This overloaded the working memory and produced poorer learning performance among the participants. The animation/narration group were able to use two different processing systems to select and process the material. Because of this reduced the cognitive load, they were then able to integrate the content more effectively and perform better. This experiment of Mayer and Moreno (1998) provides an excellent example of how three different theories (cognitive load theory, cognitive theory of multimedia learning and dual-processing theory) can be used as a framework to explain the results of empirical research.

The research of Kalyuga, Chandler and Sweller (1999) overlapped to some extent with the research of Mayer and Moreno (1998). These researchers investigated alternatives to split-attention instructional designs from a cognitive load perspective. These alternatives included narration rather than text, as well as colour to limit the load of searching for information in diagrams. The authors pointed out that dual-mode presentations did not reduce cognitive load per se, but rather increased the effective working memory capacity. In the first experiment there was evidence that the dual-mode format (audio-text) was significantly more effective than the visual only (diagram-text) and the audio-diagram-text formats. They concluded that one could use the modality effect to reduce the negative effects of split-attention, and they produced evidence to show that the redundancy effect (diagram plus text plus audio) negated the positive findings of a modality effect. The use of a diagram **and text together with** narration induced a redundancy effect which overloaded working memory and

influenced learning performance negatively. In contrast to what happened in the early research of 1991 (Chandler & Sweller, 1991), these researchers also measured the cognitive load. They found that the diagram-audio format returned the lowest cognitive load rating. In their second experiment, Kalyuga et al. (1999) were able to demonstrate that colour coding, if used with care, could also be used as a technique for reducing the cognitive load imposed in those situations where learners are compelled to integrate content from a diagram and from text. In cases where the same colour was used for text and related sections in a diagram, learners did not have to waste cognitive resources in a search for information so that they could integrate it. While a split-attention effect was still operative, its impact was reduced through the use of appropriate colour coding. In this experiment, the colour-coded version was also rated as having a lower cognitive load by the participants.

2.11.1.2 Redundancy effects

Instructional design practice in the 1990s regularly used techniques that presented information in two different formats. This is most evident in the use of both diagrams and text to explain concepts (Kalyuga, Chandler and Sweller, 1999). Some of the first experiments undertaken by Chandler and Sweller (1991) to test the predictions of cognitive load theory described the redundancy effect. The behaviourist principle that repetition reinforces learning was called into question when cognitive load researchers began to present empirical evidence that the **elimination** of redundant visual material was, in fact, beneficial for learning (Kalyuga, Chandler and Sweller, 1999).

Moreno and Mayer (2002) explored the conditions under which the addition of on-screen text would facilitate learning in a narrated multimedia explanation. Moreno and Mayer make a specific distinction between *mode* (the format used to present material, e.g. word or pictures) and *modality* (the information processing channel that is being used, e.g. auditory or visual). They defined 'verbal redundancy' as the simultaneous representation of text and narration with identical words. Their review of the literature highlighted the fact there seemed to be contradictions and a lack of congruence between the findings of the cognitive load and verbal redundancy literature. While the verbal redundancy literature had looked at using two modalities and a single mode (words as text and words as audio), the redundancy effect described in the cognitive load literature had looked at using two modalities and two modes (nonverbal (visual) and verbal). Moreno and Mayer set out to reconcile these different findings with regard to redundancy. In their first experiment, which explained the process of lightning formation by using learners with no or very little prior knowledge, they compared learning performance across four conditions:

- Explanation as a narration only versus explanation as narration and text (verbal redundancy).
- Animation before the narration versus animation before the narration and text version (their aim with this sequential presentation was to avoid split attention).

The results of this experiment indicated that verbal redundancy was the better format because the addition of an animation before the verbal learning still produced superior learning performance for the redundant verbal format. In contrast to the findings of Kalyuga et al. (1999), the learners did not

ever have to split their attention between visual and verbal material. In a second experiment, Moreno and Mayer added visual material that was displayed at the same time as the narration or narration-text format. They also presented animation before the narration and narration-text formats. In two of these formats the learner had to process a corresponding animation simultaneously, thereby creating the split-attention effect. They found that when the material was presented sequentially, the verbally redundant condition produced better learning, but that when the presentation was concurrent (a split-attention effect), the redundant conditions produced less efficient learning. In a final experiment they included non-verbal auditory materials (background sounds), and looked at how this would affect learning performance. They found that the addition of sound did not help students to understand the learning material. Moreno and Mayer (2002) concluded that the most effective technique for enabling learning with multimedia explanations was to use the auditory and visual modalities simultaneously for verbal information. In order to avoid cognitive overload this strategy would not be effective if other visual material such as diagrams and animation) was presented simultaneously. A learner cannot watch verbal and visual material and listen to verbal material simultaneously without experiencing high cognitive load.

Moreno and Mayer (2002) did not explicitly address the issue of whether or not the material needed to be integrated. The study of Leahy, Chandler and Sweller (2003) also investigated the redundancy effect in multimedia learning. This study investigated the interactions between the different effects, in this case, the modality effect and the redundancy effect. This study made use of younger participants than had previously been used, and the content was divided into high and low complexity. The researchers hypothesised that the modality effect would be greatest for the more complex learning material (higher intrinsic load). The study also attempted to determine whether audio-visual (sound and text) presentations were always beneficial to learning. Leahy et al. (2003) found that when both audio and visual information were necessary for understanding, and when the intrinsic load of the material was high, the dual mode of presentation was significantly more effective than the visual format alone. They found that when the intrinsic load was low, there was virtually no difference between the two formats when it came to the performance of participants. If the auditory information only repeats what is already present in the visual material and no new information is added, this creates a redundancy effect. They set up a second experiment to investigate this design format by making use of the same content. They changed the visual material so that the visual version became self-explanatory and once again divided the material into high and low complexity. Their findings in this case supported the redundancy effect, namely, that where material is redundant (as was the case in the audio and visual format), learning performance is significantly lower. They also demonstrated these levels of lower performance for both the high- and low-complexity conditions. It is apparent therefore that redundancy overloads the capacity of the working memory.

Diao and Sweller (2007) recently explored redundancy in foreign language reading comprehension. Their investigation focused on the use of concurrent written and spoken presentations. Previous research into the redundancy effect, which I have already reviewed in the preceding paragraphs, used scientific and technical material. Diao and Sweller were interested to know whether or not

redundancy would be beneficial in the learning of a foreign language. Since foreign language learners do not possess the well-developed schemas of first language speakers, there are differences with regard to the speed at which reading and listening develop. Since Diao and Sweller were doubtful that foreign language learners had the working memory capacity to read and listen at the same time, they hypothesised that a concurrent written and spoken instructional format would impose an extraneous cognitive load that would be detrimental to reading comprehension. They measured cognitive load by using subjective mean load ratings and exposed participants to two sessions in the experiment. The results of the study indicated that in the tests for passage comprehension and lexical knowledge, the read-only group performed better than the read/spoken instruction group, who also reported higher mental load ratings. The more difficult the task became, the more the evidence suggested that the spoken instruction interfered with comprehension and knowledge acquisition. While Diao and Sweller did mention the work of Moreno and Mayer (2002) in their introduction, they made no attempt to relate their findings to similar studies. By making use of a scientific context, Moreno and Mayer obtained evidence to show that verbal redundancy was effective provided that the split-attention effect was controlled. If one ignores the knowledge domain, these findings of Diao and Sweller (2007) contradict the findings of Moreno and Mayer (2002).

2.11.1.3 Expertise-reversal effect

A study by Yeung, Jin and Sweller (1997) was one of the first in the cognitive load stream to propose that the level of expertise of the learner moderates the split-attention and redundancy effects. They conducted a series of five experiments that explored the use of explanatory notes in reading instruction and tested both comprehension and vocabulary. Their findings across these five experiments were consistent: whenever the explanatory notes were integrated (thereby avoiding split-attention), the less experienced learners performed better in the comprehension tests. In contrast to this, the more experienced readers in the same circumstances found this integration to be redundant and they performed more poorly on the comprehension test. They measured cognitive load only in the last two experiments. In the last experiment, the learners with better reading skills indicated that the integrated format required more mental effort than the separate format.

Kalyuga, Chandler and Sweller (1998) investigated the role of learner's prior knowledge in the context of training for trade apprentices using elementary electrical engineering instructional materials. A series of three experiments was conducted. Because the learners in Experiment 1 were inexperienced, the results provided evidence for the split-attention effect. These inexperienced learners performed best when the material (diagrams and text) were integrated. Although the learners in Experiment 2 were initially inexperienced, they were given training at different points in time and tested after each stage. The researchers then compared the relative improvement of the groups for the different interventions. The findings indicated that as learner expertise and knowledge increased, the need for using integrated material decreased. Because these results yielded no clear evidence for a full redundancy effect, the researchers concluded that the level of expertise was in all probability not high enough. They therefore conducted a final experiment with more experienced learners and these

findings demonstrated a redundancy effect – experienced learners performed better when there was no redundancy. Split-attention effects in the experiment did not appear to influence performance. Kalyuga et al. (1998) made no reference to the work of Yeung et al. (1997), who obtained similar results in another context and subject domain.

Over the next few years, Kalyuga, Chandler and Sweller (2000, 2001) reported on further research that considered the role of learner experience. Each study looked at either a variation in the design or at other effects, and these slowly added to an understanding of the conditions under which design was more or less effective for both inexperienced and experienced learners. In the study which they conducted in 2000, Kalyuga et al. (2000) examined the role of experience in dual-mode instruction. The purpose was to find out whether the level of learner experience would relate the modality effect to the redundancy effect. Their expectation was that as the level of experience increased, the most effective format would be the diagram-only format and not the diagram-with-audio format. The results of these experiments continued to support the redundancy effect – even in situations where more than one modality was used. Redundant information increases cognitive load for the more experienced learner. Kalyuga et al. (2001) then looked at learner experience and the worked-example effect. Their accumulated evidence suggests that as learner experience grows, learners are better able to learn successfully from conventional problem solving. The redundant material in worked examples merely added to the extraneous load for the experienced learner, and they finally called the role of learner experience ‘the expertise-reversal effect’ (Kalyuga et al. 2003). Kalyuga et al. (2003) point out that most of the instructional effects described in the cognitive load research stream apply to learners with limited experience, and they emphasise once again the need to adjust the instructional design to the learner’s level of experience. This implies that it is necessary for instructional designers to understand a particular learner group before they recommend a particular design.

The most recent application of the knowledge about this expertise reversal effect has been to explore how a rapid assessment of learner expertise can be used to provide adaptive instruction (Kalyuga, 2006; Kalyuga & Sweller, 2005; Kalyuga & Sweller, 2004). Kalyuga and Sweller (2004) took the position that the optimisation of cognitive load in instruction needs to be predicated not only on the presentation of the appropriate information at the appropriate time, but also on the timely removal of inefficient and redundant information as the learner’s level of knowledge increases. This can only be done if the knowledge levels of the learner are continuously assessed and monitored during instructional episodes. These researchers examined research on chess expertise and found that chess masters seemed to remember sets of moves rather than the individual elements of the problem state. Kalyuga and Sweller proposed that there is a memory structure called the long-term working memory, which can be tested in the working memory (WM). If knowledge of the solution moves reduces the WM load more than the knowledge of the elements of the problem state, then a test of appropriate solution moves may be a more valid test of expertise than a test that emphasises the elements of the problems states. They proposed that learners could be presented with an incomplete solution and asked to indicate the next immediate step rather than be asked to provide all the solution steps.

The aim of the first research was to devise a rapid test of the levels of expertise on the basis of knowledge about the human cognitive architecture, to test the validity of this technique, and then to use the results of the test to determine the instructional procedures. The domain of study that Kalyuga and Sweller (2004) selected for the first two experiments was Mathematics. Their findings indicated that there was a high correlation between performance on the rapid test tasks and traditional measures of knowledge requiring complete solutions. They followed up these findings with two experiments that applied the rapid test in order to predict the instructional design procedures that should be used with students of differing levels of expertise. The study found evidence that while those learners with more expertise performed better when they used the problem-solving format, those who were less knowledgeable performed significantly better when they used worked-examples rather than problem-solving strategies. The final experiment moved from paper-based delivery to computer-based delivery and compared performance between two groups: the first group received learner-adapted instruction and the second group was randomly assigned to non-adaptive instruction. It was found that the learner-adapted group performed significantly better than the randomly-assigned format group. In 2005, Kalyuga and Sweller (2005) extended the work of the 4th experiment in the 2004 study by using a yoked control design. The learners were given the rapid test to determine their level of expertise. This was then followed by measurement of the cognitive load by using the subjective rating scale of Paas, which was then used to calculate, in real time, the instructional efficiency of the instructional method used in the rapid test. The method for calculating instructional efficiency differed from the original method developed by Paas and van Merriënboer (1993) because of the need to calculate instructional efficiency in real time. The results of this assessment were used to provide the initial learning path in the instruction. Thereafter regular rapid assessments were done, the aim of which was to determine if the learner was still on the correct learning path for their level of expertise, which changed as they progressed through the instruction. There was evidence that the learner-adapted group obtained higher knowledge and cognitive efficiency gains than the control group. In a study undertaken in 2006, Kalyuga (2006) looked at using this rapid assessment when the content involved solving arithmetic word problems. He compared the rapid assessment technique, in which the learner was required to provide **only** the first step to the solution, to the traditional assessment method that required the learner to provide the entire solution. These early results indicate that this technique could be used when learners need to solve word problems. The limitations of this study include the fact that it was conducted in an experimental setting. Kalyuga suggested that the approach would need to be adapted for an authentic environment. There is also a need to test this method in other domains. It would be interesting to see whether similar rapid assessment techniques could be used to determine the level of expertise that medical students demonstrate during the making of clinical diagnoses, which represent another type of problem-solving scenario.

2.11.1.4 Modality effects

The modality effect states that learning will be enhanced if textual information is presented in an auditory rather than (the conventional) visual format, and if such an auditory format is accompanied

by visually based information such as graphs, diagrams or animations (Ginns, 2005). Modality effects have been researched in both the cognitive load and multimedia learning research streams. In this section, I will look at the research undertaken from a cognitive load perspective, and in Section 2.11.4, I will examine the research undertaken from the perspective of multimedia learning.

Leahy, Chandler and Sweller (2003) explored the conditions under which auditory presentations might be effective or ineffective. They found that when the intrinsic load of the material is high, a dual mode of presentation (audio and text, in this experiment) achieved better results than a text-only presentation. Leahy et al. did not measure the cognitive load of the instruction. The study of Tabbers, Martens and van Merriënboer (2004), which did measure cognitive load, tested the theory of modality and cueing in a classroom setting by making use of text, audio and diagrams. They found that replacing text with audio resulted in no positive effect on retention scores. They also found that while the addition of visual cues to diagrams increased retention scores, it did not influence the outcomes of a transfer test. This limited review indicates that the findings were both unexpected and conflicting (Leahy et al., 2003; Tabbers et al., 2004). Much of the cognitive load research had been carried out in laboratory conditions. But when the same experiments were conducted in a classroom setting, the findings could not be replicated (Tabbers et al., 2004). Ginns (2005) identified 43 studies for inclusion in a meta-analytic study on the modality effect. With the exception of two studies published in 1974, the remaining studies were published between 1995 and 2004. Of these, 25 were published in 2001 and later. Studies were included in the meta-analytic study if the article contained instructional conditions means and standard deviations, or a statistic from which a d value could be computed (t or F statistic). The meta-analysis was based on the following three major hypotheses:

- The presentation of instructional material by using a combination of an auditory mode for textual information, such as spoken text, and a visual mode for graphical information, will be more effective than a presentation of all the information in a visual format.
- The strength of this effect is moderated by the level of element interactivity (high versus low) of the learning materials, even though Ginns acknowledged that there was no objective measure of element interactivity.
- The strength of this effect is moderated by the pacing of the presentation, with a strong effect for system-paced material but a lesser effect for self-paced materials.

Each study was coded for the following variables: level of element interactivity (high, low), pacing (system versus self), form of outcome variate (similar questions, transfer questions, time to solution scores, subjective rating of cognitive load), broad field of study, type of testing (individual versus group), age group, and form of modality presentation (audio-tape, computer screen, virtual reality).

The effect size used in the analysis was d , which is the difference between the means of the different conditions divided by the pooled standard deviation. Between-subject and within-subject designs were analysed separately. When the between-subject designs were compared, the overall weighted mean effect size was large $d = 0.72$ (95% confidence interval 0.52–0.92). Further analysis indicated

that the mean effect size was greater for high element interactivity materials than for low element interactivity material. The pacing of the presentation was also a significant moderator, with system-paced instruction yielding a substantially higher effect size than self-paced material. The results also suggested that the modality effect might be particularly strong for students who were learning by means of virtual reality media. Ginns's analysis supported the three major hypotheses, even though some of the studies did publish contradictory findings. Ginns (2005) concluded that there was a need for considerably more research. The reasons for requiring more research were as follows:

- The sample sizes for some of the comparisons were very small.
- Only a few studies measured cognitive load using the dual-task approach.
- More comparison was needed to test the influence of the pacing of instruction on the modality effect.
- The role of learner expertise needed further investigation, especially in the light of Kalyuga, Sweller and Chandler's finding that the modality effect was different when one compared novice and experienced learners (Kalyuga, Chandler & Sweller, 2000).

Ginns (2005) also raised the issue of cost-benefit analysis. If the modality effect is a factor in learning, then cheaper and easier methods for design and development need to be made available.

2.11.1.5 In conclusion

This concludes the section in which I reviewed how cognitive load theory has been used to guide the instructional design of instruction. While there is clear evidence for the split-attention, redundancy and expertise-reversal effects, more research is needed to explore the modality effect, and the synergies between it and the other effects. There is evidence that this modality effect cannot be ignored. My study will use two different modalities for presenting selected content: audio (auditory) and text (visual) and the learning performance for these two modalities will be compared.

In the next section, I will consider the development of the cognitive load theory.

2.11.2 Theory development and directions of research

The cognitive load theory has already been described in Section 2.6 of this chapter. Explanations of this theory and the findings of early empirical research first appeared in the literature in the late 1980s (Sweller, 1998) and early 1990s (Chandler & Sweller, 1991; Paas, 1992; Paas & van Merriënboer, 1994a). The theory addresses the way in which cognitive resources are used during learning and problem-solving and the findings from the research have helped to develop our knowledge and understanding of the human cognitive architecture. Researchers using this theory currently draw on this knowledge of the human cognitive architecture to investigate further instructional designs that reduce cognitive load and facilitate learning. Theorists have also started to link cognitive processes to the processes that underlie biological evolution (Sweller, 2006b; Sweller, 2004; van Merriënboer &

Sweller, 2005). The intention of these researchers has not been to look at the instructional implications of the biological information-processing systems, but rather to use a knowledge of the biological information-processing system to strengthen their understanding of the human cognitive architecture.

The cognitive load researchers proposed that effective instructional material facilitates learning by directing cognitive resources toward activities that are relevant to the learning rather than toward the preliminaries of learning. Early research explored the use of strategies that required learners to mentally integrate mutually referring sources of information. The outcomes of this research provided clear evidence for the split-attention effect (Chandler & Sweller, 1991). The first explanations of the theory described two types of cognitive load: intrinsic and extrinsic. Extrinsic load was then further divided into extraneous and germane load.

This new theory and its related research were soon critically reviewed by the research community (Goldman, 1991; Dixon, 1991). As recently as 2006, in fact, researchers were still deliberating about the value of the theory (Moreno, 2006).

Goldman (1991) offered several criticisms of the work presented by Chandler and Sweller (1991). These cohered around three core issues: cognitive load theory as an heuristic for instructional design, cognitive load theory as a general theory of learning, and the role between cognitive load theory and instructional strategies. Part of Goldman's critique was that the studies were insufficiently detailed to test the usefulness of the prescriptions for instructional design. She noted that there was not much clarity about which activities were extraneous to the learning task. She also pointed out that the results did not show significant support for cognitive load theory when the learning outcomes were directed towards the application of learning rather than the mere reproduction of information. She felt that the dependent measures in Chandler and Sweller's work relied too heavily on learning that was defined by the ability to master what had been presented. Goldman also questioned whether the predictions tested were unique to cognitive theory, and she was of the opinion that there were other theories that addressed presentation format that could also have been used to explain the findings.

As an instructional designer who aims to use the verifiable results of empirical work as extensively as possible, I must challenge this criticism by asking whether it is not more valuable to have several theories that address the same issue. It is my opinion that the practical application of several theories will lend a greater degree of credibility to the design. It will also, in my opinion, justify the value and importance of taking the empirical research into account when making instructional design decisions. If several theories and the combined weight of research that such theories have stimulated all point to the same conclusions, what grounds would an instructional designer have for ignoring the findings and not applying them scrupulously to the design and development of instructional material.

Goldman went on to state that the second failure of CLT was the failure to differentiate among different kinds of learners. Whether the work of Kalyuga et al. (2000, 2001, 2003) on the expertise-reversal effect was a direct result of this critique is not clear, but these cognitive load researchers **did** in fact subsequently look at different kinds of learners. She was also critical of the fact that the work of Chandler and Sweller did not actually test the theory itself, but looked more at the instructional applications that arise from the theory. It was also her opinion that there were some dissonances between cognitive load theory and constructivist theory that required further clarification.

The preamble to Dixon's (1991) critique of the work of Chandler and Sweller (1991) examined the relationship between research and application. Dixon felt that a fundamental problem in cognitive science was that the theories were developed in laboratories to account for simpler but very precise experiments, rather than addressing the tasks and variables that are important in the real world. While it has taken the cognitive load field just under ten years to respond to this critique, there has been an increase in the CLT research since 2000 that examines the more complex learning that takes place in longer learning programmes rather than that which occurs in brief experimental situations (van Merriënboer, Kirschner & Kester, 2003; van Merriënboer & Sweller, 2005). I will discuss the results of this research in a later section of this chapter.

Dixon proposed that the problem (namely, addressing tasks and variables that are important in the real world) can be solved in two ways. His first solution was to develop a theory based on conditions in the real world, and then to test such a theory in the real world. Dixon referred to this process as 'theory development' because it takes account of cumulative progress in creating a defined body of knowledge, even if individual experiments fail to support a theoretical position.

Dixon says:

In theory development, the experiments provide the basis for distinguishing among different possible accounts of the cognitive processing that occurs in the situation of interest.

(Dixon, 1991, pg 345).

The second approach he called 'theory application'. In this case, the theory is used to solve a particular problem and the goal of the research is to demonstrate the efficacy of the proposed practical solutions. In theory application, the theory is applied to a particular situation and the experiments only demonstrate the predicted effects of the theory – they do not test the theory itself.

Like Goldman, Dixon maintained that the work of Chandler and Sweller (1991) was an example of the theory application approach. Dixon maintained that because cognitive load theory had been generated as a consequence of investigations into problem-solving in learning, he remained unsure about whether it could also be applied to other kinds of instructional materials with the hope of

returning similar results. He felt the work of Chandler and Sweller had failed in five distinct areas that he characterised as follows:

- Cognitive load theory (CLT) does not predict whether the integration of material will have a large or a small effect on performance.
- CLT does not indicate when redundancy is harmful to learning and when it can be ignored.
- Neither the theory or the experiments provide a good description of what integration means, and there are many options for integrating material that have not been explored.
- CLT does not adequately address the problems with the conventional formats and how these could be overcome.
- The concept of cognitive load is not sufficiently explicit to indicate what kinds of manipulations will reduce cognitive load.

In responding to the commentaries by Goldman (1991) and Dixon (1991), Sweller and Chandler (1991) took the position that while scientific theories have many goals, there is only one goal for cognition and instruction, and that is to generate new and useful instructional techniques. They provided a brief history of the origin of CLT and pointed out that since their original work, researchers had avoided *post hoc* explanations of results. This, in their opinion, constituted the distinctive strength of cognitive load research. The theory had rather been used to generate novel applications in a wide variety of areas, and the findings offered direct instructional applications. Their responses to the various criticisms from Goldman could be summed up in the following statements:

- Psychological theory cannot produce accurate quantitative predictions at the level that Goldman requires because the nature of the field makes this near impossible. There are simply too many factors that cannot be controlled or even adequately and precisely described.
- There is as yet no empirical evidence that teaching students how to integrate material is effective.
- The work of the CLT theorists **had** in fact demonstrated that superior transfer takes place when one uses new formats.
- Unlike CLT, which had always generated the prediction and then tested it, theories that cover issues similar to CLT often did so by means of 'after-the-event predictions'.
- CLT had never been considered a global theory of learning. It is a narrower theory that was designed solely to generate novel instructional applications.

Their response to Dixon (2001) was that since theory development and theory application are in fact compatible, they should be used serially, although this was seldom the case. The theory development approach generates process models which do not have instructional implications. The next step is to take the proposed process and generate hypotheses about the instructional procedures that are needed to facilitate the process. This is theory application and it represents a natural progression from theory development to theory application. Dixon's critique (presented earlier) simply gave

Chandler and Sweller some ammunition to use, and they used the opportunity to state that Dixon's critique not only pointed to the need for more research, but it also provided the direction for that research.

So where did the theory go from there? Figure 2.7 illustrates the CLT research timeline. This timeline indicates when a specific focus of the research began. Thus, for example, the research effects focus began between 1994 and 1998, but it did not necessarily end in 1998.

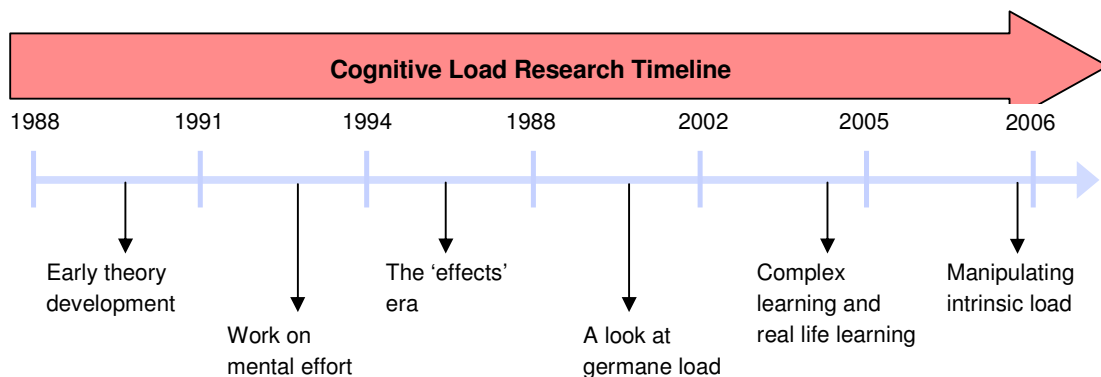


Figure 2.7: Cognitive load research timeline

Paas (1992), who devoted more attention to the concepts of mental load and mental effort, was one of the first researchers to measure mental effort in the learning environment. He viewed the absence of cognitive load measurement as one of the limitations of the research that had been carried out up until that point. The first visual representations of cognitive load theory appeared in the literature in 1993 and 1994, and the different elements in the theory finally began to be described in substantial detail (Paas & van Merriënboer, 1994a, 1994c).

Sweller et al. (1998) provided a very concise overview of instructional designs based on CLT and the empirical work related to these designs. I have called this the 'effects' era, and I discussed this work in Section 2.11.1.

Since 2000, the cognitive load research community has stopped regularly to recap on the theory, to summarise the broad scope of the field and trends in the research, to highlight the implications of the empirical work or to map future directions for the field.

Kirschner (2002) proposed that designing for a competency-based paradigm, for which there are growing calls in education, calls for new approaches to instruction. While he was of the opinion that cognition, meta-cognition and transfer were the most important variables in this new paradigm, he still regarded CLT as a good tool for helping designers to understand and to take into account the limitations of the human mind. He felt that designers needed to make use of this knowledge in designing for transfer of learning. In his introduction to the special issue on cognitive load theory,

Kirschner (2002) described the trend in the following way: it is a movement from research that had, as its sole focus, sought ways to decrease extraneous load toward research that had started to consider germane load in instructional design. While he left the mapping of future directions to Valcke (2002), he highlighted several consequences for CLT that arise out of the research published in the special edition.

I highlight four of these consequences because they are relevant to my study in various ways:

- The role of prior knowledge. Kirschner asked whether the lack of such prior knowledge would promote deeper processing or limit learners because they would only then be capable of superficial processing.
- The issue of the amount of time needed to study.
- A need to look at the principles proposed by the multimedia research stream. These researchers did not measure cognitive load but used the theory to design instruction and explain results.
- The question of whether intrinsic load could really not be manipulated (as first suggested by the theory).

In his commentary on the publications of the 2002 special edition of *Learning and Instruction*, Valcke (2002) mapped an agenda for future research that needed to take a closer look at the following three areas:

- How can CLT be used to provide answers to the call for instruction to move away from the mere presentation of information (cognitivism) to a format in which learners are given the opportunity to construct their own knowledge and understanding (constructivism)?
- The relationship between cognitive load and meta-cognition, and the possibility that CLT should include a fourth type of load, namely, meta-cognitive load, which is linked to germane load.
- The question of the role of prior knowledge and the suggestion that the CLT framework needs to be updated so that it can more clearly indicate where prior knowledge fits in.

Although I was unable to trace any research that specifically addresses Valcke's suggestion to look at this meta-cognitive load, there have been several studies that have specifically focused on various aspects of prior knowledge (Ayres, 2006a; ChanLin, 1999; Clarke, Ayres & Sweller, 2005; Reisslein, Atkinson, Seeling & Reisslein, 2006; Schnotz & Rasch, 2005; Seufert, 2003; Wallen, Plass & Brünken, 2005).

In yet another special edition from 2003, Paas, Renkl and Sweller (2003) offered an overview of research that had investigated the more dynamic approaches to instructional design. These approaches to design were based on new understandings of cognitive load theory, which proposed that changes occur in cognitive load as the learner moves from being a novice to being an expert

within a particular domain of knowledge. The cognitive load research stream had started to look at manipulating more than just extraneous load in order to reduce the total cognitive load, and was focusing more on the various design strategies that accommodated a novice or expert learner's ability to cope with cognitive load (Renkl & Atkinson, 2003). Cognitive load research was also beginning to consider the role of the goal of instruction from the point of view of the teacher and the learner (Gerjets & Scheiter, 2003).

Gerjets and Scheiter (2003) were of the opinion that one of the shortcomings of cognitive load theory was the assumption of a one-to-one mapping between instructional design and cognitive load. Their illustration of this mapping is displayed in Figure 2.8 below.

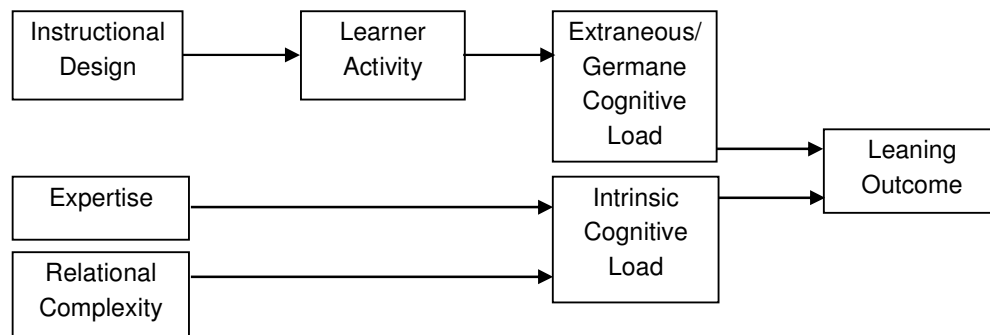


Figure 2.8: Assumptions of cognitive load (Gerjets & Scheiter, 2003, page 35)

Gerjets and Scheiter asserted that both the teacher and learner's goals for the instruction, in conjunction with the activities of the learner (i.e. what they did with the material), moderated the relationship between instructional design and cognitive load. They felt that the variability with which learners process instructional material had not been given sufficient attention in cognitive load theory. While they did not mention cognitive style per se, the variable ways in which learners process information could well include cognitive style because cognitive style is all about the processing of information. This concern indicates that it might be worthwhile to look at the relationship between cognitive load and cognitive style more closely.

Gerjets and Scheiter then proposed the augmentation of cognitive load theory that is illustrated in Figure 2.9.

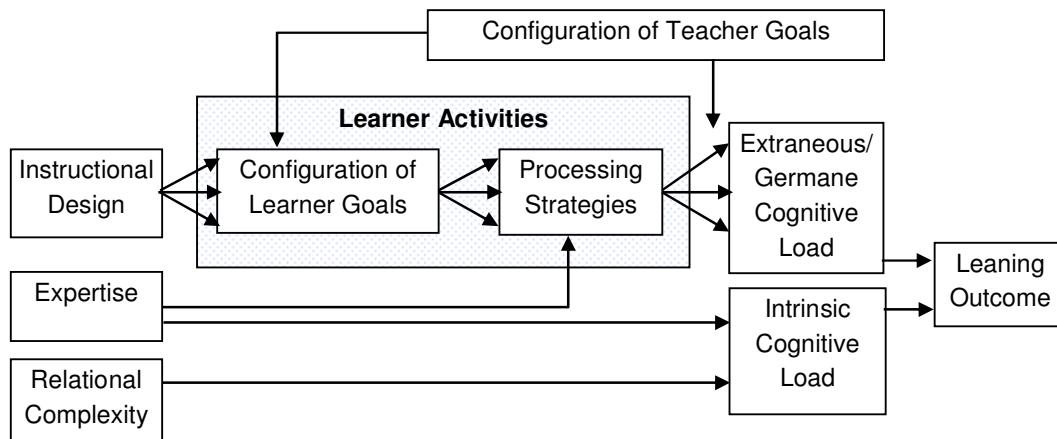


Figure 2.9: Goals and strategies as moderators between instructional design and cognitive load (Gerjets & Scheiter, 2003, page 36)

They substantiated this hypothesis with evidence from their hypertext-based instruction research. They were careful to point out that this research had not initially been designed to test a new version of cognitive load theory. They compared two approaches to learning with worked-out examples in a basic arithmetic lesson: surface versus structure emphasizing. Learners needed longer time to process the structure-emphasizing approach. In the test situation, those learners who had used the approach that emphasised structure performed better when the test items were not equivalent to the ones used during the instruction. Gerjets and Scheiter explained their findings by using cognitive load theory as the framework, and proposed that if the goal was to complete tasks as quickly as possible, then the processes in the structure-emphasizing approach imposed an additional extraneous cognitive load. If, however, the goal was eventually to solve transfer tasks, then it became important to process information more deeply. In such a case, the load is no longer extraneous, but germane, and that has a positive effect on learning outcomes. Gerjets and Scheiter (2003) also manipulated the goal of the instruction for the learner. As soon as additional goals were imposed on an initial task, the problem-solving performance of the learner was impaired. These researchers were of the opinion that they could not necessarily use cognitive load theory to explain their findings if they did not consider the new augmentations that they had proposed for cognitive load theory. I could find no research that has investigated this augmentation of the cognitive load theory any further.

The emergent empirical research of 2004 could be divided into two streams: the research that extended the investigations carried out by previous work, and new directions in research. The new theoretical perspectives of that year included drawing an analogy between evolution by natural selection and human cognitive architecture (Paas, Renkl & Sweller, 2004; Sweller, 2004). The

research field, while continuing to increase in diversity, was still engaged in investigations into the worked-example effect, the role of learner expertise, and the synergies between these two areas.

Van Gog, Paas and van Merriënboer (2004) explained from a theoretical point of view why it was necessary to consider using worked-examples with process-based information. They argued that when designers add process-orientated information, they add a 'why' and 'how' dimension to the learning. These 'why' and 'how' answers are part of an expert's schemas. Providing the learner with such a process-based view will increase the germane load of the learning event, and this will in turn make a positive impact on the transfer of learning and performance. Van Gog et al. (2004) recognise that the addition of the process perspective does not guarantee that the learner will pay attention to it. Because of this, they recommend that such a design strategy be used in combination with the strategies for teaching with worked examples. The elaboration of previous research with worked examples included an investigation into fading as a strategy in worked-example designs (Renkl, Atkinson & Große, 2004) and the reduction of intrinsic load during worked-example instruction (Gerjets, Scheiter & Catrambone, 2004).

Fading is a technique that can be used to facilitate the transition from worked-example instruction to problem-solving instruction. Renkl et al. (2004) reported that empirical evidence already existed to suggest that fading was an effective strategy. It was the underlying mechanisms that were not clear, and their study undertook to investigate these mechanisms. The results of their research indicated that the actual position of the 'faded step' (whether at the beginning or end of the worked example) did not influence the outcome. They also produced evidence that learners learned most about the steps that were faded, thus providing support for the value of germane load. Wherever a step is missing, learners must process more deeply to complete the worked example, and this in turn leads to improved performance. But the researchers were not able to elucidate the underlying cognitive processes. Although they intended to look at self-explanation, the frequency with which it was used by participants was so low that they were unable to come to any useful conclusion.

New areas for research included the use of different feedback strategies for the purpose of decreasing cognitive load (Moreno, 2004) and using knowledge about mental effort and performance to dynamically select tasks in instruction (Salden, Paas, Broers & van Merriënboer, 2004). Even though neither of these studies was intended to extend cognitive load theory as such, they have opened up new areas for research. The context of Moreno's study (2004) was discovery-based learning in a multimedia environment. Moreno's study focused on finding a solution for the dilemma posed by the need to get learners actively to engage in the learning process. This gives rise to situations in which the cognitive load can prove to be too high for the learner, and too high for the limited capacity of the learner's working memory. Such a situation therefore requires designers to look at ways of reducing the cognitive load. Learners had to make use of a gaming, discovery-based format to design a plant that would survive under different conditions. Help, which was either explanatory or corrective in nature, was provided by means of a software agent. Moreno found evidence in both experiments that the provision of explanatory feedback to novice learners resulted in

significantly better performance on both retention and transfer tests if one compared their performance to the performance of those learners who had received only corrective feedback. The explanatory feedback format of the program was also instructionally more efficient than the corrective feedback format.

Salden et al. (2004) replicated some of the earliest research that used cognitive load as a basis for adapting instruction. The earlier study was slightly modified by the inclusion of a different technique for measuring instructional efficiency, by increasing the number of tasks, and by reducing the complexity between the tasks. Salden et al. used learning performance, mental effort or mental efficiency (performance and effort) to determine the task with which the learner would be confronted in the following instance. The results of this study were as follows: (1) Salden et al. confirmed the hypothesis that dynamic task selection leads to more efficient training than a fixed task selection, and (2) they were unable to confirm the second hypothesis that dynamic task selection based on mental efficiency will lead to more efficient training and a better transfer than selection that is based on mental effort and performance alone. An fruitful path for future research would be to base this strategy of dynamic task selection on mental efficiency or performance and mental effort scores and test it in the field of health sciences education. It would be possible to devise and design a series of case studies of incrementally increasing difficulty to teach clinical diagnostic skills, and to expose learners to these case studies on the basis of calculation of mental efficiency.

In their commentary at the end of the special edition devoted to cognitive load theory, Rikers, van Gerven and Schmidt (2004) emphasised the value of the research contributions to the existing body of knowledge about learner expertise. They are of the opinion that cognitive load theory is capable of guiding designers when they have to make decisions about the instruction needed to develop expertise in a domain. By taking mental effort, performance and mental efficiency into account, designers will be in a better position to assess whether or not tasks are at an appropriate level of difficulty for the learner, and teachers will be in a better position to select appropriate instructional approaches and problems that address the specific needs of their learners.

In 2005, the *Educational Technology Research and Development* journal devoted a special edition to cognitive load research. This addition confirms that studies were increasingly looking at cognitive load in the e-learning environment at that time. There had also been a move to using real courses rather than shorter laboratory studies as the context for the research. In their review of the current research, van Merriënboer and Ayres (2005) divided the research into three main streams:

- Design strategies that focus on manipulating intrinsic load
- Measures that simulate learners to invest more effort (germane load) into learning
- The use of learner expertise as the basis for adaptive instruction

Van Merriënboer and Ayres (2005) also noted that most of the research incorporated established findings from cognitive load theory into the experimental designs. Although the research published in

that special edition on cognitive load did not propose any changes to the existing cognitive load theory, it did, however, make an effort to establish a connection between the theoretical frameworks of cognitive load and expert performance research (van Gog, Ericsson, Rikers than & Paas, 2005). The establishment of the expertise reversal effect forced cognitive load theorists to broaden the scope of their work and start to look at the design implications for more expert users. When van Gog et. al. (2005) looked at the connection between these two areas, they identified the following new directions for research: the need to consider adaptive, individualised instruction that is based on authentic tasks and that gradually allows learners to take control of the process, and the relationship between motivation, mental effort and instructional efficiency in learning and in ways of increasing germane cognitive load.

A special issue of *Learning and Instruction* in 2006 devoted more attention to the empirical work that had used worked examples within the cognitive load framework, with the emphasis on germane cognitive load. Although there is a whole range of strategies that reduces extraneous cognitive load, instruction becomes even more effective when the germane load is increased. Since learners are unlikely to engage spontaneously in the activities that increase germane load (Paas & van Gog, 2006), the cognitive load research stream has turned its attention to looking at strategies and activities that will deliberately increase this germane load. Even though the studies reported in this special edition are highly relevant for fields that benefit from using worked examples as a strategy, I will not discuss them in detail here because they have no new contributions to make to the cognitive load theory per se.

2.11.3 Measurement of cognitive load

There are several methods of measuring cognitive load, apart from performance measures, that are described in the literature. They are:

- Self-report ratings (Paas, van Merriënboer & Adam, 1994b)
- Physiological measures (Paas, van Merriënboer & Adam, 1994b; Gevins, Smith, Leong, McEvoy, Whitfield, Du & Rush, 1998)
- A direct method that uses a dual-task approach (Brünken, Plass & Leutner, 2003)
- Subjective time estimation as an index of cognitive load (Fink & Neubauer, 2001)

The assumptions that underlie these different methods of measuring cognitive load are summarised in Table 2.4.

Method	Underlying assumption
Self-report ratings	A learner can introspect on his/her cognitive processes and accurately report on the amount of mental effort that he/she has spent.
Physiological measures	Changes in cognitive function are reflected in physiological changes which include changes in pupil diameter, heart rate, and certain brain wave formations that are demonstrable in EEG readings.
Direct method using dual-task	If the primary task is to learn from the content being presented, and a person is asked to respond to a secondary task, the time it takes for this response will be determined by the cognitive load imposed by the primary task. The longer it takes the person to respond, the higher the cognitive load of the primary task.
Subjective time estimation	If the cognitive load of a task is high, a person does not have enough capacity to monitor the time they are spending on the task, and subjective estimations will become more inaccurate as the cognitive load increases. The estimation of time spent is typically shorter than the actual time that is spent, and it decreases in length as the cognitive load increases. This is based on the attentional model of prospective timing.
Neuroimaging	Active areas of the brain consume more oxygen. This effect can be detected by pulses rebounding in the magnetic fields generated by an magnetic resonance imaging (MRI) scanner.

Table 2.4: Assumptions underlying different techniques for measuring cognitive load

Neuroimaging (Whelan, 2007) has recently been proposed as a more accurate technique for measuring cognitive load. The argument for adopting this method addresses two problematic issues in cognitive load measurement. The first issue relates to the fact that current measures do not differentiate between the types of cognitive load. The second issue is the criticism that existing approaches are limited in terms of both their precision and methodology. How does neuroimaging work? Different areas of the brain mediate attentional control, working memory and cognitive workload. fMRI is a highly sophisticated technology that records the brain's hemodynamic response and activity so that it is possible to pinpoint activity in the different areas of the brain during a process such as learning. Whelan (2007) argues that because cognitive load theory has a basis in functional neuroanatomy, fMRI technology will allow researchers accurately to observe the properties of certain brain functions that are related to different types of cognitive load. Whelan uses existing neuroscience research to support his argument for considering neuroimaging as a cognitive load measurement technique. While his arguments may appear to be plausible, the studies that he cited seem to indicate that the techniques he recommends are able to differentiate between different types of cognitive load. My concern at this stage is that the tasks used in the experiments that he reviewed were far removed from real-life, practical learning tasks. Since this current research also favours the neuroscience research perspective rather than the educational research perspective, a definite effort will need to be made to reconcile these two perspectives. Although Whelan (2007) provides an outline for a possible

study that looks very interesting, he simultaneously points out the challenges that confront the research methodology when proposals to use this scanning technique for measuring cognitive load. Some of these challenges are concerned with the actual measurement and interpretation of the results. Other challenges relate to the practical difficulties of implementing this technique on a regular basis during computer-based instruction.

But these are questions that will be dealt with in the future. Let us return to the past, and then move back into the reality of our present state of affairs.

The cognitive load theory, which was illustrated and discussed in Section 2.6 of this chapter, indicates that there are three factors that are influenced by cognitive load: mental load, mental effort and performance. Mental load is imposed by the instruction (i.e. by the task structure and the sequence of information), and mental effort refers to the amount of capacity that is allocated to the instruction (Paas, 1992). Mental effort consists of all three causal factors in the theoretical model: task characteristics, subject characteristics, and the interaction between these (Paas et al., 1994b). Before 1992, cognitive load had already been determined primarily by looking at performance and task-based measures, and the suggestion was also made that time could be an indicator of cognitive load (Chandler & Sweller, 1991). It was not common then to determine cognitive load during instruction (Paas, van Merriënboer, 1994a).

Paas was one of the first cognitive load researchers to look into determining the cognitive load of instruction by applying mental effort measures (Paas, 1992, 1993). Paas developed a 9-point rating scale that has been widely used since 1994 (Kalyuga, Chandler, Tuovinen & Sweller, 2001; Paas & van Merriënboer, 1994c; Van Gog, Paas, & van Merriënboer, 2006), although some researchers have modified this scale to a 7-point scale (Ayres, 2006b; Mayer & Chandler, 2001; Pollock, Chandler & Sweller, 2002). The 9-point scale was derived from the work of Borg, Brattfisch and Dornič (1971), who were working in the area of the concept of perceived difficulty. Perceived difficulty was originally explored in the context of physical work. In the 1970s the researchers started to look at the concept as it related to mental work, and particularly as it applied to learning. They were of the opinion that psychological testing at that time was too focused on looking at achievement (the objective measure) and that it had neglected to consider the subjective cost at which performance is achieved. This research has concentrated on constructing and refining category scales with ratio properties, and its results have produced the development of what Borg (2004) calls 'level-anchored ratio scaling'.

After the first study (Paas, 1992), Paas et al. (1994b) took a more rigorous approach to investigating the measurement of cognitive load using mental effort techniques. Mental effort may either be measured subjectively by using rating scales, or objectively by using physiological parameters. By using the data from two previous studies (Paas, 1992; Paas & van Merriënboer, 1994c), the researchers investigated the reliability and sensitivity of the subjective rating scale. In the 1992 study, the coefficient of reliability (Cronbach's alpha) was 0.90 with the use of 28 measurements, and in the 1994 study, it was 0.82 with the use of 6 measurements. 'Sensitivity' was defined in these studies as

the capability of a measurement technique to reflect differences between training and transfer conditions. When defined in this way, the subjective rating scale was found to be a 'sensitive' measure. The physiological method that used heart-rate variability (Paas et al., 1994b) proved to be a less reliable method. Although all the measures, except one, correlated significantly, the correlations were so low that Paas and van Merriënboer concluded that the technique was not reliable. This method could also not be defined as a sensitive measure in terms of the definition described above. The subjective rating scale is very easy to use and implement in a classroom situation. It is not an intrusive or invasive test. When it was used together with performance measures, it proved to be a promising technique for providing additional information about cognitive load. This subjective rating technique remains widely used today and, when reported, demonstrates good reliability since it returns a Cronbach alpha of more than 0.80 (Kester, Kirschner & van Merriënboer, 2004; Stark, Mandl, Gruber & Renkl, 2002; Tabbers, Martens & van Merriënboer, 2004).

Gevins et al. (1998) used a secondary task, subjective rating (not Paas's scale) and EEG tests to differentiate between the cognitive load of computer-based tasks. All three of the techniques that were used demonstrated that as the mental load increased, so did the mental effort that was required. Reaction times (secondary task) became longer as the load increased, ratings became higher as the load increased, and EEG patterns for the theta, alpha and beta bands (different wave patterns in the brain) changed as the load increased. None of the researchers attempted to determine how the three different techniques were correlated. When one looks at the research methodology and the subsequent analyses of results, the complexity of the method that uses EEG and neural pattern recognition is evident because it requires sophisticated skills and an ability to analyse that exceeds the skills set of most instructional designers. This technique is therefore not viable for the instructional designer who is looking for an efficient and effective heuristic to determine cognitive load. EEG techniques do, however, offer a viable method for determining exactly which areas of the brain are affected by specific tasks with different cognitive loads. I did not find any other empirical work that used physiological measures of this kind to measure cognitive load in the general educational psychology literature.

It was Paas (1993) who introduced the idea of instructional efficiency. Instructional efficiency combines measures of cognitive load with measures of test performance to derive information about the relative efficiency of instruction. Paas suggested that using mental efficiency measures could be used to serve as a safeguard against the possibility that subjective ratings measure some other subjective characteristic unrelated to the instructional outcomes. If learners find something easy to learn and then perform better on tests, there is an increased likelihood that it is the cognitive load that is being rated rather than some other factor. And while it is also possible for a learner to invest more mental effort in order to compensate for increasing cognitive load, this will not necessarily be reflected in performance measures. Learner A can rate the cognitive load of instruction as low, invest little mental effort, and obtain 50% in a test. Learner B, on the other hand, can rate the cognitive load of the same instruction as high, invest considerable effort, and also only obtain 50% in the same test. By combining performance and mental effort measures, we can obtain an index of mental efficiency

(Paas & van Merriënboer, 1993; Paas et al., 2003). This then makes it possible to compare the mental efficiency of instructional conditions.

There are also different approaches in the methods that are used to calculate mental efficiency (Paas et al., 2003). Researchers have used:

- mental effort during a test and performance in the test (Paas, 1993, 1994c).
- mental effort during training and performance in the test (Kalyuga, Chandler, Tuovinen & Sweller, 2001; Pollock, Chandler & Sweller, 2002).
- mental effort during training **and** the test, and performance in the test (3-D approach) (Tuovinen & Paas, 2004).

Tuovinen and Paas (2004) propose that this third method of calculating instructional efficiency is perhaps the more sensitive measure for determining comparative instructional efficiency because learning effort and test effort are not necessarily equivalent. They note that it is possible for two learners to achieve the same performance score when the one learner must work very hard at the test to achieve the same number of correct answers while the other learner invests very little effort in the completion of the test. They compared the three methods of determining efficiency, learning, test and 3-D, for two instructional strategies (worked examples versus exploration) by using the data from a previous study (Tuovinen & Sweller, 1999). While the results for some of the comparisons all pointed toward the same direction (namely, the superiority of the worked example strategy), this was not the case for the results of other comparisons. Tuovinen and Paas (2004) propose that the use of all three methods to determine efficiency makes it possible to compare the effort required at different stages in the total learning process. This knowledge could enable better choices of regard to design strategies. Even though they recommended that this 3-D approach be further investigated, no researcher seems as if to have taken up the challenge.

The next method that received attention in the cognitive load research stream was use of a dual-task technique for measuring cognitive load. This is an objective measure of cognitive load, and it was pioneered in the multimedia learning field by Brünken, Steinbach, Plass and Leutner (2002). Brünken, Plass and Leutner (2003) describe two methods of using the dual-task approach. In their first approach, the secondary task is added to the primary task and performance is then measured. Performance should decrease when a secondary task is added to the primary task. In their second method, the secondary task is used to measure the load of the primary task. The variable measured is the secondary task. If the load of the primary task is high, the 'performance' in the secondary task varies. In all cases this performance is typically a reaction to some sort of stimulus: an auditory and visual stimulus followed by writing down a result (Chandler & Sweller, 1996), auditory stimuli followed by pressing a foot pedal (Marcus, Cooper & Sweller, 1996), and a visual stimuli followed by hitting a key on a keyboard (Brünken et al., 2002). There are three advantages to using such dual-task techniques:

- The cognitive load is measured at the exact point in time when the load is induced. This avoids the interference of confounding variables caused by the time delay between induction of load and response.
- The researcher can identify the exact step in information processing in which cognitive load is induced.
- Since this technique is usually carried out in within-subjects designs, the measurement of cognitive load is independent of the individual differences that are known to affect between-subject designs (Brünken et al., 2002).

But this technique is not without its challenges. The secondary task must require the same cognitive resources as the primary task. In other words, if the primary task is a visual one, then the secondary task should also be visual. Another concern is that the secondary task adds to the cognitive load. To reduce this risk, the secondary task has to be very simple (even as it remains both reliable and valid), and it needs to be able flexibly to use up the available free cognitive capacity. A measure that meets these requirements is reaction time to a stimulus, and the learner must react to a secondary task as quickly as possible.

Brünken et al. (2002) found that the reaction time to the secondary task was significantly longer in the case of the visual-only format of instruction than in the case of the audiovisual format. Apart from demonstrating the feasibility of the dual-task method, the results of this study provided additional empirical evidence for the modality effect and the theory of dual-system processing. Differences in performance could be directly aligned with differences in the cognitive load induced by the different presentation formats: the audiovisual format produced better performance and the reaction time to the secondary task was shorter, indicating lower cognitive load.

Brünken, Plass and Leutner (2004) also extended the modality effect and dual-task methodology research to investigate auditory cognitive load. They argued that while the audiovisual format in their previous study (Brünken et al., 2002) was the more effective presentation format, they had used a visual secondary task, and this, in effect, only gave an indication of the processing taking place in the visual sub-system of working memory. The results obtained were therefore a measure of visual cognitive load. They were also of the opinion that research into auditory load had been neglected.

Using an auditory secondary task, they looked at the auditory load under the following three conditions:

- Visual only material with no background sound. (Their prediction in this case was that performance on the auditory secondary task would be the highest for this condition.)
- Visual only material with irrelevant background sounds
- Audiovisual material with irrelevant background sounds. (Their prediction in this case was that performance on the auditory secondary task would be the lowest for this condition.)

These predictions were duly confirmed. The study demonstrated the feasibility of using an auditory stimulus for the secondary task when the primary task included the use of the auditory sub-system of working memory. An interesting finding in this study was that performance in the secondary task did not differ significantly between the condition of visual material/no background music and visual material/background music. It would appear that as long as total cognitive load is kept within the limits of the working memory, the addition of irrelevant background sound to learning material does not produce a detrimental effect on learning. This seems to contradict the findings of Moreno and Mayer (2000b), who concluded that entertaining but irrelevant auditory material in a multimedia message can be detrimental to learning. Differences in the methodologies of the two studies only permitted Brünken and his colleagues to propose a few possible reasons for these different findings. The contradictions that emerged will need further investigation. The moment a narration was added to the learning material, the performance on the secondary task decreased significantly, thereby indicating that once the learner had to process relevant auditory material, a higher load was imposed on the auditory sub-system of working memory. This study also demonstrated that performance was better in the audiovisual format than in a visual-only format, even if an auditory secondary task was added. Once again, the load is divided between the two systems.

Fink and Neubauer (2001) investigated the potential of using the subjective estimation of time as an index of cognitive load. What they found was that as the cognitive load increased, so the subjective estimation of the time duration on task decreased. They also explored this subjective time estimation in participants who registered different levels of intelligence. They found that although the subjective time estimation decreased for both groups as the cognitive load increased, the group with higher intelligence produced more realistic time estimations. They also found that speed of processing was related to intelligence, and concluded that since the more intelligent participants processed more quickly when performing cognitive tasks, they had more attentional capacity left to estimate time compared to the less intelligent participants who processed more slowly and therefore had no spare capacity left to estimate time accurately. These findings raise the question of cognitive load and intelligence. Fink and Neubauer (2005) extended their work to consider explanations other than mental processing speed to explain the relationship between subjective time estimation and intelligence. They looked at whether the number of tasks completed played any role, but could find no evidence for this explanation. They then conducted two different studies in which they gave the

participants tasks that required task processing in working memory and tasks that looked at the speed of processing. They confirmed yet again that subjective time estimation can be used an index of working memory load even though there were also still differences in subjective time estimation that were dependent on intelligence. Although these researchers conclude that there is not sufficient evidence to suggest that the subjective estimation of time is a valid index of task difficulty or cognitive load, I would be cautious in using this technique to measure cognitive load for at least the following three reasons:

- Fink and Neubauer (2001, 2005) found evidence of a relationship between subjective time estimation and intelligence. The cognitive load research stream had, as far as I could determine, not considered the relationship between cognitive load and intelligence in depth, and until such time as there is more clarity on this issue, intelligence could be a confounding variable in an effort to measure the cognitive load of learning material and instructional strategies.
- The tasks used in the studies of Fink and Neubauer are far removed from the kinds of tasks with which learners are confronted in the real world of the classroom. I would like to see more research that uses different kinds of tasks before I, as an instructional designer, embrace this method of measuring cognitive load.
- Speed of processing and accuracy were important elements of the study. The participants were asked to complete the tasks as quickly and accurately as possible. Once again, in the real world of the classroom, speed and accuracy are not always the issue. In the real world of the classroom, understanding, the application of knowledge and learning itself are fundamentally important. I would like to see this research being replicated in environments that resemble the world of the classroom more closely.

In this section, I have reviewed the literature that covers the measurement of cognitive load in some detail. While several techniques were used, and several studies used more than one technique, I did not find any studies that determined whether or not there was any correlation between the different techniques. The question ‘Which technique is the most effective and efficient for use in the authentic learning environment?’ therefore remains largely unanswered. My own study will begin to answer this question by investigating the correlation between a subjective rating scale and the dual-task method of measuring cognitive load.

2.11.4 Learning from multiple representations

If multimedia are to be used effectively within the learning environment, then designers and instructors need answers about the potential of multimedia to influence learning outcomes. Mayer (1997a) in the article, “Multimedia learning: Are we asking the Right Questions?”, refers, on the one hand, to the media debate that was conducted between Clark and Kozma in the early 1990s, but, on the other hand, discusses how the consensus among educational/cognitive psychologists has moved in the direction of calls for reframing the research questions that are applied to media. Multimedia

learning research (Mayer, 1997a) had focused on establishing whether one medium is better than the other (the so-called media effect)(Chang, 2002; Chong, Balen & Jewesson, 2003; Frith, Jaftha, & Prince, 2004; McDonough, & Marks, 2002; Rutherford & Lloyd, 2001; Sinclair, Renshaw, & Taylor, 2004; Stern et al., 2001). While Mayer discussed three main directions for future research, the essence of these directions can be summarised as a call for research that looks at how the instructional design/treatments within various media influence the cognitive processes and consequences.

This call was for research that will consider

- whether or not multimedia learning is effective (the multimedia effect)(Carney & Levin, 2002; Scherly, Roux & Dillenbourg, 2000; Zahn, Barquero & Schwan, 2004),
- when multimedia learning is effective (contiguity effects)(Mayer & Chandler, 2001; Mikk & Luik, 2003; O'Donnell, Dansereau & Hall, 2002; Rieber, Tzeng & Tribble, 2004; Verdi & Kulhavy, 2002), and
- for whom is multimedia learning effective (interaction effects)(Kozma, 2003; Liu, 2004).

A review of multimedia learning research reveals that there is considerable research into the attitudes towards and perceptions about the value of multimedia learning. The students were generally positive toward computer-based instruction that included includes multimedia resources in their design (Jha, Widdowson & Duffy, 2002; Ellis & Cohen, 2001). Most students reported an improved understanding of the content, and because they believed that multimedia added value to the learning experience, they wanted more resources which include multimedia elements. Not all students, however, reported positive experiences of multimedia. Those students who were highly computer literate were not impressed with programs that did not exploit the medium optimally and that were technologically deficient (Trinder, 2002).

There is a large body of research that has looked at using text, images and audio in different combinations and under different circumstances (this research answers the 'When is multimedia learning effective?' question). Quealy (1998) explored the role of delivery media in the simple recall of both declarative and procedural knowledge. The study compared text and still image, text and still image with audio and text, and video and audio. At the time when it was undertaken (1998), the research attempted to address issues of why and when different multimedia techniques could be expected to exert beneficial effects. The results showed that audio and video were more effective than still images and text.

The purpose of the study by Moreno and Mayer (1999) was to clarify and test two cognitive principles: the contiguity principle and the modality principle. The contiguity principle states that learning is enhanced when printed text and pictures are physically integrated or close to each other (spatial contiguity). It also states that learning is enhanced when visual and spoken materials are

temporally synchronised rather than presented successively (temporal contiguity). This principle is also described in the cognitive load literature under the heading of split-attention effect. According to the modality principle, words should be presented as auditory narration rather than as visual on-screen text. The question that they set themselves was: 'In multimedia learning with animations, what are the relative contributions of spatial contiguity and modality to multimedia learning, and what are the relative contributions of temporal contiguity and modality?' They hypothesised that if the advantage of narration over on-screen text resides in a modality principle, then the advantage for auditory-visual presentations should not disappear when the presentations are sequential rather than contiguous. Their first experiment provided support for both the modality and spatial contiguity principles. A second experiment, in which they used sequential presentations, allowed the researchers to distinguish between the modality and the contiguity effects. Their findings provided more evidence for the modality effect than for the temporal-spatial contiguity effect.

Hall (2002) used two approaches to investigate the utilisation of on-line materials within the curriculum. In the one approach, he used learning materials from the web as extra resources, and in the second approach, he used the web to change the educational structure of the course. Hall concluded that stand-alone resource banks do not necessarily support deep-level learning because the educational media were seen in isolation from the rest of the course. When the new media were integrated with other learning resources, there was greater processing of the resources and integration of these resources into the essence of the course.

Mayer and his colleagues have consistently obtained results that indicate that text and images are superior to text alone. Mayer, Mautone and Prothero (2002) investigated the effectiveness of different types of guidance or scaffolding in a discovery-based learning environment by using computer-based simulation. The scaffolding they provided ranged from no guidance, to the provision of pictorial guidance, to verbal descriptions, and to a combination of both verbal and pictorial guidance. The learning task was primarily visual in nature, and Mayer et al. (2002) found that verbal scaffolding did not seem to be effective. Learners who received pictorial scaffolding solved significantly more problems than students who were not provided with pictorial scaffolding. Although overall pictorial modelling seemed to have a positive effect on the learning, they attributed this to the largely visual nature of the task.

More recent research has investigated the effects of adding audio (Moreno & Mayer, 2000a; Moreno & Mayer, 2000b; Mayer, Sobko & Mautone, 2003) to multimedia instruction. Moreno and Mayer (2000a) tested the hypothesis that personalised explanations promote deeper understanding because the learner becomes actively engaged in the elaboration of the materials. They proposed that learners used less cognitive effort to process verbal information when it was presented in a familiar style. In a series of five experiments, they found that deeper learning took place when personalised messages were used. A design guideline that flowed from this research was that the learner should be addressed as a participant rather than an observer. Apart from motivating the inclusion of audio from a dual-processing/cognitive load perspective, additional reasons for

considering adding audio are that a speaker's voice in multimedia lessons carries important social cues that can influence the process and outcome of learning. This promotes the idea that the inclusion of voice adds a social dimension to learning in a computer-based environment, the lack of which in the use of computers for education and training has been occasionally criticised.

Mayer et al. (2003) found that the retention rates for the learning material remained the same, irrespective of whether a foreign accent (a non-American accent in the context of this research) or a standard American accent was used. There were, however, significant differences on the transfer test and the speaker rating, with the non-accented group outperforming the accented group. When comparing the use of a human voice to a machine-generated voice, the study found that the participants using the intervention with the human voice performed better across all the measures (namely, retention, transfer and rating).

Mayer, Fennell, Farmer and Campbell (2004) investigated the use of audio by comparing two styles of voice – the conversational voice and the formal voice. They looked at the effects on learning when they merely changed the 'the' (*student, learner*) to 'you', thereby creating a more conversational style. The theory behind personalisation is that using self as a reference point increases learner interest, and this in turn encourages a learner to use available cognitive capacity (germane load) for the active cognitive processing of the incoming information during learning. Deeper processing results in more meaningful learning, and this is indicated by better transfer test performance. The content used narrated animation as a strategy to explain how the respiratory system works. While Mayer et al. (2004) found that performance in the transfer test across three experiments was significantly greater for the personalised group than for the non-personalised group, there were no significant differences between the groups as far as their interest and difficulty ratings were concerned. Interest is a very individual concept, and this finding is not surprising if one considers that the participants were Psychology students who were learning content that was not really relevant to their own field of study. This use of participants to learn content completely unrelated to their own learning programme is a methodological issue that needs to be addressed in multimedia research. Mayer et al. (2004) did in fact acknowledge that the participants might have rated the content as 'not interesting' because they perceived that it as being irrelevant and extraneous to their studies. They also rated the difficulty, using Paas's scale, as less difficult for the human-voice condition.

Moreno and Mayer (2000b) investigated the use of other sounds in multimedia such as background music and sounds. There are two theoretical positions in this regard: the one position claims that adding audio makes the learning task more interesting and increases the learner's enthusiasm. The other position holds that the addition of audio can overload the auditory processing channel. The corollary of the second position is that sound that is not required to make the lesson intelligible or that is not integrated with the rest of the material will reduce the effective working memory capacity and interfere with learning. Moreno and Mayer devised two experiments whose results demonstrated that when they added unnecessary background music that was irrelevant to the content, the performance of participants on recall and problem-solving transfer tests dropped significantly. One of their

experiments showed that relevant environmental sounds that have been integrated into the learning material facilitate processing and exert no negative influence on learning. They were not, however, able to replicate this finding in the second experiment. These findings are consistent in general with accumulated knowledge about the limitations of working memory and the cognitive load and dual-processing theories.

Spickard et al. (2004) determined the impact of adding an audio-feed to an online lecture on post intervention test scores. There was a trend among students who had received the audio version of the lecture to show higher posttest scores than those students in the group who had not been given the sound. It was also demonstrated that those students who had received it were satisfied with the audio-format and that they spent more time using the material than those who had only been given access to the non-audio format.

Studies have also been undertaken that suggest that certain combinations of media do not facilitate the achievement of learning outcomes. Mikk and Luik (2003) investigated the characteristics of multimedia textbooks that affect posttest scores. What was found was that a high percentage of three-dimensional graphics reduced posttest scores, and that illustrations and graphics presented with redundant text also reduced posttest scores. These findings seem to indicate that not all media facilitate learning. They also seem to show that students could learn as well from good graphics that stand alone in their ability to illustrate a concept as they would from a passage of text that describes the concept. These findings reinforce the questions about the best combination of media.

In their introduction to a special edition of the journal *Learning and Instruction*, Schnotz and Lowe (2003) considered the conditions under which multiple representations foster learning, and noted that there seemed to be the assumption that multimedia rich learning environments would automatically result in extensive cognitive processing that would, in turn, lead to the creation of elaborate knowledge structures. They expressed a concern that technologies were being pushed to the limit and that instruction often included multiple media representations that were not necessarily always beneficial for learning. They expressed the view that there was a need to understand exactly how specific features of multimedia are able to help learners to learn complex subject matter.

Schnotz and Bannert (2003) proposed an alternative cognitive model of multimedia learning. Their critique of Mayer's model was that it assumed a one-to-one mapping between the text and picture images in the working memory. They argue that text and pictures use different sign systems and that these result in different forms of representation, which they refer to as 'descriptive' and 'depictive' representations. Each form of representation is used in different ways for different purposes. They were of the opinion that mental models are not sensory specific, but are far more abstract. They explained this by proposing that while a mental model can contain less information than the original image because irrelevant details are left out, it can also contain more information because it also includes prior knowledge information. The specific form of visualisation used in an image can affect the structure of the resulting mental model and the model's computational efficiency for specific tasks.

Picture comprehension is therefore a process of structure mapping between a graphic surface representation and a mental model representation. A designer needs to decide which visualisation is best suited to the circumstances (which includes the learning task that needs to be accomplished). Their study considered the influence of the form of visualisation on the mental model structure. Their hypothesis in this regard was that pictures with task-appropriate visualisation would support mental model construction and that pictures with task inappropriate visualisation would interfere with mental model construction. The task in the experiment was to work out time differences and answer questions about circumnavigation by making use of three different formats of content. The participants in the intervention were homogenous with regard to their prior knowledge, verbal ability and spatial ability. For the time difference task, the text-only group performed the best. The circumnavigation task showed that there was a significant main effect for the kind of task ($p = 0.029$), and a significant interaction effect between the kind of representation and the kind of task ($p < 0.001$). Schnotz and Bannert (2003) found no evidence for the dual-coding hypothesis. They concluded that their findings supported their structure interference hypothesis, namely, that pictures with task inappropriate visualisation may interfere with mental model construction. They concluded that pictures facilitate learning only if individuals have low prior knowledge and if the subject matter is visualised in a task-appropriate way. Adding pictures to text may therefore not always support learning.

As an instructional designer I have often been involved in debate about the use of multiple representations within the same program. Such practice is often defended with claims along the lines of 'It accommodates different learning styles', or 'It is a good thing for learners to see the topic from different perspectives' or 'Repetition reinforces learning'. I have already discussed the issue of redundancy (the cognitive load perspective) in this chapter. A study undertaken by Kozma (2003) looks at the issue from a multimedia learning perspective. Kozma looked at the material features of external, multiple representations and the cognitive and social affordances they provide in support of understanding chemistry. He also looked at the different ways in which scientists and students used these multiple representations. In a laboratory, experimental setting Kozma (2003) found that experts were able to cluster apparently dissimilar problems or situations into large meaningful groups on the basis of underlying principles. He found that expert chemists were able to use conceptual terms to label their clusters and that they also used a greater variety of representations in their clusters. In contrast to this, he found the novices (the students) labelled their subject groupings on the basis of surface features. Another task that participants were required to perform was to transform each representation into another form. They would have to, for example, transform animation into a corresponding graph and a video into an equation. The experts were significantly better at this than novices, who tended to use the surface features of the displays in attempts to build an understanding of the chemical phenomena they represented. Kozma then replicated this research in a naturalistic setting. Observation of the environment in which the chemists worked revealed that representations of chemistry were everywhere in the laboratory: there were, for example, diagrams on whiteboards and posters on the walls. One of his observations was how frequently the chemists used these diagrams to explain a point or to seek clarification while solving problems. These experts were able to

make effective use of different representations to understand the chemical phenomenon that they were examining. They used these representations to support their arguments and to explain and justify their reasoning. The students, on the other hand, did not function in this way at all. Kozma was of the opinion that the study conducted in the naturalistic setting confirmed the findings of the laboratory experiment.

Kozma proposed the following three design principles to guide the use of multiple representations:

- Provide at least one representational system that has features that explicitly correlate to the entities and processes that underlie physical phenomena. The processing capabilities of computers can be used to give material substance to these phenomena.
- Use multiple representations in the context of collaborative, authentic laboratory sessions, but make sure they are explicitly linked to one another so that it is easier for the student to make the connection.
- Give the students collaborative tasks which require them to generate representations and use them to confirm and explain their findings.

Seufert's (2003) review of the literature reported on several studies that indicated that while the use of multiple representations was hypothesised to have synergistic effects on knowledge construction, such synergy did not happen automatically. Seufert investigated the use of different types of help that would assist learners to form a coherent picture of the content while using representations that both complemented and constrained one another (one representation is used to understand another one). She proposed the following two kinds of help:

- Directive help, which provides explicit hints about the elements and relations that are relevant to each representation.
- Non-directive help, which enables learners to discover what they need to know in a self-directed manner.

Seufert's study was designed to investigate how these types of help influence coherence formation in learners who have different levels of prior knowledge. The first research question looked at the relative value of these two kinds of help. Directive help was found to be significantly superior to non-directive help for recall tasks, while there was no significant difference with regard to the different types of help when it came to comprehension tasks. When she looked at the influence of prior knowledge on performance, Seufert found that no help and directive help were more useful for low prior-knowledge learners, and that non-directive help exerted a detrimental effect on performance. She also found that learners with a medium level of prior knowledge benefited the most from directive help, which in fact seemed to be the best type of help to provide all round for recall tasks. Seufert (2003) concluded that the type and amount of help to give depends on the learning goal. Directive help seemed to support recall performance because of its summarising and repeating function. Directive help was also more effective for comprehension, contrary to the expected finding that non-

directive help would be more effective. Prior knowledge also mediated the effects of help. Low prior knowledge learners, for example, cannot use help effectively at any point: their need is for additional instruction on the content. A notable observation made during this study was that the high knowledge group, who did not seem to need help, did not reach the maximum level of performance. Seufert describes this as the 'illusion of knowing'. It means that learners with high prior knowledge overestimate their abilities and do not seem to realise that help can enable them to perform even better. She recommended that instructors might need to introduce help to this group in another way – and to point out to them that help could very well enable them to improve their performance.

In his commentary in this special issue of *Learning and Instruction*, Reimann (2003) highlighted the important instructional design guidelines that followed from the empirical work that had been reported. He also commented on the direction that multimedia research had been taking and that it should take in the future. Four of the seven studies in this issue focused on the modifying effects of prior knowledge on the use of multiple representations (Kozma, 2003; Lowe, 2003; Schnotz & Bannert, 2003; Seufert, 2003), and two of the studies investigated the use of animation (Lowe, 2003; Lewalter, 2003). There are three important themes in this research, which need further investigation, but which also provide thought for some new direction. These themes are concerned with:

- The role of external representations, how they are transformed into internal mental models, and the strategies that need to be employed to foster this internalisation.
- The continued significance of the dual-coding theory, which really only addresses the issue of modality but which fails to provide direction on some other important issues such as:
 - ♦ the kinds of representations that need to be developed in order to visualise something, and
 - ♦ the sequencing of representations.
- The effect of animation on learning, the circumstances under which animation should be used, and principles that guide the design of animation.

Reimann encouraged the research community to look forward, so that multimedia research did not fall into the same trap that man-interaction research did.

...once psychologists knew almost all about the psychology of the command-drive interface, there were no command-driver user interfaces in use anymore

(Reimann, 2003, pg 251).

Van der Meij and de Jong (2006) investigated the conditions that provided the best support for learners who use multiple presentations in a dynamic simulation-based learning environment. This study aimed at extending the work of previous research which had investigated the integration of images and text (Chandler & Sweller, 1991) or images, text and audio (Mayer & Moreno, 1998; Tabbers, et al. 2004). Since the learning environment had also been designed to be more complex,

simulation was being used as an instructional strategy rather than an interactive tutorial. These researchers compared the strategies of integration and dynamic linking. The simulation interface had anywhere between three and five representations, depending on the format. Integration combined the five different representations. In the dynamic linking condition, actions performed on one representation were shown simultaneously in the other representations.

There were three different representation formats:

- separate, non-linked format
- separate, dynamically linked format
- integrated, dynamically linked format

Participants who used the integrated, dynamically linked format scored significantly better ($p < .05$) than the separate, non-linked presentations on test of domain knowledge. Participants who used the integrated, dynamically linked format found the domain easier than those who used the separate, non-linked presentations format. The benefits of integration and dynamic linking did not extend to transfer learning. The authors concluded that this could have been due to the fact that the time spent on learning the material was very short. Van der Meij & de Jong (2006) also found that the group that used the integrated, dynamically linked format scored significantly better ($p < .05$) than the other groups on the high complexity domain items in the posttest. These findings were different from those of Bodemer, Ploetzner, Feuerlein, & Spada (2004), who found significant differences for low complexity domains.

2.11.5 Animations, including animated pedagogical agents

Since the use of animation is becoming widespread, multimedia and cognitive load research are including this strategy in their research stream.

Chandler summed it up well:

.....despite this seemingly endless potential and unbridled enthusiasm for technology based instruction, there is little empirical evidence to indicate that the widespread use of dynamic visualisations has resulted in any substantial benefit to learners.

(Chandler, 2004, page 353).

In this section, I will provide an overview of the most recent research that investigates the use of animation in learning.

Animation has several uses. It can support 3-D perception by showing an object from different perspectives; it directs attention to important aspects of a display; it conveys procedural knowledge; it demonstrates the dynamics of the subject matter; it enables exploratory learning through the

manipulation of a displayed object. Animation can help a learner to perform a cognitive process that they could not otherwise have performed without this kind of external support (Schnotz & Rasch, 2005). These researchers are of the opinion that the conditions under which animated pictures really enhance comprehension and learning is still an open question.

Mayer and Anderson (1992) explored the use of animation in instruction as early as 1992, in a study that focused on the conditions under which animations were effective for learning. They looked at the contiguity principle which states: given the limits on working memory, learners may make the necessary connections more easily when words and pictures are presented contiguously. They found that while animation per se did not necessarily improve understanding, there was an improvement in problem-solving transfer when the contiguity principle was applied to the design. Mayer and Anderson were cautious in their conclusion and pointed out that learner characteristics (novice learners with no prior knowledge), type of instructional material (demonstration of how a system works), and the final assessment (problem solving and retention), all influence the outcome of any study. They made the point that other conditions may not provide the same results. Cognitive style, for example, is another learner characteristic that could be considered.

ChanLin (1999) compared control strategies (user-controlled versus system controlled) for learners with different cognitive styles when learning included animation. Using the Field Independence (FI)–Field Dependence (FD) classification of style, ChanLin's overall finding was that the field independent students were able to learn better with animation than were the field dependent students. Students in the self-controlled group performed better than those in the system-controlled group. Performance was significant for the treatment but not for style, and there was no interaction between treatment and style. ChanLin also looked at the strategies that students with different styles adopted, and found that there was a difference in the performance of students who used different styles. Those who used a field dependent style work their way through the whole lesson first in order to obtain an overall picture, and they then returned for second and third reviews of the problems. Field independent students, by contrast, tended to stay at a specific point for a long period in time, and often replayed the animated visual over and over again when the system controlled the animation. I used this finding to guide the design of my study and placed the animations under user control.

In 2002 Mayer and Moreno (2002) reported that there was consensus among media researchers that the potential of animation to promote learning was not yet clear. The benefits of using animation were still largely dependent on how instructors used animation in the learning environment. They concluded that pursuing answers to questions about the media effect (questions such as, 'Will using a multimedia CD-ROM for learning produce better outcomes than attending lectures?') was a waste of time. They called for researchers rather to explore the conditions under which various media, such as animation, affect the learning process. When and how does animation (or any other media) affect learning? They laid down seven research-based principles for the design of multimedia presentations that use animation, on the basis of a decade of research that examined the conditions under which animation promotes learner understanding: the multimedia principle, the spatial contiguity principle,

the temporal contiguity principle, the coherence principle, the modality principle, the redundancy principle and the personalisation principle (Mayer & Moreno, 2002).

Lewalter (2003) used think aloud protocols to investigate the learning processes and strategies that take place when learners used either static images or animation.

She identified the following three categories of strategy:

- **Rehearsal strategies:** Use techniques such as memorising by means of recitation and recapping.
- **Elaboration strategies:** Build connections between new information and prior knowledge.
- **Control strategies:** Plan and regulate further learning, and control the actual level of comprehension.

Lewalter found that while rehearsal strategies were used most frequently by all the groups, they were used to a significantly greater extent by the static visuals group. And while the use of elaboration strategies was low, this was attributed to the fact that the participants possessed very low prior knowledge. Lewalter found that while the use of visuals improved learning outcomes in comparison to the outcomes obtained from the use of text-only presentation, there was no difference in performance of participants using static images and animation versions. She concluded that arrows, which are conventional symbols for motion, and a series of images, may be sufficient if the goal is to obtain factual knowledge. One of my concerns with Lewalter's study is that the design created a split-attention effect because text and illustrations (both static and dynamic) were presented on separate pages. Since there is clear evidence that the split-attention effect is detrimental to learning (Kalyuga, Chandler & Sweller, 1999; Sweller, 2006c), this might well have influenced the results.

Lowe (2003) examined what learners actually extracted from an animation. He used animation in his study to assist learners to make weather predictions from static weather maps. Since their knowledge base had been found to be insufficient, it was thought that the use of animation would strengthen their mental models of the changes that occur in weather patterns over time and that they would therefore be able to make better predictions when confronted later with only static weather maps. One of Lowe's findings was that novices looked at the features of the animation that were perceptually dominant even though they were not meteorologically important. They were also able to extract information more easily when there was a substantial dynamic contrast with regard to form and position as the animation progressed. But he found that the learners did not extract enough of the subtle information that was important for building a good mental model of weather maps. In order to prevent this superficial extraction of information, Lowe recommended designers to include the use of specific cues to assist the learner to focus on important information. He also recommended that additional instructional enrichment be provided rather than the mere provision of an animation that displays the process.

In 2004, the journal *'Learning and Instruction'* produced a special issue that was devoted to the use of dynamic visualisations (animation and video) and learning. This issue reported on the early beginnings in developing a research-based understanding of the perceptual and cognitive processes involved in learning with dynamic visualisations.

In their introduction to this issue, Ploetzner and Lowe (2004) raised the concern that since it has become relatively easy to include dynamic visualisations in multimedia learning environments, doing so can be justified on the basis of the assumption that such visualisations will automatically improve the learner's understanding. They cautioned that the view that dynamic content should be presented dynamically is a very simplistic assumption. It is possible that dynamic visualisations require even greater information processing on the part of the learner than do static visualisations. Learners need to process large amounts of information that change quickly, and the extent to which animation can be made interactive varies greatly. In its most elementary form, learners can play, stop, rewind and replay a sequence of visualisations. In its most complex form, it is possible to change parameters and data sets even while the visualisation is running or interactively to construct additional visual components. This may demand the use of additional cognitive processes and skills such as planning and decision-making. Ploetzner and Lowe point out that these demands have not been empirically explored in sufficient depth to provide guidance in the form of principles about how to design and use dynamic visualisations.

Ainsworth and van Labeke (2004) described how time needs to be portrayed in dynamic representation, and, in so doing, they provided a glimpse of the complex issues that need to be considered before deciding whether or not animation is the most effective strategy. There are multiple ways to represent information which changes over time, and there are specific cognitive tasks associated with each. In any animation screen, images are transient, and information needs to be held in memory if the learner needs it for making interpretations later. Ainsworth and van Labeke (2004) suggested the following reasons why learning with animations had produced mixed results:

- Learners may focus on the more obvious conceptual events rather than on those that are conceptually more important (Lowe, 2003).
- Activities are shown in a particular sequence, even if that sequence is not relevant for learning.
- Animation does not always allow the learner to retain sufficient control.

They propose that there are three types of dynamic representation, and each has distinct informational and computational properties and different uses. **Time-persistent** animations display a relation between at least one variable and time. For example, in a health care context a table comparing the incidence of HIV/AIDS and Tuberculosis (TB) over a number of years could be animated. This would provide a visual representation of the increases and/or decreases of HIV/AIDS and TB over a specific time period and how they might or might not be related to each other. This is similar to static images except that the dynamic presentation displays the data incrementally rather than presenting the whole set at once as in a static display. While no new information is added

because of the animation, it might make the features of that information more salient. This type of animation displays the greatest amount of information, and this can place a high cognitive load on the learner. While **time-implicit** animations (which are often used in simulations) show a range of values, they do not show the time when these occurred unless you are looking at the animation. Using the same example, a graph could be created with HIV/AIDS on one axis and TB on the other. The information appearing in this graph is then animated, using time as the variable that determines the sequence of the visual display in the animation. When the simulation has been stopped, the learner needs to invoke internal representations to compare current values to a particular previous state in order to answer questions about the timescale. **Time-singular** animations display one or more variables at a single instant in time. For example, a pie chart could be created to show the relative proportions of HIV/AIDS and TB for each year. The series of pie charts is then animated, but when the animation is stopped the learner only sees the representation for one time period and must hold the other views in memory if any comparison is needed. This can place a high load on the cognitive resources of the learner.

The underlying requirement for learning with complex animation is that the learner must actually extract certain information from the animation and incorporate it into an existing schema. The typical changes in an animation over time include changes in form, changes in position and the appearance or disappearance of elements.

Lowe (2004) extended his earlier work (Lowe, 2003) that looked at how learners extracted information from animation. In this later study, Lowe explored how participants interrogated an interactive animated display in order to extract the information required to complete a prediction task. What he found was that there was considerable variation in how the participants interrogated the animation, and that those who performed the learning task poorly did not go back to the animation as often as the more successful learners. Participants were very selective about which part of the animation they chose to look at. Most of them went through the animation once to obtain an overview of the content, and then they went back to selected segments in order to complete the task. It appeared that the participants limited their interrogation of the animation in order to make a complex and demanding task more manageable. In both studies (Lowe, 2003, 2004) Lowe found that the learners tended to use low-level strategies that addressed isolated temporal and spatial aspects of the animation. They tended to explore one feature at a time and looked more at changes in position than changes in form.

Evidence is accumulating that learners' extraction of information from an animation can be deficient in at least two ways: firstly, learners engage in under-processing because of a lack of prior knowledge, and, secondly, in spite of intense engagement with the animation, learners engage in under-processing because the cognitive demands are simply too high. The subject matter is often complex and the temporally distributed nature of the information increases the load. What seems to happen in these circumstances is that learners then apply their attention to a subset of information. Clarification

of the relationship between form and position becomes an important issue for the designers of animations.

Bodemer, Ploetzner, Feuerlein and Spada (2004) reported that dynamic visualisations are usually presented in combination with other kinds of representations such as text and static visuals. While these may complement one another, it is possible that learners will **not** systematically relate these to each other and integrate them into one coherent whole. During the process of problem-solving, learners may switch back and forth between these multiple representations, and not come to any conclusion about which one is most helpful for solving the problem. Learners often need support to cope with the specific requirements of these visualisations. This particular study looked at how the design of visualisation could both reduce cognitive load and increase germane load by manipulating various combinations of instructional strategies. In the first experiment, the content was concerned with how a tire pump works. It was, in fact, similar to the content used by Mayer in previous experiments (Mayer, 1997a). Two factors in the experiment were manipulated: the format of representation (split source versus integrated) and learner activities (mental integration versus external and mental integration). The results indicated that learning with the integrated format was more successful than learning with the spit-source format. This finding replicated the work of Sweller and Chandler (Chandler & Sweller, 1991; Sweller & Chandler, 1994). Actively producing an integrated format appeared to be more successful than merely reconstructing an already integrated format. Since the observed differences between the groups were small and since the task was relatively easy, the researchers set up a second experiment that made use of a more complex and difficult task with content that dealt with statistical concepts. The research design manipulated three formats of content (non-integrated, integrated and active integration) and two types of interaction (free exploratory versus structured interaction) with the dynamic visualisation phase of the learning. The participants first tackled the content with the aid of static visualisations before using the dynamic visualisation to engage in the learning task. This study found that learners who actively integrated different representations outperformed those who were provided with split-source or pre-integrated formats. The structured testing of hypotheses by the participants in the dynamic visualisations were more beneficial with respect to verbal understanding than use of the free exploratory technique. There was no statistically significant interaction between the two factors: integration of information and structuring of interactions.

In a commentary at the end of this special issue of *Learning and Instruction*, which focused on new and potentially viable approaches to dynamic visualisations, Chandler (2004) called for a better research base to explain how people cognitively process and learn from such resources. He was critical of a practice followed by many designers, namely, using the technology to generate the learning experience rather than using their growing knowledge of cognitive processes to guide them in how best to use the technology. He was of the opinion that any approach to instruction that ignored cognitive processes was likely to be deficient.

Hegarty (2004), in a second commentary at the end of this special edition, pointed out that although it seemed intuitive that there should be an advantage of dynamic over static media, the first wave of research (which was not reported in the 2004 special issue) showed no clear advantages accruing from dynamic displays. While these strategies certainly increased motivation and levels of student interest, they exerted no significant effect on conceptual learning. In the 2004 special issue, Hegarty discussed the following three factors that she identified from the research:

- The need to go beyond the simplistic categorisation of static and dynamic displays. This need arises because there are many different kinds of dynamic display. She sounded a note of warning that cautioned the reader not to assume that the results of one study can be applied to all situations that use dynamic displays.
- The fact that dynamic visualisations place heavier demands on the human cognition than do static images.
- The need to assist learners to develop metacognitive skills that will enable them to actually use the interactivity provided in certain types of dynamic visualisations.

Hegarty pointed to two areas that she felt needed more attention:

- An analysis of the task that needs to be learned, followed by a mapping of the aspects of the task that can be taught effectively by means of dynamic visualisation.
- Studies that determine the extent to which learners internally visualise the content of the external visualisation. If little or no internal visualisation takes place, it might be time to question whether the expense of this development is justified.

Schnotz and Rasch (2005) explored the enabling, facilitating and inhibiting effects of animations in multimedia learning. Within the cognitive load framework, an animation can have two functions. Firstly, if animation reduces cognitive load by allowing cognitive processing that would otherwise be impossible, then it has an enabling function. Secondly, if animation reduces the cognitive load of tasks that could otherwise only be solved with high mental effort, then it has a facilitating function. Schnotz and Rasch (2005) conducted two experiments that analysed how these assumed functions of animations affected cognitive processing and learning results. In the first study they compared learning from animation to learning by means of static images. They made a general assumption that learners would learn more from animated pictures than from static pictures, albeit for different reasons, depending on their learning prerequisites. They devised two formats for the content that utilised animation and static images respectively. The content itself dealt with the topics of time and date differences and the consequences of this when circumnavigating the globe. The study used a pretest-posttest design and the participants were given unlimited time in which to study the learning material. A learning prerequisite score was calculated for each participant and the group was then divided into high and low prerequisite learners. For learners with high prerequisites (high cognitive ability and high prior knowledge), animations seemed to have an enabling rather than a facilitating

function. For learners with low learning prerequisites, animation seemed to have a facilitating rather than an enabling function.

There were two types of questions in the study: questions relating to time-difference and questions relating to circumnavigation. The results for the time difference questions supported the hypothesis that learners would learn more easily from animation than from static pictures. The results for the circumnavigation questions were unexpected: the learners obtained better results when they studied by making use of the static pictures. Schnotz and Rasch (2005) looked at the cognitive processing required to answer each type of question and concluded that the facilitating function of animation exerted a negative effect on learning in this case. They suggested that the external support made the cognitive processing so easy that the learners spent less mental effort learning from the animation than they did when learning with the assistance of the static pictures. In this case the animation unnecessarily reduced the germane cognitive load associated with deeper more meaningful cognitive processing. Since the researchers found that they could still not distinguish between different effects of different kinds of animation, they devised a second study in which they compared different kinds of animations. Their findings in the second study were once again that certain types of animation can exert a less-than-beneficial effect on learning because the processing involved is too easy. These researchers were not, however, prepared to recommend that animation be replaced with static images. Their argument was that because it is all relative and because low prerequisite learners may never even try to perform mental simulations, the animation can be helpful because anything is better than no learning at all – even in cases where the processing is easy.

Chan and Black (2005) proposed a format-support hypothesis of learning, which was then tested in a study that compared the performance of learners when using either static images, a system – controlled animation or a version where the user could manipulate the animation. The rationale behind this hypothesis was that comprehension and learning would be enhanced if the presentation format supported the learner's need to construct a dynamic mental model of a particular phenomenon. They argued that the benefits of using animation could only be tested and ultimately realised if there was a match between the format of the animation and the learning outcome. The learning task in their study required an understanding of the dynamic interchange between kinetic energy, potential energy and total energy. The target group was 7th grade students. They used a roller coaster ride as the context for explaining the phenomenon. There were three assessment tasks: written recall, create a diagram of the phenomenon and problem-solving tasks. Participants who used the user-controlled / manipulation version of the animation performed significantly better on the recall and drawing tasks than the groups using the other versions. Participants who used the static visual version performed significantly better than the groups using the other versions on the transfer tasks. These findings might seem surprising. Why would the group who used static images outperform the groups using animation? The explanation for this finding is that when learners use a static visual they must animate the process mentally. This requires deeper cognitive processing and engagement with the content, which helps develop a solid understanding of the content. Similar findings have been obtained in several other studies (Garg et.al. 1999, Schnotz & Rasch,2005). These findings are also

similar to those of Hegarty, Kriz and Cate (2003), who found no evidence of better understanding of a dynamic process after the use of animated diagrams compared to static diagrams.

Zahn, Barquero and Schwan (2004) investigated the conditions under which different forms of interactive video, another form of dynamic visualisation, are beneficial for learning. These researchers report on apparently conflicting design guidelines, with multimedia research calling for hyperlinks to be integrated sequentially in the video (spatial and temporal contiguity) and hypermedia research calling for hyperlinks to be presented as clusters at the end of the video. The question of how many links to add also needed investigation. Their study investigated these inconsistencies. The content was about lakes as ecosystems. Two formats of hyperlinks were compared: links presented sequentially during the video versus links presented in clusters at the end of each section. The strategies the participants used when reviewing the videos were also captured in a log file. Contrary to expectations none of the analyses yielded a significant effect. Neither the position of the link in the video (sequential versus clustered) or the number of links had a different impact on learner's knowledge acquisition. The researchers did find that the longer the participants engaged in reading the text in the links the more their knowledge increased. More frequent use of the dynamic links also resulted in an increase in knowledge acquisition for the text that was reviewed. Zahn et. al. (2004) explained that the fact that learners had sufficient time to study the video, and were given considerable control over how they approached the learning task were most likely the major contributors to the unexpected findings. The success of learning from hyperlinked video may therefore rely more on the learner's ability to actively manage complex information structures than on the design principles themselves. The influence of user strategies may override the influence of single design variables.

2.11.6 In summary

This section has examined cognitive load and multimedia learning under the following themes:

- Instructional design issues
- Theory development and directions of research
- Measurement of cognitive load
- Learning from multiple representations
- Animations, including animated pedagogical agents

The first theme (Instructional design issues) focused on the effects identified in the cognitive load research stream that are relevant to my study. There is solid evidence that design which forces the learner to divide his/her attention between sources of information that are not integrated, but which need to be, leads to poorer learning performance. This holds true for both the design of paper-based learning material and the computer-based multimedia environment. The modality effect was introduced and investigated as a alternative strategy aimed at reducing the negative effects of split-attention. The use of both audio and text has proved to be successful in reducing cognitive load.

Behavioural learning theory, which emphasises repetition, was put under the spotlight when the cognitive load researchers proposed that redundancy increased cognitive load. When the expertise-reversal effect was described the researchers took a new look at redundancy. The research demonstrated that the more expert learner experienced the negative effects of redundancy more than did the novice learner. The expertise-reversal effect resulted in work that has considered the role of prior knowledge in more detail, not only for the role it plays in expertise reversal, but also in split-attention conditions. Knowledge about the expertise-reversal effect has been used in creating rapid assessments that drive adaptive learning environments, particularly in the domain of mathematics.

Cognitive load theorists have always proposed that the prime purpose of CLT was to generate new and useful instructional techniques. Many of the early criticisms of the cognitive load theory have been addressed over the past 25 years as researchers have used the theory to test a variety of instructional strategies and techniques. The theory has not changed much since its conceptualisation in the late 1980s and early 1990s. Cognitive load has been categorised into intrinsic, extraneous and germane load. The early research was focused on seeking ways to reduce extraneous cognitive load. Extensive research has been conducted on the use of worked-examples in problem-solving, which have proved to be successful strategy to use for novice learners. More recent research is now considering ways to increase germane load, which is thought to contribute to the development of sound mental models and schemas. Research has considered techniques such as fading, different forms of help and support and different feedback strategies. Initially the theory proposed that intrinsic load could not be manipulated by design, but it appears as if this idea is being challenged in some of the recent research. The research has also moved from the experimental laboratory setting to the more authentic environment, which is characterised by longer, more complex courses. I could find only one attempt in the literature to re-consider the theory. This has been a call to consider the role of instructional goals and strategies and their influence on cognitive load.

The third theme considered the measurement of cognitive load. The two most researched methods are the self-report method developed by Paas and the direct method using a dual-task approach developed and researched by Brünken and his colleagues. The easiest and least invasive method to implement is the self-report rating method. Neuroimaging, a more recent technique, has been proposed as being more accurate, but cannot be easily used in everyday learning environments. Very few studies have compared the findings of more than one measurement method using the same learning material. The self-report rating method was found to be more reliable than the physiological method that used heart-rate variability. Flowing out of this measurement research is the concept of instructional efficiency and the literature now describes three different ways of determining this instructional efficiency.

The last two themes considered learning from multiple representations and animation. There has been extensive research into the use of text, images and audio, in various combinations, with solid evidence that designs using the principles of dual-coding result in superior learning performance. Multiple representations can also result in cognitive overload. The more expert the learner the better

able they are in the use of multiple representations, both in the experimental laboratory and naturalistic setting. The findings regarding the use of animation as an alternative to static images & text are still mixed. The evidence for the superiority of animation over static images and text is not overwhelming. Learners are generally unable to extract the important information from an animation, especially if their prior knowledge is poor. They often tend to focus on the superficial qualities of the animation. This poses a design challenge to ensure that there is nothing in the animation that distracts the learner. Researchers still query whether the benefits of using animation justifies the cost of development and production.

2.12 Multimedia in health sciences education

This literature review will conclude with a section that considers the use of multimedia in health sciences education in general and the use of multimedia in Physiology education in particular.

A subject field that seems to benefit from the use of visual and dynamic illustration of content is the health sciences. Since subjects such as Anatomy and Physiology are visually rich, the ability to make an effective clinical diagnosis depends on cognitive and psychomotor skills, including the ability to observe small changes in physical appearance. Multimedia that uses case studies and simulation includes photographs of patients (DxR Development Group, Inc., 2005) and other visual material such as digitised images that enable students to see macro images of clinical features and X-rays. Multimedia resources are now also increasingly available as supplements to textbooks or on the web (Gómez-Arbonés et al., 2004; Harden, 2002; Health Education Assets Library, (n.d.); The Virtual Labs Project, 2003).

Many of the subjects that form part of the Health Sciences curriculum use computer-assisted instruction that makes use of multimedia presentation. These include:

- Anatomy and Physiology (Garg, et al., 1999; Lewis, 2003).
- Basic Sciences (Issenberg & Scalese, 2004).
- Clinical Skills (Issenberg et al., 2002; Stern et al., 2001).
- Community-based Medicine (Sturmburg, Crowe & Hughes, 2003).
- Dentistry (Aly, Elen & Willems, 2004; Schultze-Mosgau, Zielinski & Jürgenf, 2004).
- Dermatology (Morton, Foreman, Goede, Bezzant & Albertine, 2007).
- Emergency Medicine (Westendorp & McGraw, 2002).
- Oncology (Hulsman, Ros, Winnubst & Bensing, 2002).
- Psychiatry (McDonough, & Marks, 2002; Williams, Aubin, Harkin & Cottrell, 2001).
- Paediatrics (Treadwell, de Witt & Grobler, 2002).
- Rheumatology Nursing (Vivekananda-Schmidt, Hassell & McLean, 2004).

- Surgery (Jha, Widdowson & Duffy, 2002; Kneebone & ApSimon, 2001; Lynch, Steele, Palensky, Lacy & Duffy, 2001; Steele, Palensky, Lynch, Lacy & Duffy, 2002).

I analysed all of the studies listed above as well as others by making use of Mayer's (1997a) four effects (media, multimedia, contiguity and interaction) as categories for analysis. I also looked at studies that addressed perceptions of users attitudes towards multimedia instruction and meta-analyses. My analysis of a selection of the empirical studies is summarised in Tables 2.7 to 2.12. I have deliberately summarised these in tabular format so that I would be able to include as many studies as possible and so present a bird's-eye-view of the research field. This kind of format makes it easier to compare studies than does a text composed of narrative presentation and critique. Each section of analysis is followed by a critical review of the research.

2.12.1 The media effect

The media effect addresses the question, 'Which medium is effective for learning?'

* indicates that this study also addressed perceptions and attitudes

Discipline & Intervention	Study	About the study	Findings
Clinical skills in cardiology (auscultation of heart sounds) Intervention included a program with 3 multimedia case histories plus follow-up seminars or a CD-ROM with 20 mini cases.	Issenberg et al., 2002	It compared clinical bedside teaching with and without deliberate practice using a simulator and multimedia tutorials.	Improvement between pretest and posttest was significantly larger in a relatively short time for the group who had been exposed to the simulation and simulation/multimedia material.
	Stern et al., 2001	It compared clinical rotation only (control groups) to clinical rotation plus exposure to the multimedia.	The group that received clinical rotation plus 20-case mini series performed significantly better than the control group. Additional clinical teaching together with 3 longer case studies improved knowledge but not auscultatory skills. While the knowledge had diminished 1 year after the intervention, the clinical skills had not.
Undergraduate orthodontic curriculum Intervention was computer-assisted multimedia modules.	Aly, Elen & Willems, 2004	It compared the intervention with traditional lectures for their effectiveness in improving knowledge, understanding, transfer of content and problem-solving skills.	While both groups improved their scores in posttest, there was no significant difference between the two groups in relation to answers to questions about knowledge, understanding and application.
Psychiatry	McDonough, & Marks,	It compared face-to-face instruction in small groups with	Students who attended the tutorial scored slightly better in the posttest than those who used the CAI. The

Discipline & Intervention	Study	About the study	Findings
	2002 [*]	computer-based instruction.	tutorial group enjoyed the teaching significantly more than the CAI group.
	Williams, Aubin, Harkin & Cottrell, 2001 [*]	It compared teaching in a lecture with a computer-assisted program that taught the same content within the same time frame.	There was no difference between the two groups with regard to knowledge assessment. Students who used the CD-ROM showed significantly higher acquisitions of skills (which were measured with the use of an objective assessment method) even though the lecture group rated themselves higher on knowledge and skill acquisition than those who used the CAI. Perceptions still existed that CAI is not as effective as traditional methods.

Table 2.5: The media effect in health sciences multimedia education

Table 2.5 shows that there is ongoing research into the media effect in the health sciences, in spite of the exhortations of Mayer and Moreno (2002) to researchers to move on and to look at the conditions under which various media, such as animation, affect the learning process. That the health science tertiary education field is under-researched is evidenced by the very few studies that I was able to report in the previous sections of this chapter. Like the other fields that look at the media effect, the reported studies have produced mixed findings. While some studies have reported significant improvements in learning that utilises a technological medium (Issenberg et al., 2002, Stern et al., 2001), others have reported the all-too-common 'no significant difference' that is regularly found in media effect research (Aly, Elen & Willems, 2004). Studies also often combine looking at learning outcomes with a survey of perceptions and attitudes towards computer-based instruction (CBT). An interesting observation is that in several studies that combined CBT with traditional approaches, there was a definite improvement in learning (Issenberg et al., 2002, Stern et al., 2001). I am in no way suggesting that CBT should replace traditional instruction in cases where better learning was achieved with the CBT and traditional instruction together. The findings of Issenberg et al. and Stern et al. provide evidence for the approach that utilises a combination of suitable and appropriate media within the learning environment.

2.12.2 The multimedia effect

The multimedia effect addresses the question, 'Is multimedia learning effective or not?'

* indicates that this study also addressed perceptions and attitudes

Discipline	Study	About the study	Findings
Oral and maxillo-facial surgery	Schultze-Mosgau, Zielinski & Jürgen, 2004	This study developed 27 virtual multimedia lectures that covered selected topics. The study was followed up with online assessment.	97% of the students passed the online assessment. This may be compared to the previous pass rate of 85% that followed tuition by means of traditional lectures alone.
Emergency medicine & radiography Intervention was a multimedia module on interpretation of wrist radiographs.	Westendorp & McGraw, 2002	This study used a convenience sample of 36 volunteers across the education spectrum (students, residents). The effectiveness of the module was evaluated by comparing the radiographic interpretation skills of students who had used the module with students who had not used it.	The scores of the students were similar to those of the residents who already had the skills and received the training. The scores of the students who completed the multimedia module were also better than those students who had NOT completed the multimedia module.
Oncology Intervention was a CAI programme on communication skills.	Hulsman, Ros, Winnubst & Bensing, 2002	This study assessed the communication skills of participants at four points: T1 – The base line T2 – Control period T3 – After intervention T4 – To determine retention over time. Patient satisfaction was also assessed.	Results indicated that the intervention seemed promising because it was known to be an effective strategy for improving communication skills. The authors were cautious in their interpretation of the results. Motivation appeared to have played a role for those who made effective use of the multimedia lessons. But patient satisfaction did not seem to improve (this was contrary to expectations).
Paediatrics Intervention was CD-ROM followed by practice in a simulated laboratory.	* Treadwell, de Witt & Grobler, 2002	This study compared how students acquired 14 skills after partaking in the new intervention with those who in the past had only been involved in clinical rotations. The study assessed the students' perceptions of a new instructional strategy.	Those who used the new intervention performed as well in the final skills assessment as those who had followed the older approach (except for 1 of the 14 skills). This difference was not significant. In 2 of the skills, the students performed significantly better than those who had followed the older curriculum.
Surgery Intervention was CAI	* Lynch, Steele, Palensky,	The aim of the study was to determine if learning preferences and attitudes	There was no evidence to suggest that learning preferences and attitudes towards computers

Discipline	Study	About the study	Findings
program on angiogram.	Lacy & Duffy, 2001	toward computers had influenced the acquisition of knowledge during the use of the CAI program.	and the CAI program had influenced the learning outcomes. There were significant improvements in knowledge after using the program, and this improvement was sustained after 6 to 8 weeks.
Interventions – CAL to teach knowledge and simulation to teach skills.	Kneebone & ApSimon, 2001	This study observed the individual teaching and the group sessions and interviewed participants about the usefulness of the strategies.	This study found that all participants enjoyed the approach. It also found that group teaching followed by individual practice was the preferred approach. The CAL was used in the teaching sessions. Because all the students wanted more time for study and practice, the CAL was made available in a self-directed format.

Table 2.6: The multimedia effect in health sciences multimedia education

All the studies summarised in Table 2.6 seem to indicate that, in the context of the particular studies undertaken, multimedia learning was effective. Schultze-Mosgau, Zielinski and Jürgen (2004) did not report on a direct comparison of methods and it was not possible to scrutinise the questions that they asked. These questions led them to draw their conclusion that the use of multimedia produced superior learning. I am therefore cautious about affirming that ALL the studies demonstrated a multimedia effect. Lynch et al. (2001) found that students scored high on the concrete and teacher-structured learning scales of their learning preferences instrument. Students seemed to prefer to learn in situations where they were given clear and specific instructions for completing practical tasks and skills. While these authors reported that their findings were consistent with previous research findings, a review of their list of references reveals that the studies to which they were referring were all conducted in the 1970s and 1980s. Their conclusions would have been more useful to the field if they had reviewed more recent research as well. There is certainly plenty of it available.

2.12.3 The Contiguity effect

The contiguity effect addresses the question, 'When is multimedia learning effective?'

Discipline	Study	About the study	Findings
Anatomy	Garg, Norman, Spero & Taylor, 1999	This study assessed carpal bone spatial ability by making use of a 36-item test.	Familiarity with more views made no difference to student performance. When they were briefed, the students remembered only certain key views, and then rotated the image mentally so that they could answer questions presented from the non-standard views. The authors concluded that it was possible that the multiple views were redundant and/or confusing.
	Garg, Norman & Spero, 2001	This study assessed carpal bone spatial ability by making use of A 50-item test.	In this study there was a significant difference in spatial carpal ability in favour of the key view format. Spatial ability also affected performance. Students also first considered the key view and then mentally rotated the view before answering the question. Those who used the multiple view model spent more time looking at the key views in this model than the other views.

Table 2.7: The contiguity effect in health sciences multimedia education

There is currently an alarming lack of research into the contiguity effect and other the related effects described in the multimedia and cognitive load research literature. If learner characteristics, type of instructional material and type of test/assessment really do influence the outcome of any study (Mayer & Anderson, 1992), then this is where health science education field should be focusing its research effort. They should be replicating the studies from the general multimedia learning, cognitive style and cognitive load research streams and opening up new avenues of research by challenging findings and by providing new challenges. The work of Garg, Norman, Spero and Taylor (1999, 2001) is more in line with the direction that contemporary research in this field should be taking. There has certainly been enough development in the technology concerned to justify widespread use and application. One need only think of the developments that have taken place in more advanced technologies such as 3D animation, 3D modelling, simulation and virtual reality (Donnon, DesCôteaux & Violato, 2005; Jha, Widdowson & Duffy, 2002; Mantovani et al., 2003).

2.12.4 Interaction effect

The interaction effect addresses the question, 'For whom is multimedia learning effective?'

Discipline	Study	About the study	Findings
Gynaecological surgery Intervention was interactive CAI on how to perform a vaginal hysterectomy.	Jha, Widdowson & Duffy, 2002	The aim of study was to pilot the program, establish who would benefit from this CAI, and to evaluate the multimedia in terms of content, interactive features and outcome.	While the CAI was useful for students and specialists in training, the suggestions from the specialists in training on how to improve the program indicated that the programs needed more depth in the depiction of anatomy and level of detail. Both groups were of the opinion that their understanding had been improved.

Table 2.8: The interaction effect in health sciences multimedia education

There are a few publications in the health science education field that ask the question, 'For whom is multimedia instruction effective?' at the level suggested by Mayer (1997a). The study of Jha et al. (2002) used a self-report instrument to gather the data. There was no investigation into learning performance across the two groups. Other variables such as cognitive style and prior knowledge were not included in the analysis. If one compares the study to similar research, one may judge it to be a rather superficial attempt to answer the question, 'For whom is multimedia instruction effective?'

2.12.5 Perceptions and attitudes

NOTE: Italic text indicates author's emphasis.

Discipline	Study	About the study	Findings
Basic Sciences	Issenberg & Scalese, 2004	This study considered the use of computer-based assessments, which tested integration of basic science knowledge and clinical skills, and making use of multimedia stems.	This study assessed attitudes towards this form of assessment. 93% of the respondents indicated that they found this format to be user-friendly.
Oral and maxillo-facial surgery	Schultze-Mosgau, Zielinski & Jürgen, 2004	This study developed 27 virtual multimedia lectures that covered selected topics. This was followed up with an online assessment.	75% of the students indicated that they considered this method to be superior to traditional teaching.
Dermatology	Morton, Foreman,	This study developed 4 eBooks using TK3 for use	Students regarded the eBook as an effective way for distributing course

Discipline	Study	About the study	Findings
	Goede, Bezzant & Albertine, 2007	in a dermatology course for second-year students. It also carried out a qualitative assessment of the usefulness and effectiveness of this methodology when compared to using paper-based lecture notes.	content. While they enjoyed using the eBook as a self-study tool, they preferred to take paper-based notes during a lecture instead of using the eBook (which could also be used for note-taking).
Neuroscience	Brueckner & Traurig, 2002	This study investigated the extent to which students accepted and used the digital neuroscience guide that had replaced the paper-based guide.	Students used the guide primarily as a study tool at home. There were significant gender differences in usage patterns, with male students using it as an independent study tool and female students using it for independent and group study. All of the students agreed that it was an efficient study tool.
Surgery Intervention CAI program about angiogram for surgery rotation.	Steele, Palensky, Lynch, Lacy & Duffy, 2002	This study explored the relationships between learning preferences (<i>instrument more aligned to assessing learning style than cognitive style</i>), attitudes towards computers and the students' evaluation of the CAI program.	Learning preferences and attitudes towards computers had no significant influence on the attitudes about the program. Even though students evaluated the program positively, they still had reservations about the place of CAI in medical education, and expressed the fear that CAI would replace the student-teacher relationship.

Table 2.9: Perceptions of and attitudes towards multimedia education in the health sciences

Since there are numerous studies that look at attitudes and perceptions, the list in Table 2.9 is by no means exhaustive. It merely provides a broad picture of the range of scope. Five subject disciplines have been listed here. I have excluded studies from Nursing Science and the other allied health sciences. The media included CD-ROMs, e-books, digital study guides, and virtual lectures. An interesting finding from the study by Morton et al. (2007) was that only half the participants saw the availability of images in the eBook as an advantage. Even though this might indicate a cognitive style preference, they did not investigate it in their study. While most students are generally positive about using the technology, even the recent studies (Treadwell et al. 2002) report that participants have reservations (and perhaps even fears?) about the new technologies and strategies replacing the traditional lecture-based approach that is so deeply entrenched in medical education.

2.12.6 Meta-analyses

Discipline	Study	About the study	Findings
Anatomy & Physiology	Lewis, 2003	This was a meta-analysis of studies that investigated the utility of computer-assisted instruction (CAI) for teaching Anatomy and Physiology in subject allied to medicine.	Only 10 studies met the criteria for inclusion. Most of the studies reported an improvement in performance for those students who used CAI when compared with the performance of those students who used more traditional methods of learning.

Table 2.10: Meta-analyses of research into multimedia education in the health sciences

I could find no other meta-analyses in the health science education literature that were relevant to my study. The analysis of Lewis (2003) in reality only addressed the media and multimedia effect.

The large volume of research into media effects, student perceptions, attitudes and the evaluation of usefulness might well indicate the ongoing ambivalence that exists toward CAI in health science education (Brueckner & Traurig, 2004). Even though there is a great need for research that examines contiguity and interaction effects, these studies cannot be ignored because attitudes and perceptions are indicative of motivation and we know that that motivation plays a definite role in learning (Paas et al., 2005).

The section provided a broader view of the status of health science educational research in general. In the following section, I will narrow the focus to look at multimedia, cognitive style and cognitive load research in Physiology education.

2.12.7 Multimedia in Physiology Education

Sefton (1998) illustrated how complex the field of Physiology is.

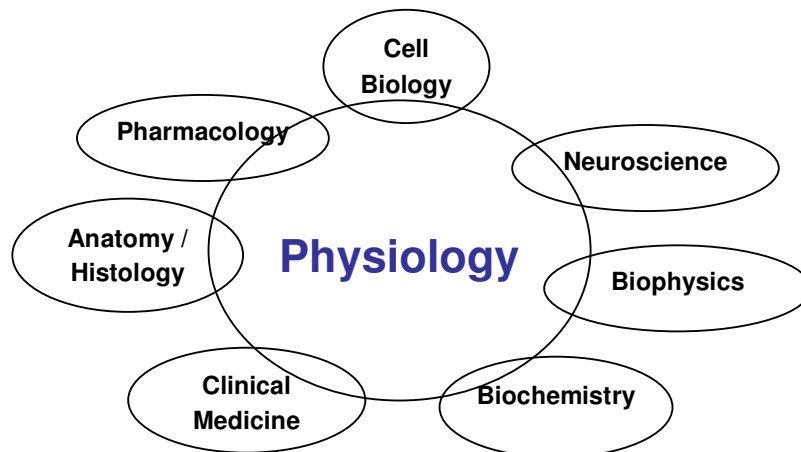


Figure 2.10: Relationships of physiology with related Medical Science (Sefton, 1998, pg. 54)

Sefton (1998) outlined some of the global challenges in Physiology education. The one issue she addressed that is relevant to my study is the use of computers in Physiology education. While Sefton was of the opinion that the early promise of computer-assisted learning in Physiology education had not been generally realised by 1998, she acknowledged that the use of computers could be effective provided that the educational goal – and not the technology – determined the direction. In 1998 Sefton had already noted that because the explosion of resources on the Internet could have a profound impact on Physiology education, it was necessary to devise skills development programmes for Physiology lecturers so that they would be able to search for relevant and appropriate materials and integrate those materials into their existing courses. Finding such resources has now become relatively easy. But deciding on whether or not they will facilitate learning, and whether or not multimedia instruction is in fact effective, is more difficult. This suggests the need for a sustained empirical research programme in the field of multimedia education in Physiology. My study seeks to make one small contribution to this field.

In 2007, Michael (2007) surveyed faculty to determine their opinions about what makes Physiology hard to learn. He first asked the participants for an open-ended response to the question, 'What do you think makes physiology hard to learn?', and then followed this question with a series of questions that were grouped under the following three category headings:

- The nature of the discipline
- The way in which the discipline is taught
- The issues that students bring to the learning environment (prior knowledge, study skills)

Michael reported that the majority of responses to the open-ended question could be mapped to one of these three factors. Respondents indicated that the nature of the discipline and the nature of the student contributed more to the fact that the students found the subject hard to learn than the actual teaching methods that they used. The top five factors that Michael identified were:

- Learning requires the ability to reason casually.
- Students believe that learning is the same as memorizing.
- Learning presupposes the ability to think about dynamic systems.
- Understanding is communicated graphically or in other mathematical ways.
- The content of the learning needs to be understood at a number of different organisational levels simultaneously.

Both Rawson and Quinlan (2002) and Griffin (2003) address the nature of the subject and the requirement that students have to integrate multiple sources of information in order to understand some of the concepts and processes in Physiology. Another interesting factor in Michael's 2007 study (Michael 2007), but lower in the ranking, was the fact that students tended to ignore the graphs, tables and figures in their learning material. Is this indicative of a Verbaliser profile (which would focus on using textual rather than visual material) in the student group, or does it point to a lack of

appropriate study skills? I did not find any reference to the cognitive styles of students in the Physiology education literature, and a discussion of study skills is beyond the scope of this review.

Although Michael (2007) asked some very interesting questions, he could not find the answers in the health science education literature. He asked, 'Is there objective data about the difficulties of learning a science?', and concluded that there is very little data that directly addresses this issue. Cognitive load literature does, however, address the issue. If the factors listed above are correct, they point to a subject that appears to have high intrinsic cognitive load, and Michael would find some of this objective data by measuring the cognitive load of the learning material. There is ongoing empirical research (and this review has referenced many of the published results) that is building the knowledge base of effective instructional design strategies that aim to reduce cognitive load and facilitate learning. These guidelines, that aim to minimise cognitive load, are as relevant to Physiology education as they are for the subject domains in which the research was conducted.

Michael (2007) also discusses the issue that students find the transfer of learning difficult. He reviewed some of the studies that indicated that students have serious misconceptions about the phenomena that they encounter, that students need help in causal reasoning, and that they misinterpret visual material. Michael felt that these factors indicated that serious attention needs to be paid to the instructional strategies that are being used to teach to teach Physiology – in spite of the fact that the faculty that he surveyed did not think that the way in which they taught actually contributed to the fact that their students experienced Physiology as a 'difficult' subject.

The theme of my study considers the use of multimedia as a teaching tool and strategy. There is evidence that Physiology textbooks are visually rich because Physiology textbooks not only make substantial use of static illustrations, but they also include CD-ROMs that make extensive use of animations and more dynamic media. There is a need to investigate the quality of these products and the impact they make on learning outcomes. Is the inability of students to correctly interpret graphs, figures and tables a cognitive load effect or a cognitive style effect, or both? And how should such content be designed and presented so that learning will be more effective? My study will compare animation and the use of static images.

A review of the recent empirical research (1997-2006) in Physiology education is summarised in Table 2.11.

Research stream	Researcher	Participants	Subject area in the field
Approaches to studying – no multimedia involved	Abraham, Kamath, Upadhya & Ramnarayan, 2006	1 st year medical students (n=223)	Endocrine, renal and reproductive Physiology
Effectiveness of computer-assisted learning and evaluation of the usefulness and design of the program	Azer, 2005	1 st year medical students (n=106)	Bile salts and bilirubin metabolism
	McGrath, Kucera & Smith, 2003	Students taking Human Bioscience 1. Their numbers included students from medical disciplines (n = 1044).	Neuron membrane potentials
	Rawson & Quinlan, 2002	1 st year veterinary science students	Acid-base physiology
	Kukulja Taradi & Taradi, 2004	2 nd year medical students	Several topics in Physiology and Immunology
Comparison of multimedia and lecture as methodologies	Buzzell, Chamberlain & Pintauro, 2002	Mixed group (n = 32)	Human body composition analysis
Attitudes to, perceptions of, and general use of technology-based resources	Davis, Wythe, Rozum & Gore, 1997	1 st year medical students	Various topics used on the Web such as CD-ROM resources and online self-assessments

Table 2.11: Review of recent research in Physiology education

There are several publications that describe only the use of technology in Physiology education (Brann & Sloop, 2006; Dwyer, Fleming, Randall & Cohen, 1997; Griffin, 2003). The pedagogical motivations for using technology for teaching and learning in Physiology include:

- Simulations can be designed that include temporal elements that demonstrate processes that develop over time and that cannot be understood by means of the rote memorisation of facts.
- Dynamic events can be explained by means of simple mechanisms.
- The computational capabilities of the computer make it possible to demonstrate concepts and processes to the learner, without them needing to understand the complicated mathematical

calculations behind the process. The learning objective in these cases is understanding of the process itself – and not the Mathematics behind it.

- Strong visual images that support the learning process can be presented.
- Because of the complexity of body systems, real-life experiments do not always turn out as expected, and this can be very confusing for novice learners. A simulation can simplify the content and focus on essential details rather than on peripherals (McGrath, Kucera & Smith, 2003).
- Technology is useful for correcting misconceptions.
- It is far more expensive to stage laboratory experiments than it is to demonstrate such experiments by means of computer simulation (McGrath, Kucera & Smith, 2003).

There is an interesting trend in the use of web-based courses which includes discussion groups about the content to supplement traditional lectures in Physiology (Davis, Wythe, Rozum & Gore, 1997; Kukulja Taradi & Taradi, 2004; Kukulja Taradi, Taradi, Radić & Pocrjac, 2005). This might seem strange as one tends to think of Physiology as a subject that consists mainly of factual knowledge that has to be learned. While this is certainly the case, the increasing use of discussion as an adjunct to understanding the subject domain supports the methodological assumptions of a problem-based curriculum. In a PBL curriculum, the basic sciences, including Physiology, are integrated into the clinical disciplines, and this approach is far more successful than the traditional approaches that first teach the basic sciences and then only the clinical sciences.

Not one of the studies reviewed in this section makes any reference to the role that cognitive load and cognitive style play in Physiology education. In the study of Davis, Wythe, Rozum and Gore (1997), one of the comments about the usefulness of the strategy indicated that cognitive style played a role in that participant's use of the resources, but since it was only one of many comments that supported the qualitative analysis of the findings, its significance was lost in the plethora of other comments.

While it is still possible to understand the lack of focus on cognitive load and style, what I do not understand (and this is cause for concern for me) is the finding that not one of the studies reviewed in Table 2.11 (or those listed in the previous paragraph) used a theoretical framework or model from the cognitive science field (or any other field for that matter) to guide the research. Not one of these studies so much as referred to the substantial body of research in multimedia education that has contributed so much to our knowledge of what constitutes sound instructional design and an appropriate and effective use of technology.

While calls continue to be made for research-based evidence for the need to change teaching and learning practices in health science education (Michael, 2006), the evidence is there and it has been extensively and thoroughly discussed in this literature review. Health science educators need to look beyond their own field of education and take note of this evidence. Health science education needs to

move beyond research into students' attitudes toward and perceptions of multimedia programmes (see Table 2.9), and it needs to focus more on finding effective instructional strategies that meet two important challenges: firstly, that of a field where the knowledge base that is growing faster than people can keep up with, and, secondly, the life-long learning needs of the health science practitioner.

The issues that Michael (2007) addresses in his study suggest to me a powerful argument for extending the cognitive load research stream into the tertiary health science education field, and into Physiology in particular. There seems to be no research that uses cognitive theories of learning as a research framework.

2.12.8 In summary

This section reviewed the use of multimedia in health sciences education, including Physiology education. Multimedia is used in a large number of the disciplines in the wider health science education field. The field seems to be dominated by media effects research, and the findings are typical of this type of research, in that many studies report the 'no significant difference' finding. Most of the studies that considered the question from the multimedia effect perspective concluded that, in the context of the particular studies undertaken, multimedia learning was effective. There has been limited research that considers the question 'When is multimedia effective?' The research considers some of the more advanced techniques offered by technology, for example the ability to rotate images in 3D. Results indicated that the learners used the simpler views and mentally rotated and manipulated the images rather than using the computer-generated views. There is a large body of research that investigates perceptions and attitudes to computers in education. While attitudes are generally positive, many learners in the health sciences disciplines still have reservations about the use of computers as tools for learning.

Some of the challenges in Physiology education include the fact that the intrinsic load of the content is often high. Learners are required to integrate multiple processes and concepts in order to understand the subject. Computer-enabled learning environments are becoming more common, but the research is very narrowly focused within the health science education discipline. Researchers would do well to look to other disciplines for some direction for the effective design of these learning environments.

2.13 Conclusion following the literature view

In this chapter I have discussed my review, analysis and critique of numerous studies. I have made use of Mayer's (1997a) four effects (media, multimedia, contiguity and interaction) as categories for analysis when reviewing the health science education research. I detect a logical progression in these effects, and would like to propose that the research maturity of the field can be determined by noticing where in this progression the field is concentrating its research focus.

My own observation after reading close on 600 publications over the past three years is that early research in a discipline often starts by looking at the media effect. In view of the mixed findings in this area, and my own alignment with the recommendations of the expert researchers (Mayer & Moreno, 2002), the next logical area would be to consider the multimedia effect, which, in my mind, provides direction for the position that can be encapsulated in the comment, ‘Well, if we are going to use technology, we had better make sure that it is effective.’ Once this question has been answered, the research needs to progress to asking, ‘When is multimedia learning effective?’. We could then extend this last question to include the questions, ‘Why is it effective’ (which would lead us to look at the processing demands/cognitive load) and, ‘For whom is multimedia learning effective?’, which would lead us to examine issues surrounding individual differences.

The literature and the research findings suggest that there might be an relationship between cognitive load and cognitive style. While the possibility of overloading the processing capacity of the working memory is inherent in all multimedia learning, we need to ask whether this is a characteristic of the media only or whether individual human differences in processing information also play a role. Then there is the matter of choice. Current technology has enabled choices which did not exist ten years ago. Learners like to be able to choose, and the cognitive style of learners can influence the choices that they make. Will the choices that learners make result in cognitive overload? How can we design the correct combination of media so that such a combination will support individual choices and needs – especially if there is a relationship between cognitive load and style?

How do these concepts relate to the causal and assessment factors that make up the cognitive load construct? I have not been able to find any research in the literature that set out to explore the link between cognitive load and cognitive style that is illustrated in the cognitive load theory. Paas et al. (2003) report that research in the late 1990s took a more dynamic line and that it provides an opportunity for researchers to consider cognitive load as a property of the task-subject interaction – which is open to instructional control. This line of thought aligns with the overall aim of this study.

In conclusion to this literature survey I now point out how it has guided the formulation of the research questions. The research question is 'What is the relationship between cognitive style and cognitive load as factors in the achievement of learning outcomes when someone learns the same content by means of different multimedia formats?'

I drew this question	from the literature reported in
What were the cognitive styles of the participants who took part in the study?	Section 2.7 and 2.10
How did the participants rate the cognitive load of selected multimedia content?	Section 2.11 in general and section 2.11.3 in particular
What was the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?	Section 2.11.3
To what extent did the presentation formats influence cognitive load?	Section 2.10.5, Section 2.11.2, Section 2.11.4, Section 2.11.5 and Section 2.12
How was learning performance influenced when content with different cognitive load was studied by learners with different cognitive styles?	Section 2.10.4, Section 2.10.6, Section 2.11.4 and Section 2.12.

Table 2.12: The literature review and the research questions



Research Methodology and Design

3

3.1 Introduction

This chapter describes the research methodology and design of this study. The research methodology is illustrated in Figure 3.1. The research problem determines whether a quantitative or qualitative approach should be used. The overall research approach informs the design and the design determines the nature of the data. Understanding the nature of the data influences the decision about which instruments need to be developed for data collection. This methodology needs to be constantly interrogated for validity and reliability.

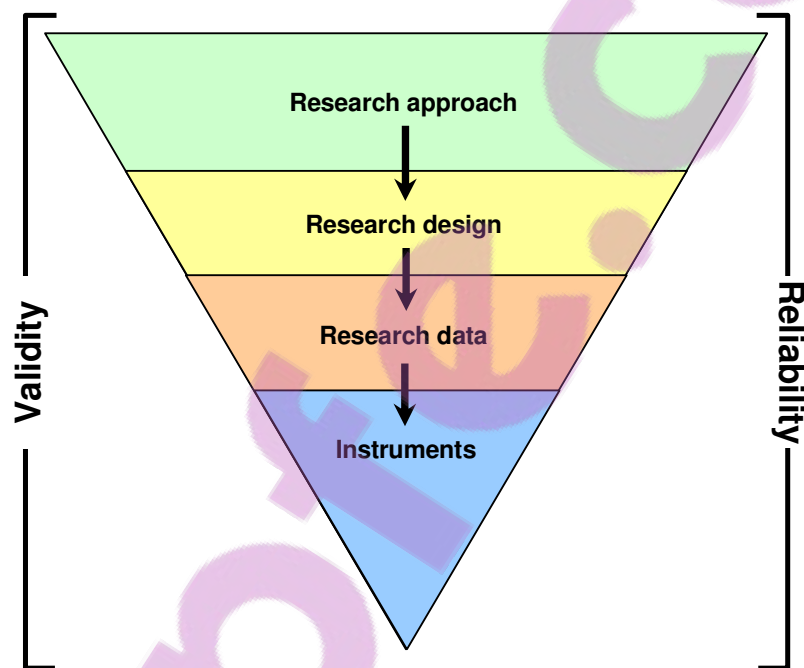


Figure 3.1: A visual representation of the research methodology

3.2 Overview of this chapter

This chapter has four parts:

- **Part 1** presents the research design and methodology. This section describes how each of the component parts of the research illustrated in Figure 3.1 were designed for this study.
- **Part 2** provides a brief description of the design of the research intervention.
- **Part 3** discusses the expected findings and hypotheses.
- **Part 4** describes how the research design and methodology were implemented in both the pilot and the main studies.

Table 3.1 presents a more detailed outline of the sections into which Parts 1, 2, 3 and 4 were divided.

Part	Section	Focus of the Section
Part 1 - Planning the Design	3.3	Purpose of the study
	3.4	The research questions
	3.5	The research approach
	3.6	The research design
	3.7	The research sample
	3.8	The research data
	3.9	The research instruments
	3.10	Ethical considerations in the study
	3.11	Summary of Part 1
Part 2 – Design of the Intervention	3.12	Introduction
	3.13	Source of the content
	3.14	Design and development of the multimedia
	3.15	Instructional strategies and media used
Part 3: The Hypotheses	3.16	The hypotheses
Part 4 - Implementing the Design	3.17	Introduction
	3.18	Conducting the pilot study
	3.19	Conducting the main study
	3.20	Summary

Table 3.1: Detailed outline of Chapter 3

Part 1 – Planning the Design

3.3 Purpose of the study

The purpose of this study has been stated in Chapter 1, Section 1.2, page 1. A critical review of this purpose statement informed the design and implementation of the approach that I took in order to achieve the research purpose.

‘Exploring the role that cognitive load plays’ requires the researcher to be in a position to measure the cognitive load of different multimedia formats. There are several accepted methods for measuring cognitive load (Brünken, Plass & Leutner, 2003; Paas, Tuovinen, Tabbers & Van Gerven, 2003; Paas et al., 1994b). The research that supports these methods has already been discussed in Chapter 2. A review of these methods inevitably posed the question: ‘Will two different methods of measuring cognitive load produce similar results for the same content?’ Because of this, it was

necessary for me to establish whether the measures correlated with each other and to make a decision about which of the measures were reliable and useful.

In this study I used only one method. Smith (2007) used the second method. In order to answer the question posed in the previous paragraph I describe both methods in Section 3.9 of this chapter.

‘Exploring the role that cognitive style plays’ cannot be undertaken unless the researcher is able to measure the cognitive style of the research participants.

The **‘successful achievement of learning outcomes’** implies some gain and/or improvement in knowledge and skill acquisition as a result of the research intervention. Once this knowledge has been properly assessed and analysed, it is possible to decide whether or not a learning gain has taken place. ‘Learning gain’ in the context of this study is measured by means of pre- and post-intervention assessment.

The ability to determine cognitive load is extremely important for an instructional designer because instructional designers can use the empirical evidence obtained from such explorations to guide them to make empirically efficient design decisions. Design considerations include, among other things, selecting the best possible presentational format for specific content.

One of the most important things that instructional designers need to take into account is the distinctive cognitive load of each format. In order to be in a position to do this, it is necessary first to measure and compare the cognitive loads of the available formats. This study measured the cognitive loads of two multimedia programs: one that used mainly **narrated animation** and the other that used **static images and text**. The SAME content was taught in both formats. The questions that I asked included the following:

- What is the cognitive load of each of the formats and is any difference in the measured cognitive load statistically significant?
- Which format achieves better learning results?
- To what extent does the cognitive load of the format influence learning gain?

It was also one of the goals of this study to investigate the relationship between cognitive load and cognitive style.

During my investigation of cognitive style, I asked the following question:

- To what extent do learners with different cognitive styles achieve the learning outcomes when they use different formats, with different cognitive load, to master the same content?

Obtaining an answer to this question required me to compare the learning performance and achievement of the learning outcomes for all the selected style, cognitive load and format combinations.

The literature review in Chapter 2 highlighted the call for research to be conducted in more **authentic learning environments**, as opposed to the kind of cognitive load research that takes place in controlled experimental environments, where the study often uses smaller samples and is short in duration. In my capacity as an instructional designer I developed a multimedia program for health science students that would complement an existing face-to-face course in Physiology. It became obvious to me as I pondered alternatives for the design of the learning materials that it would be best to present content of this kind by means of a variety of multimedia strategies. I therefore developed a section of the content by using two alternative formats: narrated animation or static images & text. I then obtained permission to use this authentic learning environment for the study (See Appendix B). The time frame was authentic because the students had been allocated a certain amount of time in their course schedule to use multimedia resources, and I was able to slot my intervention into this pre-existing schedule. The learning material was also authentic because it is part of the knowledge base that health practitioners need to have before they can make clinical diagnoses or prescribe appropriate treatment regimes (both of which represent complex cognitive skills).

Once the experimental session and the data collection had been completed, the following data became available:

- Cognitive load of the formats
- Cognitive style of the participants
- Pre- and posttest scores from the assessment

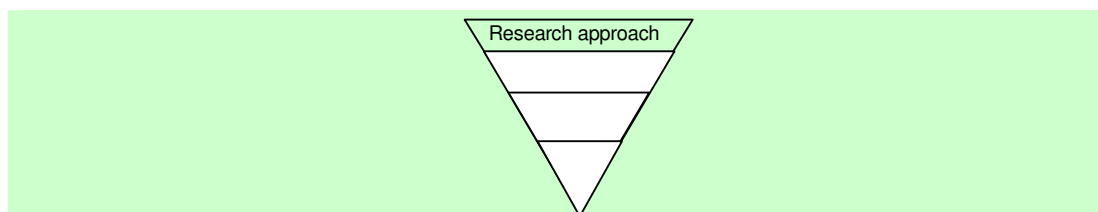
I then used these results to answer a final question: What is the nature of the relationship between cognitive load and style?

3.4 The research questions

Since I have already undertaken a literature review in Chapter 2 and analysed my purpose statement in detail in Section 3.3 of this thesis, I will now re-state the final set of research questions. I will then move on to describe and discuss the details of the research design and methodology.

Question: What is the relationship between cognitive style and cognitive load as factors in the achievement of learning outcomes when someone learns the same content by means of different multimedia formats?

- i. What were the cognitive styles of the participants who took part in the study?
- i. How did the participants rate the cognitive load of selected multimedia content?
- ii. What was the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?
- iii. To what extent did the presentation formats influence cognitive load?
- iv. How was learning performance influenced when content with different cognitive load was studied by learners with different cognitive styles?

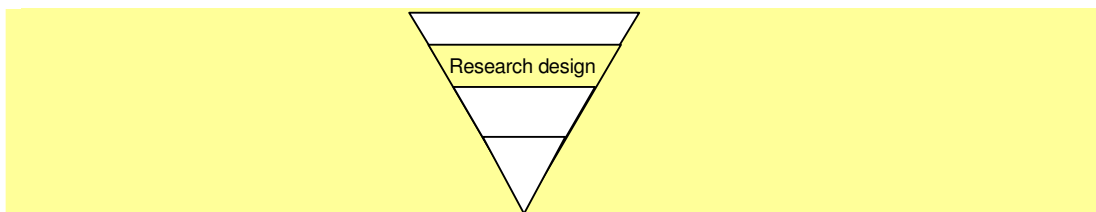


3.5 The research approach

Because this is a quantitative study, it demonstrated the following characteristics of quantitative research:

- It explained relationships among variables.
- The design of the study and data collection was preceded by a substantial review of the literature, which sought justification for the study and which provided direction for the final research questions and design.
- The final questions were very specific and narrow. This made it possible to obtain measurable and observable data.
- Deliberate efforts were made to isolate and control a few variables to study.
- The instruments were identified and designed before the study commenced.
- Data was collected from as large a number of participants as possible (Creswell, 2002).

The variables under investigation in this study were cognitive style (independent variable), cognitive load (independent variable), presentation format (independent variable) and achievement of learning outcomes (dependent variable).



3.6 The research design

This study used an experimental and a correlation design. This can be asserted because the study aimed to determine whether and how a particular intervention (multimedia learning with animation and images) influenced the learning outcomes of a group of students, and because it undertook to investigate whether a relationship existed between two factors which could have made an impact on the learning outcomes (the two factors being cognitive load and cognitive style).

This study therefore used two of the three most common quantitative research designs, namely:

- **experimental designs** – used to test whether an educational practice or idea makes a difference to individuals;
- **correlation designs** – a process of examining the association or relationship of one or more variables by various statistical procedures, and
- **survey designs** – by means of which one administers a survey or questionnaire to a small group of people (a sample) to identify and describe trends in a large group of people (the population) (Creswell, 2005).

The cognitive load of the multimedia formats used in this study were specifically designed to be different. This indicated that an experimental design would be an appropriate choice for this study. Since the study also explored the relationship between cognitive load and cognitive style, a correlation design was also an appropriate choice.

It was not possible deliberately to manipulate cognitive style since

...cognitive style is understood to be an individual's preferred and habitual approach to organising and representing information

(Riding & Rayner, 1998, p. 15).

I did, however, expect to find that the distribution of cognitive styles across the sample would be sufficiently adequate to allow me to divide the sample into groups on the basis of style so that the number in each group would be large enough for statistical analysis.

Figure 3.2 (below) presents a broad outline of the study design. This figure illustrates how the two designs were used and how they are related.

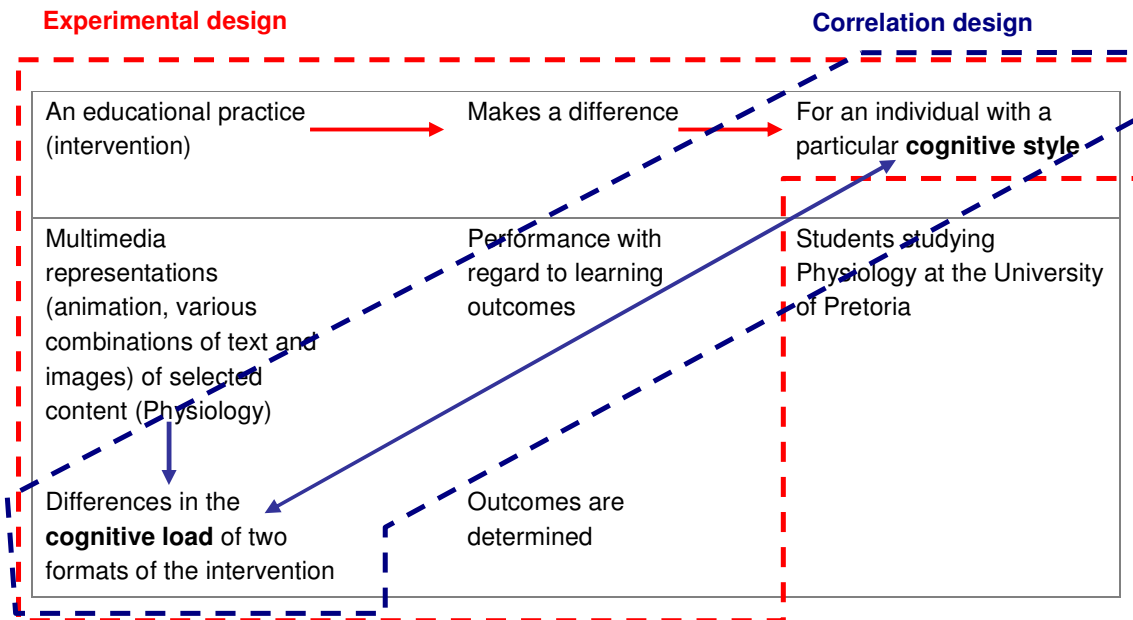


Figure 3.2: Graphic representation of the issues that define this research design

In the following section, I will examine the experimental design in more detail.

3.6.1 The experimental design

It is customary in experimental studies to randomise subjects into treatment and control groups. The purpose of this randomisation is to control variables that are not explicitly included in the study (Garson, 2006). I therefore randomly assigned the participants in this study to different groups.

Classic experimental designs are further divided into the following three groups (Creswell, 2005):

- Between-subjects design
- Within-subjects (repeated measures) design
- Matched pairs design

One design issue that required careful deliberation was the choice of whether to use a between-subjects or within-subjects design. When I looked at what needed to be taken into account for the measurement of cognitive load, I saw that at least one factor indicated that a between-subjects design would be the more appropriate experimental design.

Cognitive load is influenced by a subject's prior knowledge (Clarke et al, 2005; Kalyuga et al., 2003; Schnotz & Rasch, 2005). From the point of view of instructional design, the development of the best format (i.e. one in which the cognitive load is as low as possible) for novice learners with no or little

prior knowledge of a subject domain, requires that the cognitive load of the format be measured without interference from a prior knowledge variable.

This excluded a within-subjects (repeated measures) design because such a design is characterised by the use of the same subjects for each level of an independent variable. In a study such as this, that would have meant that the participants would have had to work their way through both versions of the multimedia. The obvious danger of such an approach was that the validity of the study might have been compromised by possible carry-over effects (such as those that occur when subjects have been exposed to earlier (or other) versions of the treatment).

Any repeated-measures design would also have influenced the assessment of the learning gain (as measured by the difference in performance between the pre- and posttest). This served to exclude the within-subjects (repeated-measures) design. It also excluded the matched-pairs design as a possible first choice for the design (the matched-pairs design is a variant form of the repeated-measures design). In this design, subjects who have similar key attributes are paired before they are assigned to a specific treatment or intervention. This kind of design avoids some of the types of invalidity that are produced by within-subjects designs such as the threat of subject fatigue across repeated tests. But one of the disadvantages of this design is that the design controls only for the matched attributes whereas same-subject within-subjects designs control for both explicit and unmeasured subject variables (Garson, 2006).

Experimental research in education does not always look at only one activity or intervention that is likely to influence the outcome. Experimental research designs often consider multiple variables simultaneously. They do this in an effort to replicate the field of practice where it is never only one factor that influences both practice and outcome. The control for unmeasured variables in a between-subjects design would allow the study to 'ignore' some of the variables that might influence the results – variables that cannot be controlled by the design. This study took place in an authentic learning environment, and there are many variables in an authentic learning environment that cannot easily be controlled or measured. Because a between-subjects design is the most powerful design for controlling unmeasured variables, it was the design that I eventually chose for this study.

My final decision was therefore to use a between-subjects factorial design. In terms of such a design, subjects who have different cognitive styles are exposed to a different version of the independent variable (in this study, one of the two multimedia formats). Apart from the different presentation formats of the research intervention, the formats were thought to have different cognitive loads. Because I studied three categorical, independent variables and their effect on the dependent variable the design was factorial. I also make comparisons between the subjects' reactions or effects (in this study, the achievement of learning outcomes).

But because the participants completed a pre- and posttest after they had been assigned to the different treatment groups, the design also possessed the characteristics of a within-subjects

(repeated measures) design of the kind in which the repeated measure is performance in a knowledge and comprehension test.

In summary, I randomly assigned the participants to one of the two presentation formats of the multimedia program. These formats were the narrated animation version (Version 1) or the static images & text version (Version 2). Both groups took a cognitive styles analysis test. Both groups also completed a pretest before studying the content so that I would be able to determine the extent of their prior knowledge. The pretest tested for recall and comprehension of knowledge. On completion of the instruction, I administered the same posttest to both groups. While the posttest allowed me to assess recall and comprehension of knowledge, it also included an additional section that tested for application of knowledge.

3.6.2 Managing threats to validity

Because there are always threats to the validity of an experimental design, I use this section to discuss the threats to this particular study. While Garson (2006) identifies six broad categories of validity, he cautions researchers to be less concerned with defining the types of validity and more concerned about questions that should be asked to test validity. Cohen, Manion and Morrison (2000) identify 18 different types of validity. The two greatest threats to experimental designs are internal and external validity.

Table 3.2 summarises the threats to validity that are relevant to this study, together with the measures that I used to control and/or eliminate these threats.

Threats to validity relevant to study	Measures to control these threats in this study
Internal Validity – relates to design and procedures used in an experiment	
Attrition from the sample	I used as large a group as possible (more than 130 participants). I scheduled sessions in a time period that was acceptable to participants (i.e. not a day before a test or after hours).
Attrition during the experiment. (Because of technological problems and errors in capturing the data logs, participants 'forget' to complete certain portions of the electronic questionnaire.)	The following measures were put in place to manage this threat: <ul style="list-style-type: none"> • I selected a large sample. • I tested the technology and electronic data collection process by means of a pilot test. • I gave the participants clear instructions about how to participate in the research component of the study. They were therefore thoroughly briefed before the session commenced. They also received notes that explained to them how they should log on, and research assistants were available to assist them

Threats to validity relevant to study	Measures to control these threats in this study
Internal Validity – relates to design and procedures used in an experiment	
	with their queries throughout the session.
The 'people factors' were implicated in the selection of participants for the different treatments.	I randomly assigned the participants to the groups, and randomly assigned the groups to the two versions of the research intervention.
Diffusion of treatments if the groups can communicate with each other	I appealed to the participants to work as individuals and not to communicate with one another during the session. The research assistants were briefed to be on the lookout for contraventions of this request.
Compensatory equalisation	The design of the multimedia ensured that extraneous load known to be detrimental to learning (such as the split attention effect) would be kept to an absolute minimum. I adhered to the principles of good design throughout the development of the intervention.
Carry-over effect from pretest to posttest: the possibility that participants might anticipate the questions in the posttest.	I changed the order of the questions in the posttest . Two questions were similar (but not exactly the same) in the posttest.
External validity - threatens ability to draw correct inferences from the sample data	
The independent variables were not explicitly described.	I described all the independent variables together with the method I used to measure them.
Invalid or unreliable instruments	I tested the reliability of the instrument. I used (wherever possible) instruments that had been developed and tested by other researchers.
The dependent variable was not adequately operationalised.	Since the test was checked by the subject matter expert, it could be used with confidence in any assessment setting. The test was relevant because it tested the knowledge that was covered in the multimedia program.
Content or face validity – instruments are fair and comprehensive and cover domain	
Pre- and posttest failed to cover the domain of learning adequately.	All questions were reviewed and approved by the subject matter experts before the experiment commenced. Questions addressed the learning outcomes for the program.
Statistical validity – conclusions can be accepted as the correct statistical procedures were used.	
Significance levels were not appropriate.	I established significance at $p < 0.05$, the level that is most commonly used in social science research.
Introduction of Type 1 errors (namely, thinking that a relationship exists where one does not exist at all)	I established significance at $p < 0.05$. I also developed the hypotheses <i>a priori</i> – i.e. before I undertook the analysis.

Threats to validity relevant to study	Measures to control these threats in this study
Internal Validity – relates to design and procedures used in an experiment	
Introduction of Type 11 errors (namely, thinking that no relationship exists where one does in fact exist)	I used powerful statistical procedures such as multiple regression analyses and analyses of variance. I assembled a large sample.

Table 3.2: Managing threats to validity

3.7 The research sample

The multimedia program designed to teach this topic **could** well be used by all health science students who study the topic of the Autonomic Nervous System (the population) for the first time (novice learners). Initially it will only be used by the students at the University of Pretoria (the target population). The second-year medical, dentistry and physiotherapy students (the sample) were all engaged in studying the topic of the multimedia – the Autonomic Nervous System (ANS) – at the very time during which I was undertaking this research intervention.

Sample size is often a problem in experimental research because the final size is often dictated by very practical issues such as the number of participants who actually volunteer for the study and the number available to the researcher (Creswell, 2002). There are two methods for calculating the sample size: by using a sampling error formula or by using a power analysis formula (Creswell, 2002).

When one uses a power analysis formula, one calculates the sample size by considering the level of statistical significance, the amount of power desired in the study and the effect size. There are various tables that one can use to calculate the sample size from these factors. Creswell (2002), for example, used one of Lipsey's sample size tables. Lipsey's table specifies the approximate sample size of the experimental group that one would require to arrive at various criterion levels of power for a range of effect sizes at $\alpha = .05$. The power that one needs to reject a hypothesis when it is false is typically set at 0.80. The effect size (which is the expected difference in the means between the control and experimental groups) is expressed in standard deviation units and is typically set at 0.5 for education research. By using these parameters I found that I needed a total of 130 participants (65 in each group). My plan was therefore to make use of the second-year medical and dental students from the University of Pretoria where the class size in 2006 was 262.

Difficulties imposed by funding and time limitations made it difficult for me to use the more rigorous probability selection technique. Creswell (2005) notes that this is a common problem in educational research. I therefore used convenience selection, which is a non-probabilistic selection strategy, to create the sample for this study.

Convenience selection was thought to be appropriate for this study for the following reasons:

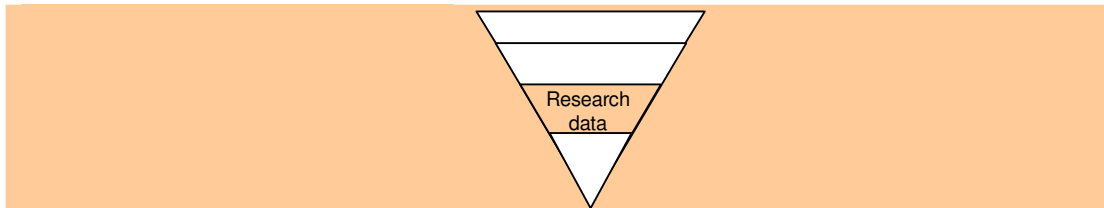
1. The potential number of participants was available.
2. The group identified for selection had characteristics common to both the wider research population and the target population.
3. The number of students in the group enabled me to assemble an adequate sample size.

Once I had identified the sample, I assigned the identified participants randomly to one or other of the treatment interventions. This procedure is in line with the rigorous approach required in experimental research and

...distinguishes a rigorous, 'true' experiment from an adequate, but less-than-rigorous, 'quasi-experiment'...

(Creswell, 2002, p. 284).

All the participants for the first pilot study and the main study were required to participate in the study as part of their scheduled class requirements. This is in contrast to the number of other studies in which students were required to learn material that was not directly relevant to their chosen field of study (Dwyer & Moore, 1991; Ford & Chen, 2001; Kiewra, & Mayer, 1997a; Quealy & Langan-Fox, 1998; Whelan, 2006).



3.8 The research data

This section describes the data that I collected for each of the three independent variables and the dependent variable, and the process that I used for calculating the final values for each variable.

The variables for this study are:

- Cognitive style (independent variable)
- Cognitive load (independent variable)
- Presentation format (independent variable)
- Achievement of the learning outcomes (dependent variable)

In Chapter 1 I presented basic definitions and explanations of the core concepts and terminology that are used in this study. The operational definition for each of the variables is presented here in Table 3.3.

Variable	Operational Definition
Cognitive style	Cognitive style has two dimensions. The first dimension (Wholistic-Analytic) refers to a person's preferred approach to organising information and it is measured on a bi-polar scale. The second dimension (Verbaliser-Imager) refers to the way in which a person represents information during thinking and is measured on a second independent bi-polar scale.
Cognitive load	Cognitive load refers to the burden imposed on the cognitive system when someone performs a particular cognitive task or activity. Two different techniques are used to measure cognitive load: (1) a subjective rating of mental effort, and (2) a direct method that makes use of an on-screen dual-task.
Presentation format	'Presentation format' refers to the use of specific multimedia strategies to deliver instruction. In this study the multimedia strategies are predominantly narrated animation (in Version 1), and mostly static images & text (in Version 2).
Achievement of learning outcomes	'Achievement of learning outcomes' refers to the measured difference in the scores of a pre- and posttest that test recall and comprehension of knowledge. It also refers to learner performance in a test that is designed to measure application of knowledge.

Table 3.3: Operational definitions of the variables

3.8.1 Cognitive style data

I collected two separate datasets for cognitive style. These two datasets contain ratios that indicate the position of each individual on each of the two dimensions of Riding's Cognitive Style Model, namely the Wholist-Analytic (WA) and the Verbaliser-Imager (VI) dimension.

3.8.2 Cognitive load data

I have already described the methods that are used to measure cognitive load in Chapter 2. In this study, the self-report rating (SR) method was used. Smith (2007) measured the cognitive load using the direct measure dual-task (DM) methodology. It is far easier to obtain measures of self-report of cognitive load than it is to use some of the techniques that measure cognitive load directly. Since there is no published research that examines the correlation between the measures obtained by these different techniques in the same study, and because Smith (2007), using the same sample and learning material that I used in my study, applied the direct method to measure cognitive load, I was

able to investigate this correlation. Other indicators of cognitive load include looking at performance outcomes and time (Chandler & Sweller, 1991).

3.8.2.1 Self-report rating of cognitive load

I measured mental effort, an indicator of cognitive load, by using a subjective rating scale. Each measurement provided a value between 1 and 9, where 1 indicated that *very, very little mental effort* had been required to understand the content and where 9 indicated that *very, very great mental effort* had been required. By using the mental effort scores from the screens (SR_n , where n is the number of the screen from which the subjective rating was obtained), I was able to calculate a total mental effort score for the entire program for each participant (SRCLV). The mean of the individual scores for each intervention provided a total score for the subjective rating of the cognitive load of the intervention (Version 1 = SR1 and Version 2 = SR2). These calculations provided a dataset of continuous variables.

Self-report of cognitive load for each participant:

$$SRCLV = \text{MEAN} (SR_1, SR_2 \dots SR_n)$$

Self-report of cognitive load for each version:

$$SR1 = \text{MEAN} (SRCLV_1, SRCLV_2, \dots) \text{ and } SR2 = \text{MEAN} (SRCLV_1, SRCLV_2, \dots)$$

3.8.2.2 Direct measurement of cognitive load

A value for cognitive load was obtained using the direct measurement method. This score was obtained in the study conducted by Smith (2007). Cognitive load was directly measured 10 times in Version 1 of the program and 13 times in Version 2. A total cognitive load score was calculated for each participant (MEANCL). These scores were then averaged to calculate a cognitive load score for each of the two interventions (Version 1 = DM1 and Version 2 = DM2). These calculations provided a dataset of continuous variables.

Direct measure of cognitive load for each participant:

$$\text{MEANCL} = \text{MEAN} (CL_1, CL_2 \dots CL_x)$$

where CL_1 is the cognitive load score for each screen where it was measured.

Direct measure of cognitive load for each version:

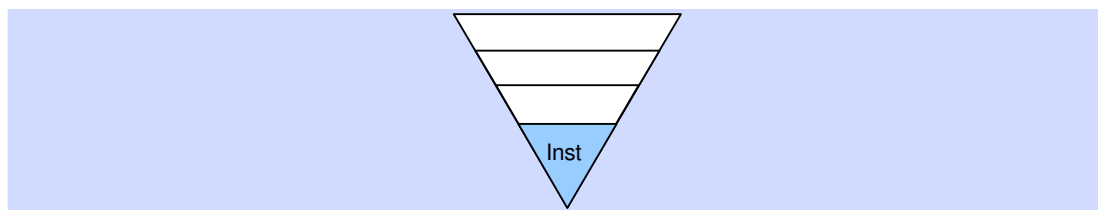
$$DM1 = \text{MEAN} (\text{MEANCL}_1, \text{MEANCL}_2, \dots) \text{ and } DM2 = \text{MEAN} (\text{MEANCL}_1, \text{MEANCL}_2, \dots)$$

3.8.3 Data for the presentation format

This was categorical data. I allocated a value of 1 for Version 1, the version with narrated animation, and a value of 2 for Version 2, the static images & text version.

3.8.4 Learning performance data

I measured learner performance by using a pre- and posttest design. The computer-based test items were the same for the pretest and posttest, except that the questions were presented in a different order in the two tests. I used a computer to calculate a score for each participant for both the pre- and posttest and then allocated a score of 1 for each correct answer. The tests were not marked negatively. The posttest also included a pencil and paper test, which tested application of knowledge. There were two questions in this section of the posttest. The maximum score for each question was 10.



3.9 The research instruments

This section deals with the development and use of the research instruments. In this study, I used the first two options listed below – options that are described in the research methodology literature for obtaining research instruments:

- A self-developed instrument
- An existing instrument that the researcher has not modified in any way
- A modified existing instrument (Creswell, 2005)

Table 3.4 summarises the instruments used in this study. A broader discussion of the development and use of each instrument follows the table.

(Appendix K provides more detailed information about how I integrated these instruments into the intervention and used them in the study.)

Instrument	Measures	Development of the instrument	Number of times administered	Source
Demographic questionnaire	Demographic profile of sample	Self-developed	Once	Not applicable
Cognitive Styles Analysis (CSA)	Cognitive style on both the WA and the VI dimension	An existing test	Once	I purchased a license to use CSA from Learning & Training Technology, United Kingdom. I used the South African version of the test (Riding, 2005a).
Pretest	Prior knowledge of learning outcomes	Self-developed	Once	Physiology textbooks (The test was validated by subject matter experts for content validity.)
Subjective rating of cognitive load	Self-rating of mental effort required to understand and learn the content	An existing instrument	Embedded in intervention at selected points in the program. Version 1 (Animation) Five (5) times Version 2 (Static text & images) Six (6) times	This method was developed by Paas et al. (1993). I obtained their permission via e-mail to use this instrument (F. Paas, personal communication, 31 July 2005).
Posttest	Achievement of learning outcomes – a performance measure	Self-developed	Once	Same test as pretest, plus two additional items.

Table 3.4: Summary of research instruments used in this study

As already indicated, Smith (2007) measured cognitive load by using the direct measurement method. While the protocol was modified for her study, the basic principles described by Brünken, Plass & Leutner (2003) did not change.

3.9.1 Measuring cognitive style

I measured cognitive style by using Riding's Cognitive Styles Analysis (CSA). This test and the literature about the CSA have already been described in Chapter 2.

3.9.2 Measuring cognitive load

3.9.2.1 Self-report ratings

The self-report measure that was used is a version that Paas adapted for his own use from the work of Borg, Bratfisch and Dornič (Paas et al., 1994b). This instrument measures perceived task difficulty on a 9-point scale ranging from *very, very low mental effort* (a score of 1) to *very, very great mental effort* (a score of 9). There are some researchers who preferred to use a 7-point scale rather than the 9-point scale (Marcus, Cooper & Sweller, 1996; Mayer & Chandler, 2001).

I decided to use this instrument rather than to develop my own because the literature describes it as a reliable and sensitive instrument that is capable of highlighting differences in the expenditure of processing capacity associated with different conditions (Paas et al., 1994b). It is also widely used in cognitive load research (Pollock, Chandler & Sweller, 2002; Kester, Kirschner, & Van Merriënboer, 2004; Tabbers et al., 2004; Clarke, Ayres & Sweller, 2005; Van Gog, Paas, & Van Merriënboer, 2006). I obtained permission from Paas to use this rating scale (F. Paas, personal communication, 31 July 2005).

Although the literature is critical of the value of self-report ratings, Paas et al. (1994b) assessed the usefulness of subjective ratings and cardiovascular measures of mental effort and found that the subjective rating scale was more useful than the cardiovascular measure. They subsequently recommended that the use of both these measurement techniques be further explored in instructional research.

Reliability was determined in the studies of Paas (1992) and Paas et al. (1994b) by using Cronbach's alpha coefficient of reliability. The coefficient for the studies was 0.90 and 0.82 respectively. I used the same test for reliability (see Chapter 4).

Figure 3.3 shows a screenshot of the instrument that the participants were asked to use to self-report a rating for mental effort in this study.

Section 3 | Questionnaire - Mental effort

A question about mental effort

Indicate how much mental effort you used to study the information on the two screens you have just reviewed. Select ONE option only.

- Very, very low mental effort
- Very low mental effort
- Low mental effort
- Rather low mental effort
- Neither high nor low mental effort
- Rather high mental effort
- High mental effort
- Very high mental effort
- Very, very high mental effort

Image cropped for illustrative purposes

Figure 3.3: Screen format of self-reporting a rating for mental effort

3.9.2.2 Direct method that uses a dual-task approach

Brünken et al. (2003) described and used a dual-task approach to measuring cognitive load. This approach requires the user to respond to a secondary task when using a multimedia program. The primary task is to learn, understand and master the content of the program. The secondary task requires some sort of response by the user to another stimulus that is superimposed on the learning content. Various strategies have been used to present a secondary task. These include *sound* (Brünken, Plass & Leutner, 2004, Chandler & Sweller, 1996), and *visual monitoring* (Marcus, Cooper & Sweller, 1996). The response time to this secondary task is then measured. The hypothesis is that the faster the response time, the lower the cognitive load. Their work with a within-subjects design led Brünken et al. (2003) to conclude that reaction time in a secondary monitoring task can be regarded as a valid measure of the cognitive load induced by multimedia instruction.

Smith (2007) investigated the use of the secondary task, which involved the use of the letter 'A', which changed colour on selected screens at pre-defined intervals. The first stimulus was presented exactly four seconds after the screen objects had all loaded, and at ten-second intervals thereafter. The participant was required to press the ENTER key on a standard QWERTY keyboard every time he/she noticed the letter changing colour. The time between the stimulus, which Smith called a trigger, and the response (record of the time at which the ENTER key was pressed by the

participant), if any, was recorded by the computer and written out to an external data file. A sample of such a data file is displayed Appendix N.

3.9.3 Measuring learner performance

Both the pre- and posttest were computer-based and were scored electronically. Each test contained 9 items. The maximum possible score was 22 (see Appendix D), and the total for each test was calculated during the analysis of the data.

The computer-based pre- and posttest measured the participant's knowledge and understanding of the content. The paper-and-pencil posttest tested the ability of the participant to apply and transfer this knowledge to an authentic clinical scenario. There were two questions (see Appendix E).

I used the content of the multimedia to draw up the questions for both tests, and obtained that content from three different Physiology textbooks (Meyer, Van Papendorp, Meij & Viljoen, 2002; Sherwood, 2004; Silverthorn, 2004). I then passed the questions to subject matter experts and asked them to review the questions for content and face validity.

In the first question –

You are on holiday at the sea. A swimmer narrowly escapes a shark attack. You go to see if you can help. Fortunately there are no injuries, but the person is very shocked.

Describe the clinical symptoms you would expect to see, and provide adequate information about what you see and why you see these symptoms. [10]

– knowledge of the physiology of the Autonomic Nervous System would enable the participant to recognise the symptoms of shock and thus to treat the shark attack victim appropriately. It was not the treatment of this shock that was assessed but merely the ability of the participant to apply knowledge to an authentic case.

In the second question –

You are assigned to the spinal unit during a clinical rotation. The rehabilitation of the paraplegic patients involves teaching them how to empty their bladder.

What neurological process makes this possible? [10]

– knowledge of the reflexes of the Autonomic Nervous System would enable the participant to explain to a paraplegic patient how they must empty their bladder.

The difference between the pretest score and posttest score for each participant was calculated during the analysis of the data.

3.9.4 Integrating the research instruments into the intervention

The CSA was administered as a separate test at the beginning of the experiment. In order to provide a more streamlined experience and to assure that the design promoted authenticity, I decided to integrate the data collection instruments and multimedia interventions into one multimedia program. I therefore divided the entire multimedia program into the six sections that are illustrated in Figure 3.4. The version illustrated here was used for the second pilot study. The main study did not have a menu. It had a title screen and the sections followed each other in the order shown in Figure 3.4 so that the presentation would be smoother. Appendix F shows the title and practice session screens. The version that was used in the main study is shown in Appendices G, H, I and J of this thesis. Appendix G presents the content common to both Version 1 and 2 of the program. Appendices H to I display the different presentations of content.

There were three differences in presentation format across the full program:

- Different strategies to display content (Appendix H)
- Animation versus static images (Appendix I)
- Whole view versus Parts view (Appendix J)

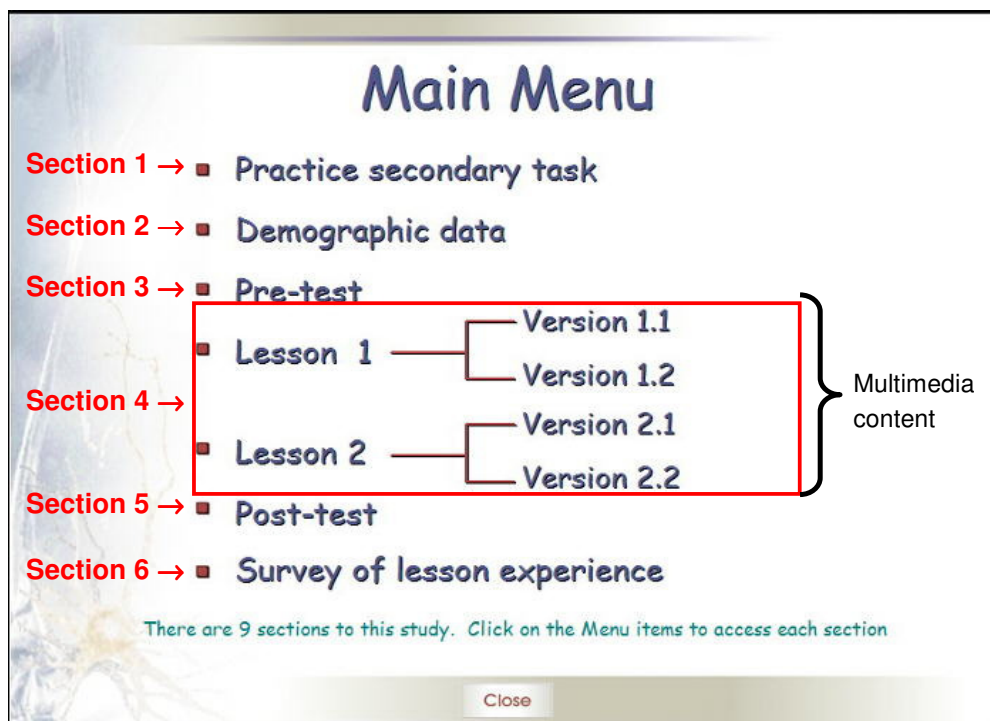


Figure 3.4: Menu for multimedia intervention used in the second pilot study

Figure 3.4 shows how the content was presented in Lesson 1 and Lesson 2. Participants were randomly assigned to one of four programs (Version 1.1, 1.2, 2.1 and 2.2). All the other sections had to be completed by each participant.

This study only investigated two formats of the content (these are illustrated in Figure 3.4 as Lesson 1 and Lesson 2). Figure 3.4, however, displays four lessons. This is because Smith (2007) used the same multimedia program for her study. Smith also investigated the influence of the **screen position** of the secondary task on cognitive load. Because of this, it was necessary for me to develop two programs for each version. The only difference between Program 1.1 and 1.2 and between 2.1 and 2.2 was the position of the secondary task on the screen – a difference not relevant to my study. The considerations that influenced the design of the two versions will be described in more detail in Part 2 of this chapter. A summary of the multimedia and the integration of the research instruments is provided in Appendix K.

Section 4 of the multimedia was the actual program with the content that had to be studied. The instrument to collect the data for the self-report of mental load and the direct measurement of cognitive load was embedded in this section at various points in the program, and is summarised in Table 3.5.

	Self-report rating of mental load	Direct measurement using dual task methodology
Version 1 (Animation)	5	10
Version 2 (Static images)	6	13

Table 3.5: Number of times cognitive load was measured in each version

3.10 Ethical considerations in this study

The following four ethical issues received attention in this study:

- Obtaining the consent to participate in the study
- Informing the participants about their cognitive style
- Ensuring the anonymity of the participants
- Making both formats of the intervention available to the participants

3.10.1 Obtaining consent to participate in the study

I first obtained consent to undertake this study from the Department of Physiology at the University of Pretoria. A copy of the letter requesting permission to conduct the study is contained in Appendix B, as is a copy of the letter of consent that was sent to me by the Head of Department of Physiotherapy (Physiotherapy students participated in the pilot study).

I met with the participants before the pilot and main study commenced, informed them about the study and explained to them what would be required of them. I then gave the participants an opportunity to ask questions about the study. After that I presented each participant with a letter that

invited them to participate in the study, and I asked them to use the letter to give their written consent to participation in the study. A copy of this letter of invitation is contained in Appendix O.

3.10.2 Informing participants about their cognitive style

The Cognitive Style Analysis (CSA) of Riding displays the scores for two dimensions of the analysis at the end of the computer-based test. While the participants were allowed to record these scores for their own interest and benefit, the test did not offer any information about how to interpret these scores.

Knowing the cognitive style of learners can facilitate decisions about effective instructional design strategies in cases where a particular strategy will empower learning because of the match with style (Douglas & Riding, 1993; Ford & Chen, 2001; Graff, 2005) and in cases where a strategy can help to compensate for the weaknesses of a particular style (Graff, 2003b; Riding & Sadler-Smith, 1992; Triantafillou, Pomportsis, Demetriadis & Georgiadou, 2004). By telling learners what their cognitive style is, we create a situation in which they are able to understand their own idiosyncratic techniques of learning. This will enable them not only to build on their own strengths, but also to adjust their learning strategies to compensate for the weaknesses that are inseparable from a particular style. While I have no intention of further discussing or debating the issue of knowing one's own cognitive style in this study, I suspected that participants would be curious about their own cognitive style. I was able to confirm this during both the pilot and the main study when several participants asked for more information about cognitive style and the meaning of their scores. Since I had already determined that none of the participants had ever taken the CSA or a similar test, I decided not to provide any information about cognitive style characteristics *before* the data collection commenced. This was one method that I used to control this variable. But since there could be no valid reason for withholding information about personal cognitive style from the participants in the long run, I did appraise each participant individually of the results of their test once the data had been analysed. I also supplied each participant with a handout that explained what cognitive style was and how cognitive style should be interpreted.

3.10.3 Anonymity

Because I had already informed the participants about their cognitive styles, I needed some way of identifying them individually. I decided that the best way of ensuring the anonymity of individual participants would be to identify them by means of their student numbers. I therefore decided to use another software application that was generally available in the university to randomly assign students in a class to experimental groups, and requested a faculty member who was not involved in this study to generate the groups for me. This application displayed the group members by name and student number. The group lists that were thus compiled were given to the participants after I had briefed them about the study. Each participant subsequently completed a consent form that displayed both the participant's name and student number. (I needed their names so that I would be able to

inform each participant about the results of the CSA once the study had been completed.) These forms were then collected and filed.

This use of student numbers effectively protected the anonymity of each participant. The participants were required to enter their student number when they logged in to the multimedia program. They were also asked to write their student number rather than their name on the paper-based section of the posttest and some of the other documentation they completed during the data collection. Since the electronic files for each participant used only the student number for the file name, it was not possible for me to identify any particular participant merely by looking at the student number.

Before I told the participants their cognitive styles, I enlisted the aid of an independent third party to help me to match student numbers to names. Each student was given an individual letter sealed in an envelope, and that letter contained the results of their CSA. The letters were given to the class representative and he then distributed these letters to his peers.

3.10.4 Making all the formats available to the participants

A decision to withhold an intervention from selected participants raises ethical questions, especially if the experimental intervention concerned turns out to be the better intervention. In order to ensure that no participant would be disadvantaged in any way by the study, I made arrangements for both formats of the program to be loaded onto the network in the computer laboratories. This gave participants the opportunity of revisiting the multimedia programs at any time.

3.11 Summary for Part 1

Part 1 has described the research methodology and design in detail by using the following framework:

- Research approach
- Research design
- Research data
- Research instruments

The chapter described a quantitative study that used an experimental between-subjects design to determine the relationship between cognitive load and cognitive style in multimedia learning. I discussed the variables under consideration (cognitive load, cognitive style and learning performance), and carefully described both the data and the instruments that I had used to collect this data. The data was collected electronically by means of the following instruments:

- Cognitive Styles Analysis
- Demographic questionnaire

- Pretest to assess prior knowledge
- Subjective Rating of Mental Effort
- Posttest to assess knowledge at level of recall, comprehension and application

Part 1 concluded with a discussion of the ethical issues relevant to the study.

Part 2: The Design of the Intervention

3.12 Introduction

This section of Chapter 3 briefly describes the intervention that was used in this study. The empirical findings from several studies about instructional design that controls cognitive load were used in the design of the multimedia used in this study. Many current multimedia design guidelines are grounded in substantial research (this research was reviewed and discussed in Chapter 2). By taking account of the findings of these studies, my intention was to create a situation in which new research would build on the solid foundations provided by previous research.

3.13 Source of the content

The topic of the finished multimedia program was 'An Introduction to the Autonomic Nervous System' – a compulsory topic of study in the Homeostasis module of Physiology. Physiology is a compulsory course for medical and dental students in their second year of study at the University of Pretoria (UP). For the students of all the other disciplines (Nursing Science, Physiotherapy, Occupational Therapy, Radiography) represented in the Health Sciences Faculty at UP, Physiology is also a compulsory course, and the Autonomic Nervous System (ANS) is a part of that curriculum. The learning outcomes for the Homeostasis module and the various units of this module have already been compiled for the curriculum of the MBChB programme. My analysis of these learning outcomes guided the development of my own multimedia program, and I devised more specific and detailed learning outcomes for the multimedia program in consultation with the subject matter experts.

I was able to use selected content from a multimedia program that had been developed for the Department of Physiology at UP. The target audience for this original program – Psychoneuroimmunology (PNI) – consisted of postgraduate Physiology and Psychology students. I obtained permission to re-use selected content from that program from the lecturer who was the subject matter expert for the development of the PNI multimedia. This particular subject matter expert had been teaching the ANS and has been undertaking research in this field for more than 25 years. I also worked closely with the lecturer who presents this topic to second-year MBChB students. She had been teaching the ANS to various student groups for close on 15 years. Both these subject matter experts provided input for the content of the program.

I developed two versions of the multimedia: Version 1, which used narrated animation, and Version 2, which used static images & text as an alternative to narrated animation. Two of the three animations used in this study were re-used in their original format. The third animation was developed specifically for this program. Some of the text-based content had to be adapted so that it would be appropriate for the target group. It was also necessary to develop some new content. The static images and the accompanying text used for the alternative presentation of content in Version 2 were entirely new.

3.14 Design and development of the multimedia

I designed this program, storyboarded it on paper, and developed it during the months between October 2005 and January 2006 inclusive. Both the subject matter experts reviewed and approved the storyboards before development commenced. The program was tested extensively by myself and another instructional designer, after which it was reviewed by both the subject matter experts. Their subsequent input and critique led me to make small changes to the content and the design. There were no programming errors in the program. My experience during the early testing of the program led me to schedule two hours for the study.

3.14.1 Design of the program

The learning outcomes for the program were couched in the following language:

- Describe the structure of the Autonomic Nervous System by using basic illustrations.
- Explain the control of the Autonomic Nervous System.
- Compare the structure and function of the Sympathetic and Parasympathetic divisions.
- Describe the function of the Sympathetic Nervous System.
- Describe how the function of the Autonomic reflexes applies to patient care management for selected problems.

Figure 3. 5 (next page) illustrates the structure of the program visually.

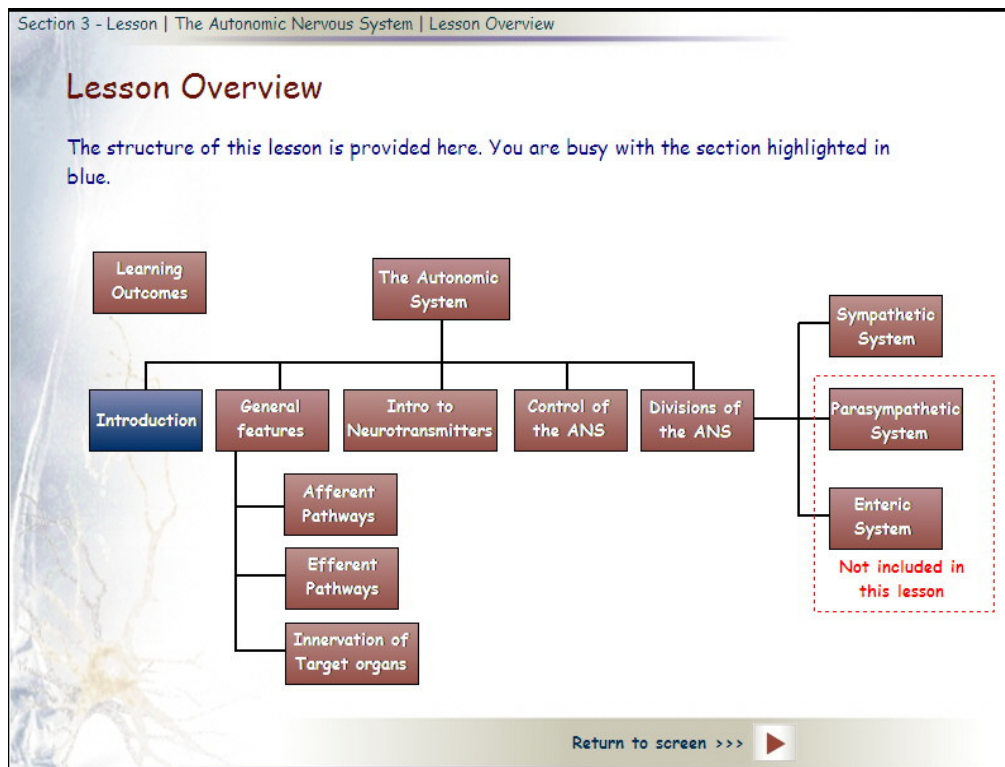


Figure 3.5: A visual representation of the structure of the multimedia program

Figure 3.5 illustrates the first screen of the program after the title screen. It provides an overview of the program. First-time users were required to study the content in a linear fashion. They were able to return to this program overview at any time by clicking on a Site Map button that was displayed at the bottom of the screen (it is not visible in Figure 3.5). The learners could use this overview screen as a menu once they had completed the program; it gave them access to any of the sections that they might need. The user’s navigation through the program was tracked and recorded in an external log file.

Table 3.6 provides a summary of the structure of the two formats of the program.

	Version 1 - Animation		Version 2 - Static images	
	Program 1	Program 2	Program 3	Program 4
Formats used to present same content	Narrated animation on one screen		Static images & text using five screens	
Number of screens in program	19		23	
Total number of animations in program	2		2	

Table 3.6: A summary of the major design similarities and differences between the programs

3.15 Instructional strategies and media used

The design of the multimedia program was informed by previous cognitive load and multimedia learning research (Chan & Black, 2005; Chandler & Sweller, 1996; Dempsey & van Eck, 2003; Gerjets & Scheiter, 2003; Ginns et al., 2003; Kalyuga et al., 1999; Leahy et al., 2003; Mayer, 2003; Mayer, 1997; Mayer et al., 2004; Mayer et al., 2003; Mayer et al., 2002; Mayer & Chandler, 2001; Mayer et al., 1999; Mayer & Moreno, 1998; Mayer, Bove, Bryman, Mars & Tapangco; 1996; Moreno, 2004; Moreno & Mayer, 2002; Moreno & Mayer, 2000a, 2000b; Moreno & Valdez, 2005; Moreno et al., 2001; Paas et al., 2005; Sweller & Chandler, 1994; Sweller et al., 1998; Tabbers et al., 2004; van Gog et al., 2004).

Table 3.7 (on the following page) summarises:

- The scope of the content
- The number of screens that were used to present the content
- The teaching strategy that was used

The actual screens are illustrated in Appendices G–K. Although there was only one screen that was incomplete for the main study, I used a different strategy to present exactly the same content that was contained in the incomplete screen on another screen. Cognitive load using the dual-task approach was not measured on this screen and the participants were not asked to estimate the mental effort that they had exerted to understand this content.

Section	Scope of the content	No of screens		Instructional strategy and media used	
		Vers 1	Vers 2	Version 1	Version 2
Introduction	Title screen, program overview and learning outcomes, functions of the Autonomic Nervous System (ANS), the physiological activities controlled by the ANS, an overview of the three parts of the ANS and the general structure of the system.	6	6	The same strategy and media were used for both versions. Instruction was tutorial in style, and pop-ups provided additional information. Learners were given control over whether to access or not access the additional information. Static images were used to instruct learners about the differences in the structure of the three parts of the ANS.	
General features – 3 sections	Afferent pathways: Describes the structure and function of pathways carrying stimuli from the environment to the spinal cord and cortex.	3	3	Afferent pathways: The instruction was tutorial in style, and pop-ups provided additional information. One of the screens included a drag-and-drop activity to stimulate interaction. A static image was used to explain the content.	
	Efferent pathways: Describes the structure of the pathway and its components.	1	1	Efferent pathways : The instruction was tutorial in style. It made use of text and a static image.	
	Innervation of the target organs: Describes how this works	3	3	Innervation of the target organs: The instruction was tutorial in style. The first screen used text. The second screen was an interactive screen that permitted users to explore the effects of the SNS and PNS on different organs. This effect was illustrated visually. Learners could toggle between the two systems and observe the differences. They could also reset the organ to a neutral state. Version 1: The third screen summarised the information in one large table. This is illustrated in Appendix J, page 389 Version 2: The third screen summarised the information in a series of smaller tables that were presented by means of pop-ups. This is illustrated in Appendix J, page 389 - 390	
Introduction to Neuro-transmitters	Basic information about the functions of neurotransmitters. Also presented in the face-to-face presentation of the course.	1	1	A narrated animation was used to explain this content in both versions.	
Control of the ANS	Basic control mechanisms are discussed, with particular emphasis on the Sympathetic Nervous System (SNS).	1	1	The same strategy and media were used for both versions. The instruction was tutorial in style and pop-ups provided additional information. Learners were able to control whether or not to access the additional information. Static images were used on both the main screen and in each of the four pop-ups.	

Section	Scope of the content	No of screens		Instructional strategy and media used	
		Vers 1	Vers 2	Version 1.	Version 2
Divisions of the SNS	The detail of the structure and function of the SNS is covered in this section. While there is a multimedia program that covers the other two sub-systems, the scope of the content covered for this study had to be limited to what learners could do in an average lecture session of between 45 and 50 minutes.	4	8	The first screen used text and a static image to introduce the topic.	
				The second screen made use of a narrated animation (which was 1 minute and 46 seconds in duration) to describe the transmission of stimuli along the neural pathway.	The second screen introduced the learner to the smaller sections and served as a sub-menu screen.
				The third screen used text and static images to explain the difference between convergence and divergence. Each concept is explained and illustrated in a separate pop-up which almost fills the entire screen when it is activated.	The 3 rd , 4 th 5 th and 6 th screens replaced the animation of version 1 (screen 2). The third screen used text and a static image, as well as a pop-up that to display a second static image for purposes of instruction. The image in this pop-up was created from the material used for the animation.
				The fourth screen used text and a static image for the instruction.	The fourth screen used text and a static image as well as text-based pop-ups.
					The fifth screen used images and animation (not narrated) with text labels.
					The sixth screen used text and a static image for purposes of instruction.
While the seventh screen used text and static images to explain the difference between convergence and divergence, its method of displaying the content differed from version 1.					
The eighth screen was the same as the fourth screen in version 1.					
TOTAL NUMBER OF SCREENS		19	23		

Table 3.7: Scope of the content of the multimedia program

Part 3: The Hypotheses

3.16 The hypotheses and expected findings

This section deals with the hypotheses and expected findings on the basis of existing knowledge about:

- how learners with different cognitive styles process and represent information, and
- the influence of the cognitive load on learning.

3.16.1 Distribution of the styles across the sample

There is not enough evidence in the literature about the cognitive style profiles of health science students to make a prediction about the findings of this study on the basis of existing literature alone. Detailed familiarity with the curriculum and nature of the field itself suggest some directions that a discussion about the expected findings might take. My own qualifications in this regard are my extensive experience as a professional nurse for close on 20 years, the fact that I have taught in this field for the past five years and my work as a multimedia instructional designer in the Faculty of Health Sciences for five years.

On the one hand, this discipline is characterised by a massive knowledge base of facts that need to be learned and integrated. Clinical skills in the Health Sciences presuppose that a practitioner has mastered both the knowledge base and a large number of processes and procedures. One may assume that a practitioner who is Analytic in style will find it easier to master the necessary knowledge and skills than a Wholistic learner, since they favour a step-by-step detailed content. Curriculum reform in Health Science education (Rendas, Pinto & Gamboa, 1998; Treadwell et al., 2002) has moved away from studying subjects in contextual isolation and has embraced a situation in which many of these subjects, skills and processes are currently integrated into problem-based teaching and learning strategies. Since there are simply so many of these facts that need to be integrated, one finds that the learning environments are characterised by an intrinsically high cognitive load. On the other hand, Health Science students need to be able to see and understand their patients holistically if they hope to acquire clinical skills, make correct diagnoses and prescribe appropriate treatments. Problem-based curricula place a strong emphasis on acquiring these skills (Schmidt, Vermeulen & Van der Molen, 2006), and students who display a more Wholistic style are more likely to benefit from a problem-based approach to learning and skill acquisition. Students who are natural communicators and who thrive on social interaction will in all likelihood fit into such an environment more easily.

In such circumstances, one may consider the following hypotheses:

Null Hypothesis 1a

There will be no difference in the percentage of the sample having either an Analytic/Intermediate or Wholistic style on the WA style dimension.

Alternate Hypothesis 1a

More than 50% of the sample will have an Analytic or Intermediate style on the WA style dimension.

Null Hypothesis 1b

There will be no difference in the percentage of the sample having either a Verbaliser/Bimodal or Imager style on the VI style dimension.

Alternate Hypothesis 1b

More than 50% of the sample will have a Verbaliser or Bimodal style on the VI style dimension.

The one variable that could interact with style and load is that of *time*. It is Riding's belief that an Analytic learner processes information more elaborately and spends more time looking at detail (Riding et al., 2003). Elaborate processing, which is typical of the Analytic learner, also implies that the learner needs more time to process the information. Not only does the Analytic learner need more time; he or she will *take* more time to learn content. I did not control for time in this study, but gave all the participants as much time to complete the program as they needed because of the authentic nature of the study. I did, however, keep track of the total amount of time that the participants spent on the program by tracking individual screens.

For the WA style dimension, the following predictions were made:

Null Hypothesis 1c

There will be no difference between the Analytic and Wholistic learner with regard to the amount of time spent studying the content of the program.

Alternate Hypothesis 1c

The Analytic learner will spend more time studying the content of the program than the Wholistic learner.

The program was also visually rich. There were very few screens that did not have an image to illustrate its content. One might expect that this kind of format and would suit learners with an Imager style, even though the literature provides evidence that learners with a Verbaliser style also enjoy and benefit from visually rich content. Two hypotheses were formulated.

After thinking about the program in general, I made the following predictions:

Null Hypothesis 1d

There will be no difference between the Verbaliser and Imager learner with regard to the time spent studying the content of the program.

Alternate Hypothesis 1d

The Verbaliser will spend less time studying the content of the program than the Imager.

There were 29 images across 19 screens in the animation version and 39 images across 23 screens in the static images & text version in the whole program. There were therefore proportionally more images in the animation version. Verbalisers are more likely to spend more time on the processing of text than images when they study content.

I also made the following predictions when I compared the time the Verbalisers spent on each version with the time Imagers spent on each version.

Null Hypothesis 1e

There will be no difference proportionally in the amount of time that the Verbaliser learner will spend in studying the two versions of the content.

Alternate Hypothesis 1e

The Verbaliser will spend proportionally more time in studying the content of the static images & text version than in studying the animation version.

3.16.2 Cognitive load of the presentation formats

I deliberately designed the program in accordance with the guidelines that I found in both the multimedia and cognitive load research streams. I paid particular attention to avoiding extraneous cognitive load attributable to poor design techniques such as split-attention and redundancy. Since I anticipated that the intrinsic load of the material would be high, I deliberately adopted a design that incorporated the advantages of the modality effect (i.e. the use of text and audio rather than just text alone). It was for this reason that the animations were narrated.

But since the learning would take place in an authentic learning environment, I developed and presented a fully integrated program that was relevant to the sample. There were specific screens within the program that were directly aligned with the purpose of the study. The cognitive load was measured at these points in the program, as well as at least one other point. This additional measurement served as a control because the content and presentation format were exactly the same for both versions. While this design enabled me to determine the cognitive load for the program as a whole (i.e. just as it would be used in an authentic learning environment), it also allowed me to examine the cognitive load at very specific points.

Null Hypothesis 2a

There will be no difference between the animation and static images & text versions in the cognitive load of the program as a whole.

Alternate Hypothesis 2a

In considering the program as a whole the animation version would have a higher cognitive load than the static images & text version

An additional explanation for this hypothesis is that the chunking of the animation into a series of screens using text and static images would reduce the intrinsic load of the material, thus resulting in a lower cognitive load.

Null Hypothesis 2b

There will be no difference in the cognitive load of the screen using animation and the alternative version using static images & text.

Alternate Hypothesis 2b

The screen using animation will have a higher cognitive load than the alternative version using static images & text.

Null Hypothesis 2c

At screen level there will be no difference in the cognitive load across the versions where the content and presentation format are the same.

Alternate Hypothesis 2c

At screen level, the cognitive load will be the same in each version where the content and presentation format are the same.

This study also investigates the correlation between the cognitive load scores for the same program by using different methods of measuring cognitive load.

Null Hypothesis 2d

There will be no difference in the cognitive load for each version when two different methods are used to measure cognitive load.

Alternate Hypothesis 2d

The two methods used to measure cognitive load will return results that are the same for each version.

Null Hypothesis 2e

There will be no correlation between the self-report method and direct measurement method for determining cognitive load.

Alternate Hypothesis 2e

There will be a positive correlation between the self-report method and direct measurement method for determining cognitive load.

3.16.3 Rating of cognitive load according to style

While Wholistic learners prefer to learn by using big picture views, Analytic learners prefer to learn by means of processing step-by-step detail. Since the animation presents the information in one session, it is expected that this will suit the learner who has a Wholistic style. And because the animations are also strongly visual, this should suit the learner who has an Imager style.

Null Hypothesis 3a

There will be no difference between the Wholistic and Analytic learner for the cognitive load of the animation version.

Alternate Hypothesis 3a

The cognitive load of the animation version will be lower for the Wholistic learner than for the Analytic learner.

Null Hypothesis 3b

There will be no difference between the Wholistic and Analytic learner for the cognitive load of the static images & text version when it is used as an alternative for the animation version.

Alternate Hypothesis 3b

The cognitive load of the static images & text version, when used as an alternative for the animation version, will be lower for the Analytic learner than for the Wholistic learner.

Null Hypothesis 3c

There will be no difference between the Verbaliser and Imager learner for the cognitive load of the animation version.

Alternate Hypothesis 3c

The cognitive load of the animation version will be lower for the learner with an Imager style than for the learner with a Verbaliser style.

Riding proposed that an Analytic learner processes information more elaborately and spends more time looking at details (Riding et al., 2003). Elaborate processing, which is typical of the Analytic learner, also implies that the learner needs time to process the information. Not only does the Analytic learner need more time, but he/she uses more time to learn content where such time is available. I did not, as I reported earlier, limit the time available to the participants in this study because of the authentic nature of the study. I did, however, track the total amount of time that the students spent on the program and the amount of time that each students spent on individual screens in each program. Doing this enabled me to confirm whether learners with Analytic style do use more time than Wholistic learners, and whether time did in fact moderate the relationship between load, style and learning performance.

Those analytic learners who spent inadequate time on the program (for reasons that I did not examine) might have not been able to process as elaborately as they would have liked, and they would therefore rate the cognitive load as being lower.

Null Hypothesis 3d

The amount of time spent on the program will make no difference to the rating of cognitive load by the Analytic learner.

Alternate Hypothesis 3d

Analytic learners who spent inadequate time on the program will rate the cognitive load lower than other Analytic learners.

How will this affect Verbaliser and Imager learners? The content in this study was presented by means of text and animated narration or text and images. The narrated animation was also visually rich. These combinations should suit both the Verbaliser and Imager. VI style should therefore not influence learning performance unless the cognitive load was high and inadequate time was spent on the program.

Null Hypothesis 3e

The amount of time spent on the program will make no difference to the rating of cognitive load by the Verbaliser and Imager learner.

Alternate Hypothesis 3e

Verbalisers and Imagers who spent inadequate time on the program will rate the cognitive load more highly.

3.16.4 The relationship between cognitive style, cognitive load and learning

Riding et al. (2003) investigated cognitive style, working memory, learning behaviour and attainment within a secondary school context and their subsequent discussion refers to the Analytic style within the context of their findings. Analytics typically have a more elaborated approach to processing. They examine all aspects of the subject and consider all the options. While such an approach is likely to result in good performance, it requires an adequate amount of processing capacity. Wherever the processing capacity of the Analytic learner is low, they perform poorly. But where it is high, they perform well. In contrast to this, Wholists, who aim for the bigger picture, generally require less processing capacity. Working memory capacity will therefore not exercise such a strongly determining influence on their performance. When one considers the Verbaliser–Imager style dimension, one sees that Verbalisers are typically more like the Analytics in their processing needs (Riding et al., 2003). Because of this, they will also perform well if they have enough processing capacity.

Is processing capacity dependent upon intelligence or load in the working memory? Since Riding et al. (2003) do not clarify this distinction adequately in their publication, I intend to interpret it from a cognitive load perspective. It is important to remember that the design of the multimedia intervention in this study aimed to keep extraneous load as low as possible. When one keeps the cognitive load as low as possible, one increases both the available working memory and the possibility to perform better. No learner with sufficient working memory capacity to process in a manner that suits his or her style should report a high mental effort. One should therefore not use cognitive style to predict

performance where the cognitive load is in the low or medium range (< 7 on the subjective rating scale).

If cognitive load is high, then cognitive style might influence performance, although this could be moderated by the amount of time spent on the program. This same group of Analytic learners (who spent inadequate time on the program and rated the cognitive load as low) would also not perform as well as the other groups of Analytic learners.

Null Hypothesis 4a

There will be no difference in the posttest results of the Analytic learner who spent inadequate time on the program and who rated the cognitive load as low and the other Analytic learners.

Alternate Hypothesis 4a

The Analytic learner who spends inadequate time on the program, and who rated the cognitive load as low, will perform more poorly on the posttest.

Once again I predicted that the interaction of time and cognitive load would influence the performance of the learners with VI style.

Null Hypothesis 4b

There will be no difference in the posttest results of the Verbaliser and Imager learner who spent inadequate time on the program and who rated the cognitive load as high and the other Verbaliser and Imager learners.

Alternate Hypothesis 4b

The Verbaliser and Imager learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest.

This concludes Part 3 of Chapter 2. The next part deals with the conduct of the studies.

Part 4: Conducting the studies

3.17 Introduction

Part 4 describes the implementation of the research design and methodology. While this study included two pilot studies and a main study, the window of opportunity for conducting both the pilot and main study was very small. This was caused by the fact that the intervention was part of their approved study programme for the participants who took part in the first pilot study and the main study.

Although the participants had been informed by their lecturer that the multimedia session was compulsory, I did not enforce participation in the research. In order to adhere to the ethical guidelines and protocol of the university, I agreed that participation in the research component of the class would

be voluntary. This explains the difference between the size of the class and the sample. Problems with the technology caused a small number of participants either not to complete the programs or to be dropped from the sample because of large amounts of missing data. Table 3.8 reflects the number of participants in the class and the final sample for both the pilot and main study.

Study	Class / Potential population	Class	Sample	Completed
Pilot 1	Physiotherapists	38	38	35
Main study	2 nd year medical and dental students	262	245	238

Table 3.8: Profile of the research sample

3.18 Conducting the pilot study

The main aims of the pilot study were to:

- conduct a final evaluation of the design of the multimedia intervention
- test the research instruments which were embedded in the multimedia program, and to refine them if necessary
- test the reliability of the electronic data collection process
- determine the average time spent on using the program so that scheduling for the main study could be finalised
- identify potential problems that might arise during the main study – problems that could contaminate data or introduce bias in any kind of way
- view and analyse some preliminary data
- finalise the research questions

Smith (2007) used the same pilot study for her study. The purpose of her study was to measure the cognitive load of the two multimedia programs by using the direct-measurement method. The intention was to have this information available before the main study commenced. In the event that there was no significant difference in the cognitive load of the different formats, the design would have been manipulated until a format with high and low cognitive load could be identified. Due to a series of technological problems during the first pilot study, it was not possible to measure the cognitive load of the two interventions.

The outcome of the first pilot study demonstrated that:

- the instructional design of the multimedia intervention was sound
- the participants understood how to use the research instruments and could enter the required data with minimal additional instruction

- the reliability of the electronic data collection process was NOT stable or correct
- it took between 60 and 90 minutes to complete the multimedia and answer all the questions

Apart from the electronic data collection process, no obvious problems appeared to arise during the main study that could contaminate data or introduce any kind of bias. Since the electronic data collection process was not reliable, a second pilot study had to be scheduled. The main aim of the second pilot study was to ensure that the problems connected with collecting the data electronically were resolved. Some additional programming was needed to improve the electronic data collection process. And because there was no longer enough time between Pilot Study 2 and the main study to manipulate the cognitive load of either version, we preserved the original design of the programs.

3.18.1 Pilot study 1

3.18.1.1 Participants

Thirty eight second-year Physiotherapy students who had registered for the 2006 academic year at the University of Pretoria participated in the study. Although this was a convenience sample, these participants covered the same learning material in their curriculum as the participants in the main study.

3.18.1.2 Materials

The learning materials that were used were described in Part 2 of this chapter.

3.18.1.3 Procedure

The pilot study was conducted on the Health Science campus of the University of Pretoria during January 2006. The laboratory was equipped with seventy computers and all the participants took part simultaneously in one session.

The participants were randomly assigned to two different groups on arrival at the computer laboratory. Each group was randomly assigned to one of the two programs and the participants were then briefed as a group about the study. The focus of the briefing was on the

- purpose of the study
- different sections in the multimedia program
- a secondary on-screen task that we used to measure cognitive load

This briefing took fifteen minutes. Participants were asked to indicate if they did not understand any of the questions in the various instruments. The instructions that governed the sections of the study that had to be completed were also projected in the laboratory for the duration of the session. Participants were helped, where necessary, to gain access to the program. There was no time limit for completing

any of the sections in the study, just as there was no time limit imposed for the actual program itself. We did this in order to ensure that the learning environment would be as authentic as possible. Each participant accessed a stand-alone version of the program that had been installed prior to the commencement of the session. The time that each participant spent on the multimedia was recorded electronically. Participants worked on an individual basis and were asked not to consult or discuss the program with their peers.

The cognitive styles of the participants (as determined by Riding's CSA) were not calculated during the first pilot study because the software and necessary license to use the test had not yet been received.

3.18.1.4 The pilot study experience

There were a number of unexpected technological problems that were beyond my power to resolve during this session. Several of them are irrelevant to this discussion. But the following problems decisively influenced the successful outcome of the pilot study.

- The computers did not seem to have enough RAM and/or a sufficiently fast processing speed to run the multimedia. Because several of the computers had to be re-booted during the session, participants had to start the program again. Three participants withdrew from the pilot study as a result of these problems.
- Most of the data was not written out to the .INI file – even though this had worked without error on all the computers used during the testing of the program prior to the pilot study.

These problems were unexpected and surprising because this laboratory is used by the Department of Anatomy who make extensive use of multimedia resources. I was also familiar with this laboratory. I had used multimedia there previously without any problem and had been assured by the IT personnel that the facility was suitable. What we eventually learned was that the maintenance of the facility should have occurred long before the study took place, but this had not yet taken place.

In spite of these obstacles, the pilot study proved to be a valuable experience because it gave us grounds for improving the design of the study.

The participants were observed using the program by myself and a research assistant. Our participation enabled us to make the following interesting observations:

- The participants seemed to spend a long time on the pretest – certainly more time than we had anticipated they would need. They were allowed to move backwards and forwards in the pretest, and were observed to do just that.
- The participants were required to use earphones to listen to the narrated animations. While they had been specifically briefed to do so, many of them were observed to have watched the animation without listening to the sound. We instructed these participants to use the earphones

and return to the beginning of the animation since it was impossible to understand the animation without the narration.

- It appeared to us that the placement and formatting of the text that accompanied the secondary task and that was used to measure the cognitive load of the format was somehow problematic. It may have been the case that the text was too small or that the contrast between the two colours used (black and purple) was too indistinct for comfortable visibility.

The placement of this secondary task on the screen is illustrated in Figure 3.6.

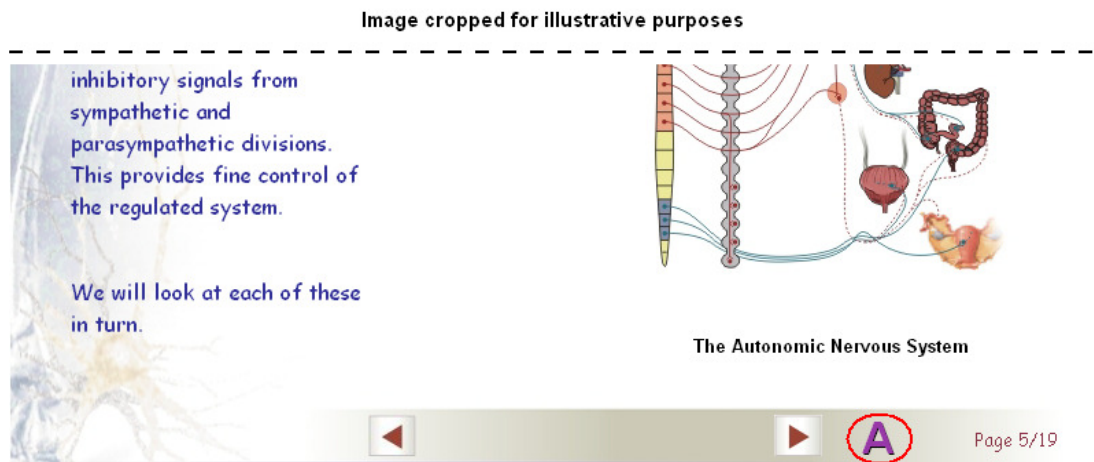


Figure 3.6: Display of secondary task in the program used in pilot study 1

The text, which was placed in the bottom right-hand corner of the screen, was formatted in 24-point Arial and the font colour was black. When the colour of the text changed, it changed to purple. Observation of the program as it was displayed on the computers used for the pilot study showed that the default screen resolution of the monitors influenced this contrast.

The participants' data logs were retrieved from the computers at the end of the pilot study session. Although very little data had been written out to the .INI file, the data that had been written out was the response time to the secondary task.

There were, however, few responses to the secondary task. A closer inspection of this data seemed to confirm our suspicion that the placement and formatting of the secondary task was partly or wholly responsible for the failure of the participants to respond to the secondary task. In spite of my early suspicion that the students had not responded because the cognitive load of the screen might have been too high, the participants did respond to this secondary task on screens with an expected higher cognitive load, but failed to respond to the secondary task on screens that clearly had a very low cognitive load.

3.18.1.5 Measures to address the problems of this pilot study

The following changes were made to the design of the multimedia and data collection instruments.

- The formatting of the secondary task used to measure cognitive load was changed. While the use of the letter 'A' was retained, the font size was increased from 24 to 36 points. The letter was also set in bold type, and lime green (RGB 0, 255,2) rather than purple was used for the colour change.
- Smith (2007) decided to extend her study and investigate whether the position of the secondary task on the screen influenced the cognitive load for the screen. Another two versions of the program were developed to accommodate this extension.
- A time limit of 10 minutes was imposed for both the pre- and posttest. (There was no time limit for the pencil-and-paper section of the posttest.)
- The option of moving backwards and forwards between the questions in the pre- and posttest was removed.
- A bookmarking facility was programmed into the multimedia. This was in anticipation of technological problems that might arise in the main study. This bookmarking feature allowed participants who had crashed out of the program because of technological problems or who had accidentally closed the program to return to the study at the point where they had left it. (The data collected up to that point was also retained and so could still be used.)

We made no changes to any of the actual questions in the various instruments because these had elicited no queries. We also made no changes to the pre- and posttest questions.

3.18.2 Pilot study 2

A second pilot study was scheduled and took place two weeks after the first pilot study.

3.18.2.1 Participants

Volunteers were sought for this study by advertising the session in two of the residences on the Faculty of Education campus of the University of Pretoria. The participants from each residence competed against each other in the posttest and the winning residence was given an extra allowance of printing paper in the computer laboratory. Forty-eight students volunteered to participate in this second pilot study. While they were all registered for different programmes, the majority of them were B Ed undergraduates. This was also a convenience sample.

3.18.2.2 Materials

The learning materials that were used have already been described in Part 2 of this chapter. No changes were made to the content or instructional strategies of the two formats. While the assessment questions remained unchanged, a time limit of 10 minutes was introduced for both the

pretest and posttest. The time for the test was displayed across the assessment, and the allotted number of minutes counted down on each screen (as illustrated in Figure 3.7). When the 10 minutes were up a message was displayed informing the participant that the time limit had been reached and that it was no longer possible to enter any answers. When that point was reached, the program branched automatically to the next section.

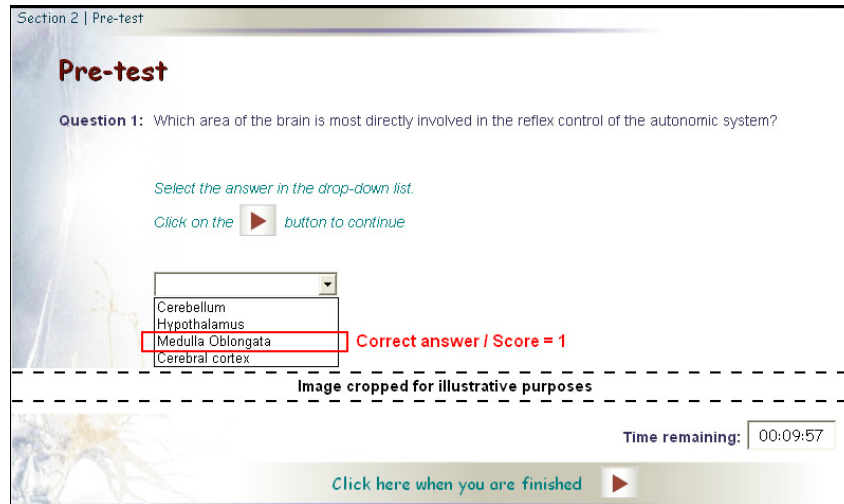


Figure 3.7: How remaining time was displayed on screen in the pretest

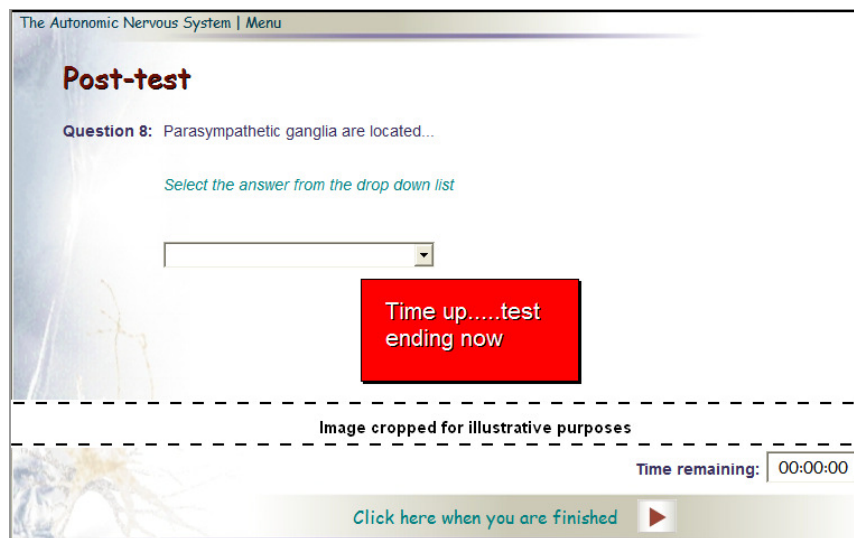


Figure 3.8: The message that was displayed in pre- and posttest once time had elapsed

The second change made to the materials used for Pilot study 2 was that four programs were used in this pilot study. Two programs were developed for each version and the only difference between the programs was the placement of the secondary task on the screen.

The formats are presented in the matrix below.

	Version 1 - Animation	Version 2 – Static Images
	Program 1	Program 3
Placement of secondary task	Bottom right-hand corner of screen (See Figure 3.9)	
	Program 2	Program 4
Placement of secondary task	Top right-hand corner of screen (See Figure 3.10)	

Table 3.9: Placement of the secondary task across four programs

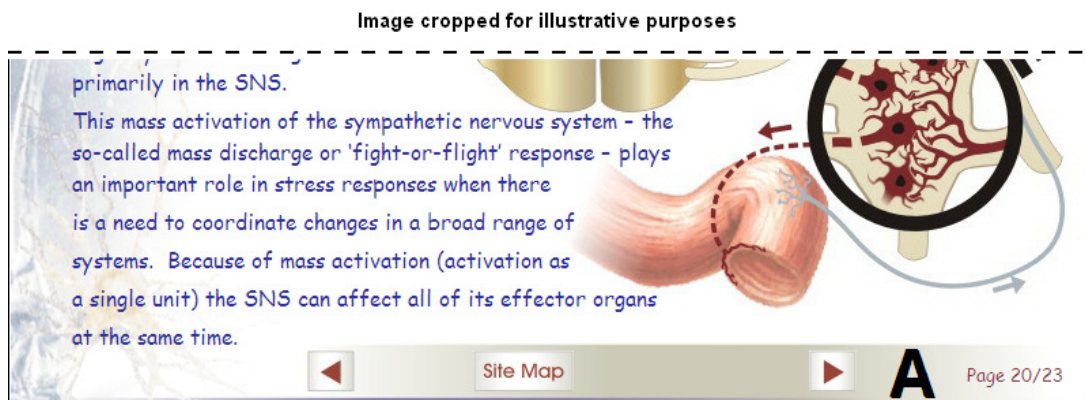


Figure 3.9: Placement of the secondary task in Programs 1 and 3

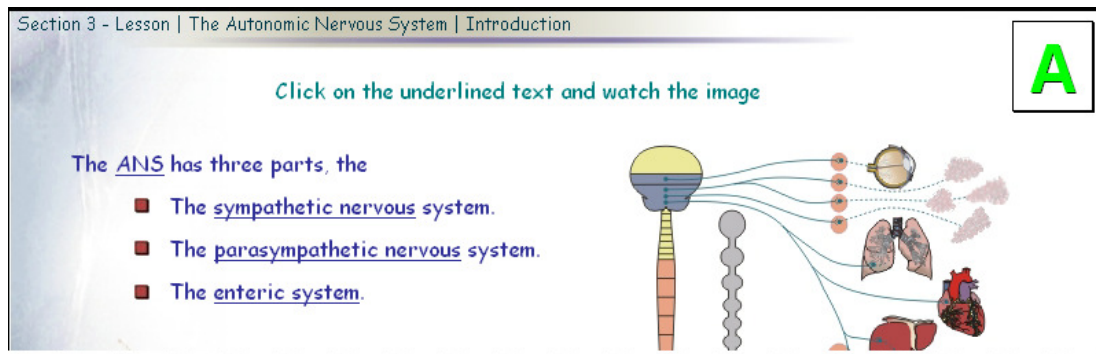


Image cropped for illustrative purposes

Figure 3.10: Placement of the secondary task in Programs 2 and 4

3.18.2.3 Procedure

The venue of the study was also relocated. The second pilot study took place in the computer laboratories on the Faculty of Education campus of the University of Pretoria. These laboratories had recently been equipped with brand new computers and their computers had been thoroughly checked and appraised by the Laboratory IT Administrator and his team. This team also tested the program on the computers before the pilot study commenced so that they could see whether the technological problems and problems with the data recording had in fact been resolved. The Laboratory IT

Administrator was trained to conduct the pilot study. Although I was not present at this pilot study (I was in hospital), Smith (2007) attended the session. The participants were randomly assigned to four different groups on arrival at the computer laboratory, and each group was randomly assigned to one of the four programs. The session was conducted in exactly the same way that is described in Section 3.18.1.3.

We were unable to establish the cognitive styles of the participants (as determined by Riding's CSA) during the second pilot study because we had not yet received the software and the necessary license to use the test.

3.18.2.4 Data collected

The problems that we had experienced in getting the data written out to an external log file for each participant were resolved during this pilot study. All the data was written out as expected, and the following data was made available:

- The demographic data
- The self-rating of mental effort
- The data used to measure cognitive load directly
- Data from the pretest and posttest
- The sequence in which the screens in the program were visited
- The time spent on each individual screen
- The time spent on the program

The transfer test (the second part of the posttest) was not given to this sample.

3.18.2.5 The pilot study experience

The implementation of the second pilot study proceeded smoothly and the participants appeared to understand the instructions for the different instruments. The one request that the users made was that they be allowed to move back to screens in the questionnaires as they sometimes wanted to make sure that they had completed all the questions. No other technological problems arose at this stage.

3.18.2.6 Measures to address the problems of this pilot study

No further revisions of the wording of the questions in the instruments were necessary. We considered the request that students be allowed to move backwards and forwards between the screens that displayed the questionnaires, and we agreed to it for all screens (except for the pre- and posttest sections).

3.19 Conducting the main study

This section describes the data collection process for the main study.

3.19.1.1 The participants

The Autonomic Nervous System (ANS) is taught in the second-year of the medical and dental curriculum in the block that deals with Homeostasis. Since this Physiology course is compulsory, the medical and dental students were grouped together for this course. The research population included all second-year medical and dental students at the University of Pretoria (there were 262 students who had registered for this course in the 2006 academic year). The Autonomic Nervous System was covered in week six of the eight-week course. This was the week that commenced on 27 February 2006.

3.19.1.2 Selection of the research sample

The multimedia session was presented as part of the normal series of lectures that covered the ANS, and some of the multimedia content was repeated in a lecture that took place after the multimedia session. Attendance at the multimedia session was therefore (as far as the faculty member responsible for this section of the course was concerned) compulsory. In spite of this requirement, I briefed the class about the purpose of the study and requested them to sign a consent form wherein they declared that they had understood the nature of the study and were willing to participate in the study.

The experimental design required that the sample be divided into two groups with an optimum sample size of 130. The sampling error formula method of determining sample size (Creswell, 2002) dictated that a group should be larger than the available research population (which was 262). The power analysis formula, however, pointed to a sample size in the region of 130. The sample for this study was 245. This is clearly smaller than the size required by the sampling error formula, but larger than that suggested by the power analysis formula. I consulted with a statistician prior to the commencement of the study (J. Fresen, personal communication, November 2005), and he advised me that a sample size of 262 would be adequate for the study.

The random sampling took place in two stages. The participants were first randomly assigned to one of sixteen groups by an application that had been developed by the University of Pretoria and that, among other features, randomly assigns class members into groups. The groups were then randomly assigned to a time period, laboratory and program. The final matrix for this assignment is displayed in Appendix M.

It was my hope that since the study covered content that was relevant to their studies, most of the students would be willing to participate. Two-hundred and forty-five students, all of whom had consented to participate in the study, arrived at the computer laboratories to work through the

multimedia program. While attrition in the sample reduced the number of participants available for analysis from 245 to 238, the number was still within the required limit.

3.19.1.3 Procedure

Participants were then randomly assigned to one of sixteen participant groups prior to a briefing about the study. The division of the sample into 16 groups was done both for logistical purposes and in order to be able to accommodate Smith's study (Smith, 2007). The actual number of participants assigned to each version was in fact initially 131.

Since the research was being conducted in an authentic learning environment where the window of opportunity for data collection was very small, the entire experiment had to be completed within a period of four hours. That was the time allocated by the course leader for the multimedia session. I had two computer laboratories at my disposal and each of these laboratories could accommodate 70 students. The pilot study had allowed us to determine that two hours would be sufficient for both the purpose of the instruction and for the data collection. We therefore scheduled two sessions back to back and used both the computer laboratories in the process.

The participants were briefed about the multimedia session and study in the larger group. This 20 minute briefing included information about the purpose of providing instruction in multimedia format, the purpose and design of the study, the different sections of the multimedia program and how each related either to the study or the instruction, and the secondary task and how to respond to it.

The students were then informed about their group allocation. They were requested not to swop groups.

On entering the laboratory each participant was provided with a handout that included the following:

- An informed consent form that they were required to sign at the start of the session. This form was collected by the research assistants.
- A page describing how to access the cognitive style test and the multimedia program.
- A page to record the result of the cognitive styles test.
- A page on which to make notes while studying the instructions. (This page was collected before the posttest was administered.)

The participants first completed the cognitive styles test. When they had completed the test, they were required to raise their hands. The result was recorded electronically by the software (CSA) and manually on the handout by the research assistant. Once the research assistant had recorded the result of the CSA, the participant was informed that he/she could continue with the multimedia session. In spite of the instructions provided in the handout, many participants still sought additional help from the research assistant. Such assistance was willingly given because the purpose of the

study was not to assess their level of computer literacy. It was for this reason that participants were allowed to ask for help during the session. The help that was given was limited to providing information about navigation and, in some cases, actual content. No help, other than help with navigation, was given while the participants were completing the pre- and posttests. The program required a password before the student could continue with the posttest. This password was entered by the research assistant who also used the opportunity to collect any notes that had been made and to provide the participant with the pencil-and-paper section of the posttest. All the handouts (including the paper-based posttest) were turned in by the participants as they left the laboratory. Participants were allowed to leave the laboratory as soon as they had finished because many of them had other commitments between the laboratory session and their next class.

3.19.1.4 The materials

I have already described the materials that were used in Part 2 of this chapter. Each participant accessed a stand-alone version of the program that had been installed prior to the commencement of the session. Participants worked on an individual basis and were asked not to consult or discuss the program with their peers. This request was fortunately adhered to in a very exemplary fashion by the participant group.

3.19.1.5 The data collected

The scores for the CSA that were recorded manually on the participant handout were entered into a spreadsheet which was then checked against the electronic logs of the test. All the other data was recorded in an .INI file that had been created for each participant. These .INI files were downloaded from the individual computers by the IT personnel in the laboratory and were given to me on a CD-ROM disk. The data collected was described in Section 3.18.2.4, and a sample of this file is presented in Appendix N.

3.19.1.6 The study experience

In general, the session went well. The IT personnel in the laboratory helped me to set up the computer laboratory prior to the two sessions. This included installing the software, testing it and labelling each computer with the version and program that had been loaded onto it. As the participants entered the laboratory they were directed to the first open computer that had the version to which they had been assigned. The research assistant could help those who had forgotten the version to which they had been assigned because they had been given a list of the randomised allocations. I did not involve myself in this activity because I did not want to be able to identify the versions to which participants had been allocated. (I was personally acquainted with some of the participants because I was working in the Faculty at that time and did not want to complicate matters or violate my ethical protocol.)

The staff of the IT laboratory also acted as research assistants during the session because there were as many as 70 students in the laboratory at once. There were a few students who experienced problems. These problems were caused mainly by the premature closing of the program. When that happened, they merely logged on again. And because of the bookmarking feature, they were able to continue exactly from where they had left off. Three of the students experienced major technological problems for reasons I cannot explain. While the IT personnel did try to resolve these problems, the data from these participants were excluded from the dataset used in the analysis. Data for 242 participants finally became available for analysis.

3.20 Summary

Chapter 3 described the research design and methodology in detail and did so in terms of the following four main sections:

- **Part 1:** A description and discussion of the research design and methodology
- **Part 2:** A description of the research intervention
- **Part 3:** The hypotheses
- **Part 4:** A description of implementation of the research design and methodology

Part 1 was summarised in Section 3.11 of this chapter and will therefore not be summarised here again.

The research intervention was a multimedia program that was developed to teach second-year medical and dental students selected content about the Autonomic Nervous System. This content formed part of the Homeostasis block that they are required to take at the beginning of their second year of study. Two formats were developed to teach the same content. Version 1 used narrated animation on selected screens and Version 2 used static images & text, which replaced the animation on one of the screens in Version 1. Version 1 had 19 screens and Version 2 had twenty-three.

Seventeen hypotheses were developed. These were written as null hypotheses and an alternative hypothesis was described and discussed for each null hypothesis.

Part 4 of this chapter described the actual implementation of the study, which, in its final form, consisted of two pilot studies and a main study. A second pilot study had to be staged because the first pilot study revealed that there were major problems that affected the data collection process.

Since electronic data collection was such an integral part of the intervention, I could not risk going into the main study without testing the data collection process again. But it all ended well because I only had to make very few changes to the research intervention and no changes at all to the wording or format of the items in the research instruments. The second pilot study and the main study proceeded without any problems at all, and, when the main study concluded, I was able to collect electronic data logs for 242 students for analysis and interpretation.

This analysis of the data will now be presented and discussed in Chapter 4.



Presentation and Analysis of Empirical Data

4

4.1 Introduction

This chapter presents the analysis of the empirical data for this study, which set out to explore the role cognitive load and cognitive style play in the achievement of learning outcomes, when using animation and static images as multimedia learning formats in an authentic learning environment. The study also investigated the relationship between cognitive load, which is influenced by both the nature of the content and the specific design strategies used, and the cognitive style of the individual, using the multimedia formats of animation and static images.

The demographic profile of the sample is described in detail, followed by the results of the analyses undertaken to find answers to the research questions. This analysis and discussion is presented and explored using the research question, the sub-questions and the accompanying hypotheses as a framework. Each major section of the chapter concludes with a summary of the process of analysis and subsequent findings.

I have already stated that the cognitive load and cognitive style research streams are each extensive in their own right. Merging these two fields increased the scope of the research beyond what is normally considered adequate for doctoral work. Since this study follows a relatively new line of research, the analysis undertaken for this study was both broad **and** deep, and extended beyond merely looking for answers that would support or reject the hypotheses. In addition, both these fields are under-represented in the health science education research. My study also addresses this shortcoming. While Chapter 4 presents this broad, deep analysis I will only interpret and discuss the findings related to the hypotheses in Chapter 5, unless the analysis results in unexpected findings, in which case such findings will also be discussed in Chapter 5.

Two hundred and forty five data logs were retrieved by the computer laboratory administrator from the individual computers used for the experiment. Once the data had been cleaned and prepared for analysis the final sample included 238 participants.

Participants were excluded from the final sample if

- they failed to answer any of the questions in the various electronic questionnaires **and** the pre- and posttests, in spite of retrieving a data file for the participant.
- data for **both** the cognitive styles analysis and the self-report of mental effort were incomplete.

4.2 Statistical analysis

The data was analysed by Dr M van der Linde, Department of Statistics at the University of Pretoria, using the SAS[®] application.¹

A review of the literature on research methodology indicates that it is generally accepted that researchers want three basic questions answered once the data has been collected and analysed:

- Is this effect real or can it only be attributed to chance?
- If it is real, how large is it? and
- Is it large enough to be useful? (Ellis & Steyn, 2003; Kirk, 2001; Vacha-Haase, 2001; Vaske, 2002; Winkleman, 2001)

These three goals also apply to this study. The first question is answered using null hypothesis significance testing, the result of which determines whether or not the observed effect is due to chance (Onwuegbuzie & Leech, 2004). The question about the size of the effects can be addressed using descriptive statistics, confidence intervals and effect sizes (Kirk, 2001; Onwuegbuzie & Leech, 2004). Determining whether the effect is real and of practical significance does involve an element of subjectivity. It requires a value judgment on the side of the researcher, which is influenced by, amongst others, the value system of the researcher, societal concerns and the assessment of costs and benefits. Kirk is of the opinion that researchers have an obligation to make this kind of judgment and that no-one is in a better position to do so than the person who collected and analysed the data (Kirk, 2001).

In this study descriptive statistics are reported as frequencies, means and standard deviations ($M \pm SD$) or standard error ($M \pm SE$). Inferential analyses of data were performed using regression analysis, the general linear model (GLM), t tests, Chi-square analysis, Pearson's correlation and effect sizes. Confidence intervals of 95% were reported. Differences were considered significant for p values ≤ 0.05 . This means that the probability to obtain a statistic as or more extreme, if the null hypothesis is true, is smaller than 5%. Of particular interest in inferential analysis is the R^2 value. This value indicates the extent to which the variables that are included in the analysis co-vary (Lowry,

¹ SAS Version 8.2 running under VM/CNS on the University of Pretoria's mainframe system

2007) or to put it in another way, R^2 is a measure of the strength of the relationship between two variables (Garson, 2006). The higher this value the better.

Effect sizes address the issue of statistical significance versus practical significance. The effect size is not dependent on sample size (Vache-Haase, 2001; Winkelman, 2001). The effect size is typically evaluated as small, medium, or large. One commonly used guideline is based on socio-behavioural research, where a small effect size is described as 0.2, a medium effect size as 0.5, and a large effect size as 0.8. An alternative guideline, based on psychotherapy research, describes 0.15 as a small effect size, 0.45 as medium, and large as 0.90 (Winkelman, 2001). The advantage of reporting effect size in an area of knowledge with a growing body of research allows the reader to determine whether there is support for the theory underpinning the research, judge the practical importance of the findings and compare findings across studies despite different analytic procedures (Winkelman, 2001).

I will use the guidelines provided by Cohen (1988) when comparing the means of two samples, where a small effect size is described as 0.2, a medium effect size as 0.5, and a large effect size as 0.8. It is reported as d . In a 2 x 2 table the phi (ϕ) coefficient (Ellis & Steyn, 2003), which in SAS is reported together with the Chi-square statistic, is the measure of the effect size and is reported as w . The guideline here is that a small effect is described as $w = 0.1$, a medium effect as $w = 0.3$ and a large effect as $w = 0.5$. I will use these guidelines in this study.

4.3 The profile of the participants

One hundred and ninety three (193) participants were registered for the MBChB programme (medical students) and 45 for the BChD programme (dental students).

The demographic issues explored in this study were

- age,
- gender,
- cultural group and
- home language of the participants.

Since prior knowledge (learner experience) influences the self-report rating of mental effort (Ayres, 2006a; Clarke, Ayres & Sweller, 2005) and learning performance (Ginns, 2003; Kalyuga, Chandler & Sweller, 2001) the issue of whether or not the participants had studied the topic previously was explored. Participants were asked to rate their prior knowledge on the topic. These findings are presented and discussed in this section on the profile of the participants.

4.3.1 Age, gender and cultural group

The gender profile of the research sample is presented in Figure 4.1.

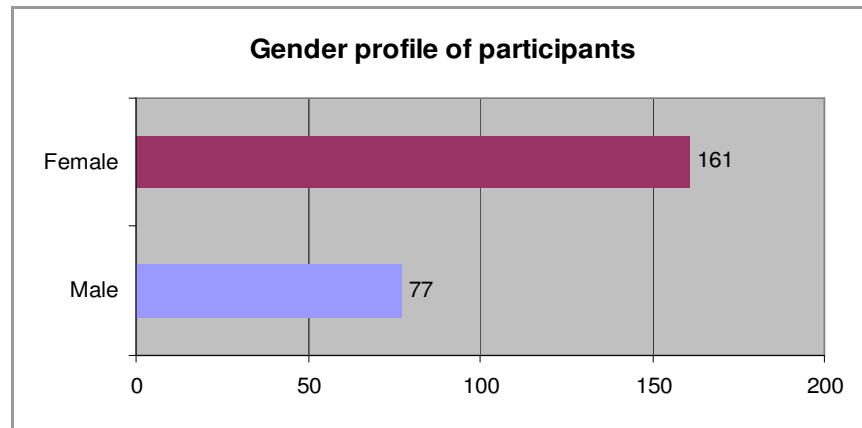


Figure 4.1: Gender profile of the sample

Looking at Figure 4.1 it is evident that the majority of the participants were female. The ratio of female to male students in this class is almost exactly 2:1. The number of female students in this sample does not reflect the gender profile in the South African population, where, using the 2001 Census data, the ratio of females to males for the 15 – 24 year group is almost equal (Statistics South Africa, 2005). The reason for this high proportion of female students reflects current strategy within the National Department of Education in South Africa. Increased enrollment of black and female students in what was historically a white institution for Afrikaans-speaking white South Africans is in line with the requirements for transformation in Higher Education, documented in the White Paper for Higher Education of 1997. One of the aims for higher education is to increase access for black, women, disabled and mature students (Department of Education, 1997).

The age range for the participants was 17 - 37 years, with the majority of the participants in the 19 - 20 year age range. A breakdown of the gender and age distribution is presented in Figure 4.2.

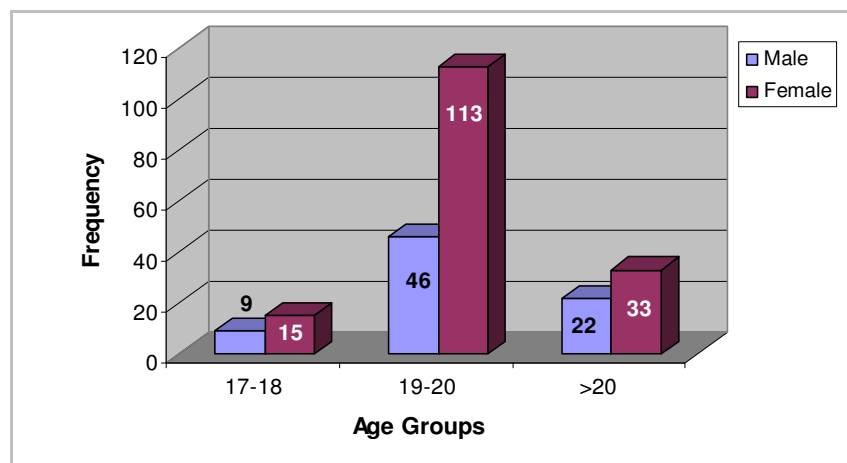


Figure 4.2: Profile of sample by gender and age

The school-leaving age in South Africa is typically between 17 – 18 years of age. The average age for the sample was expected to lie in the 19 - 20 range. Sixty four percent (64%) of the sample are in this expected age range. Figure 4.2 shows a slightly skewed distribution. Fifty five participants (23% of the sample) were 21 years or older. There are at least three factors that contribute to this distribution.

1. The demographic profile of the school leaving population has changed due to the political changes in South Africa since 1994. A larger percentage of school leavers are now in fact older than 18 years.
2. There are generally more applicants for the MBChB programme at the University of Pretoria than there are available places in the programme. Many students who qualify for entry to the MBChB programme, but are not accepted, enrol for a basic science degree (BSc). They complete the BSc degree, a 3–4 year programme, and then go on to enrol for the MBChB programme.
3. Students also transition to the MBChB programme before they complete their BSc degree. This is possible because they are given the opportunity to fill vacancies in the MBChB programme, which arise due to attrition from the MBChB programme.

Participants were randomly allocated to the two versions of the same program (the research interventions). The gender and age distribution of participants across the research interventions, after this random allocation, is displayed in Table 4.1.

	Animation version		Static images & text version	
	Male	Female	Male	Female
17-18	1	11	8	4
19-20	21	59	25	54
>20	13	15	9	18
TOTAL	35	85	42	76

Table 4.1: Distribution of sample across research intervention, by age and gender

Only one participant was 17 years of age at the time of the research intervention. Ten participants were older than 25 years, and two of these ten were older than 30 years. The oldest participant in the group was 37 years old. Factors that could contribute to this age range have already been discussed earlier in this section. They will not be explored further as they are beyond the scope of this study.

Students at the University of Pretoria are required to indicate their cultural group when enrolling. I collected this data in order to explore the effect of culture on cognitive style. Figure 4.3 displays the profile of the sample by cultural group.

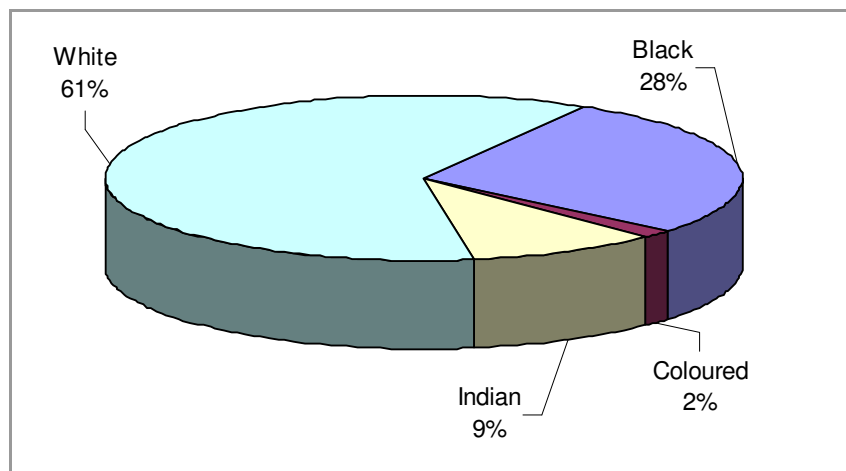


Figure 4.3: Profile of sample by cultural group

The majority of the sample (61%) were White students. Historically, Black and Indian medical students enrolled at the Medical University of South Africa (MEDUNSA) or the University of Durban-Westville respectively. Changes in the structure of South African universities (Department of Education, 1997) and the establishment of a new democracy in 1994 has seen these traditional patterns of enrollment change, with an increase in Black, Coloured and Indian students in universities that previously enrolled only White students. Further discussion of the cultural profile of the sample is beyond the scope of this study.

4.3.2 Home language of the sample

When discussing this study with different physiology lecturers a comment was made by one of them that she experienced that students who were not English first language speakers really battled to understand Physiology (L. Nagel, personal communication, 22 November 2005). Two of the factors Michael (2007) identified that contribute to the difficulty students have in understanding Physiology relate to language. Firstly, the language is mixed: by this the author meant situations where many commonly used words in the discipline take on specific scientific meanings that are different, and often opposite, from their lay meanings. The second factor was that faculty and textbook authors use language imprecisely, where imprecise refers to the fact that there is widespread use of jargon and acronyms, often to the detriment of learning. The University of Pretoria is a multi-cultural university, and tuition in the Faculty of Health Sciences is mostly through the medium of English. There are however 11 official languages in South Africa and the Black students represent a wide spread of cultural and ethnic groups, where mother tongue is not necessarily English.

Participants were asked to indicate whether English was their first, second, third or fourth (and higher) home language. The results are displayed in Figure 4.4.

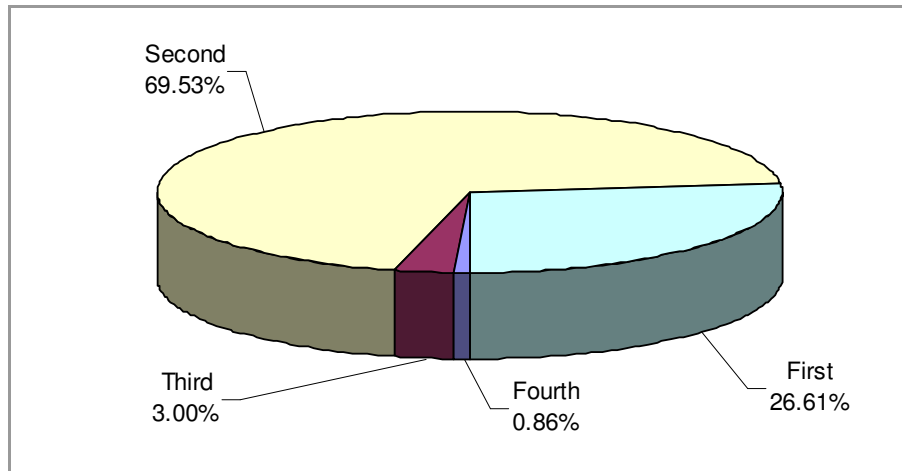


Figure 4.4: Percentage of participants where English is their home language

Only 26.6% of the sample indicated that English was their first home language. The majority of the group (69.5%) use another language as their first language at home. I did not explore which of the other 10 official languages were their first language. The possibility does exist that language plays a role in cognitive load and learning performance and this will be explored later in this chapter.

4.3.3 Prior knowledge of the participants

Prior knowledge was determined in three different ways:

- i. Asking participants whether or not they had studied the topic before.
- i. Asking participants to rate their knowledge and understanding of the physiology of the Autonomic Nervous System.
- ii. Using a pretest to determine knowledge at the level of recall and comprehension.

Data for the first two points listed above were obtained using self-report ratings. The results for these two questions will be presented and discussed in this section. The pretest results will be presented and discussed in Section 4.8 of this chapter. The responses to the question 'Have you studied this topic previously?' are presented in Table 4.2.

	Have you studied this topic previously?	
	Yes	No
Animation version	84	35
Static images & text version	83	34
TOTAL	167	69

Table 4.2: Number of participants who had studied topic previously

Nearly one third of the participant group (29.2%) indicated that they had never studied the topic before. The 70% who answered 'Yes' to this question could have studied aspects of this topic in Biology, a school subject, or while enrolled for a BSc degree in the School of Biological Sciences. It is also possible that they had touched on this topic in Anatomy and/or other units in the Physiology curriculum. I did not interview any of the participants to further explore their responses to this question due to constraints with regard to access to the sample after the session in the computer laboratory.

Participants were asked to rate the level of their knowledge about the topic (The Autonomic Nervous System), illustrated in Figure 4.5. Four options were provided. I think I know and understand

- absolutely nothing about the topic.
- the basic concepts of the topic.
- the concepts beyond a basic understanding.
- the topic at an expert level.

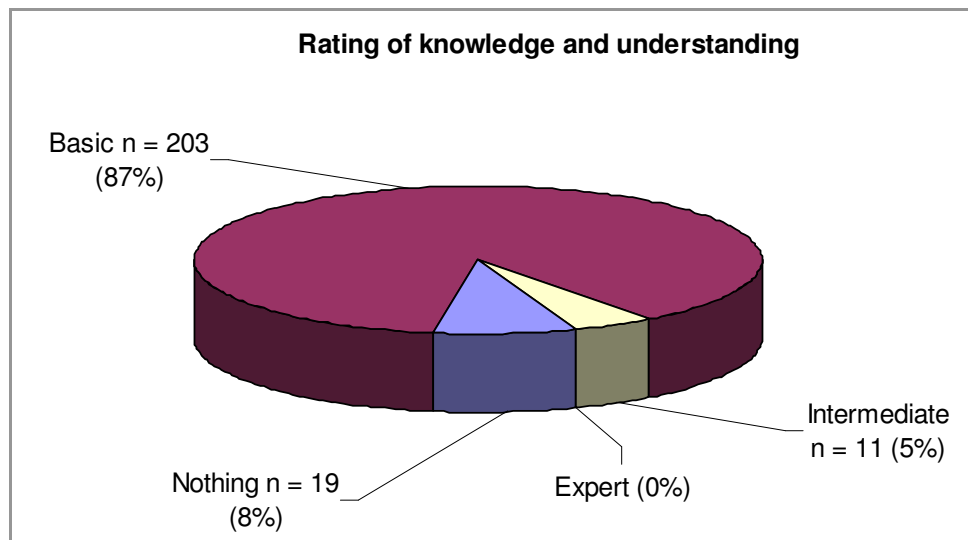


Figure 4.5: Rating of level of knowledge and understanding

From Figure 4.5 we see that only 8% of the sample indicated that they know nothing about the topic. Most of the sample (87%) indicated that they understood the topic at a basic level. None of them rated themselves as experts on the topic.

Cognitive load theory, discussed and illustrated in Chapter 2, proposes that task characteristics and subject characteristics interact to influence the amount of mental effort invested in learning. Asking the participants to rate their knowledge and understanding of the subject takes the simple question 'Have you studied this topic before?' one step further. Perceptions about prior knowledge might influence motivation to learn. Participants who feel that they already know the content might not invest much mental effort in understanding the content. However, by asking them to indicate how much mental effort they actually did invest to understand the content it is possible to start looking at

how accurate their perceptions are, particularly if their learning performance is also considered. This rating addresses prior knowledge from a subjective perspective, while the pretest, which is discussed in Section 4.8 of this chapter, addresses prior knowledge from an objective perspective.

4.3.4 In summary: profile of the participants

The profile of the sample used for this study can be summarised as follows: The majority (68%) of the participants are female. The majority (81%) of the participants are medical students. The only other learner group represented in this sample are dental students. The same Physiology course is followed by both these groups at the University of Pretoria, which explains why they are in the same class as the medical students. Due to the fact that this study does not intend to compare cognitive style, cognitive load or learning performance across these two groups no further use will be made of this difference in the learning programme being followed. The majority (64%) of the sample are between 19 – 20 years of age.

The group is culturally diverse and for the purpose of this study four cultural groups were identified: Black, Coloured, Indian and White participants. The majority of the sample are White students (61%) Only 26% of the participants use English as their home language. The majority (70%) of the participants indicated that they had studied this topic previously, but also indicated that they only had a basic understanding of the content.

I now move on to explore the time spent on each version of the program. Time is considered when discussing both cognitive style and cognitive load. The next section analyses the time spent without considering the issues of cognitive load and style.

4.4 Time spent on each version of the intervention

This section looks at the time spent on each version in general, irrespective of cognitive load or style.

4.4.1 Time spent on full program

The animation version had 19 screens. During the testing of the program (before the pilot studies) it was determined that it took users between 30 – 45 minutes to work through the animation version. The participants in this study spent between 15.03 and 76.76 minutes studying the content of the animation version, with a Mean (\pm SD) of 42.35 (\pm 11.59) minutes. There were outlier values on both sides of the range. Two participants spent just over 15 minutes studying the content (15.03 and 15.5 minutes respectively) and two spent more than 65 minutes studying the content (68.9 and 76.76 minutes respectively). The remaining times were all more than 19 minutes and less than 65 minutes.

The static images & text version had 23 screens, and during testing of the program it took users between 45 – 75 minutes to work through the program. The participants in this study spent between

22.86 and 92.76 minutes studying the content of the static images & text version, with a Mean (\pm SD) of 45.63 (\pm 14.01) minutes. There were also two outlier values, 87.45 and 92.76 minutes. The remaining times were all under 83 minutes.

The time spent on the program was divided into three groups. The three 'time' groups were labeled Inadequate (IA), Adequate (AD) and More than Adequate (MA). This categorisation used the time taken during the program testing (before the pilot studies) as a guideline for setting the time parameters for each of the three groups.

The amount of time assigned to each of these three time groups, for each version, is presented in Table 4.3.

	Time - Animation version	Time - Static images & text version
Inadequate (IA)	< 30 min	< 40 min
Adequate (AD)	30 – 50 min	40 – 65 min
More than Adequate (MA)	> 50 min	> 65 min

Table 4.3: Time groupings for each version of the program

The frequencies and mean times, for each time group, by version are presented in Table 4.4 (see next page), together with the results of the t test to determine the statistical significance of the means in each category of time.

	Animation			Static images & text			Tests for significance		
	n	Mean	SD	n	Mean	SD	t test	P value	Effect size (d)
Inadequate	21	25.4429	4.3994	46	33.0065	5.1138	-7.3100	< 0.0001	1.59
Adequate	72	41.6669	6.1519	62	49.9255	6.6207	-10.6849	< 0.0001	1.29
More than Adequate	27	57.3198	5.6094	10	77.1050	9.2193	-9,.8279	< 0.0001	0.65

Table 4.4: Mean time spent on program for each version

With reference to Table 4.4, the results of the t tests for independent samples returned differences between the two versions that were highly statistically significant ($p < 0.0001$). The effect size was also determined for each group and for the categories 'Inadequate' and 'Adequate time' a very large effect size was returned. The effect size for the 'More than Adequate' time category is a medium-to-large effect. These results, indicating both statistical and practical significance, must still be interpreted cautiously as there were more screens in the static images & text version. The results only provide a big picture view of the time spent on the program. They do not indicate exactly where in the program the time was spent. A screen wise comparison, which follows in the next section, will shed more light on this.

The range in time spent studying the content in both versions is wide. Of concern is the number of participants spending less than 30 minutes studying the animation version (17.5 % of the sample) and less than 40 minute for static images & text version (39% of the sample). Due to the time pressures for completing the data collection and the fact that the participants were not available after the laboratory session it was not possible to interview the participants to determine the reasons why some of the participants spent so little time studying the content. No assumptions can be made about these reasons.

4.4.2 Time spent on individual screens

This section looks briefly at the time spent on specific screens across the two versions of the program. The screens of interest are screen 12 in the animation version and screens 13-16 in the static images & text version. Screen 12 in the animation version presented the content using an animation that was 1 minute and 45 seconds in duration. Screens 13-16 in the static images & text version replaced the animation (screen 12). These four screens will be treated as one unit for the purpose of this analysis and discussion.

Since the participants were allowed to go back to previous screens the total time each participant spent on each of these screen units was calculated.

The participants spent between 2 and 23.85 minutes studying the content of screen 12 in the animation version, with a Mean (\pm SD) of 5.4328 (\pm 3.2353) minutes. Four of the participants spent more than 11 minutes on this screen. The participants accessed this screen between 1 and 6 times.

The participants spent between 2.466 and 16.283 minutes studying the content of screens 13-16 in the static images & text version, with a Mean (\pm SD) of 7.5558 (\pm 2.9871) minutes. There were no outlier values.

Using the t test for independent samples, the analysis indicated that the difference between the mean time spent on screen 12 and screens 13-16 was statistically significant, $t = -5.26$, $df = 235$, $p < 0.0001$. The effect size was $d = 0.710$, indicating practical significance. This finding was expected as

there were more screens to access in the static images & text version (screens 13-16). In fact I expected the time to be even longer for screens 13-16.

Sixty six participants never accessed screen 16 in the static images & text version. Access to this screen was from screen 14. Participants were required to click on the small magnifying glass, illustrated in Figure 4.6 using the red circle and arrow.

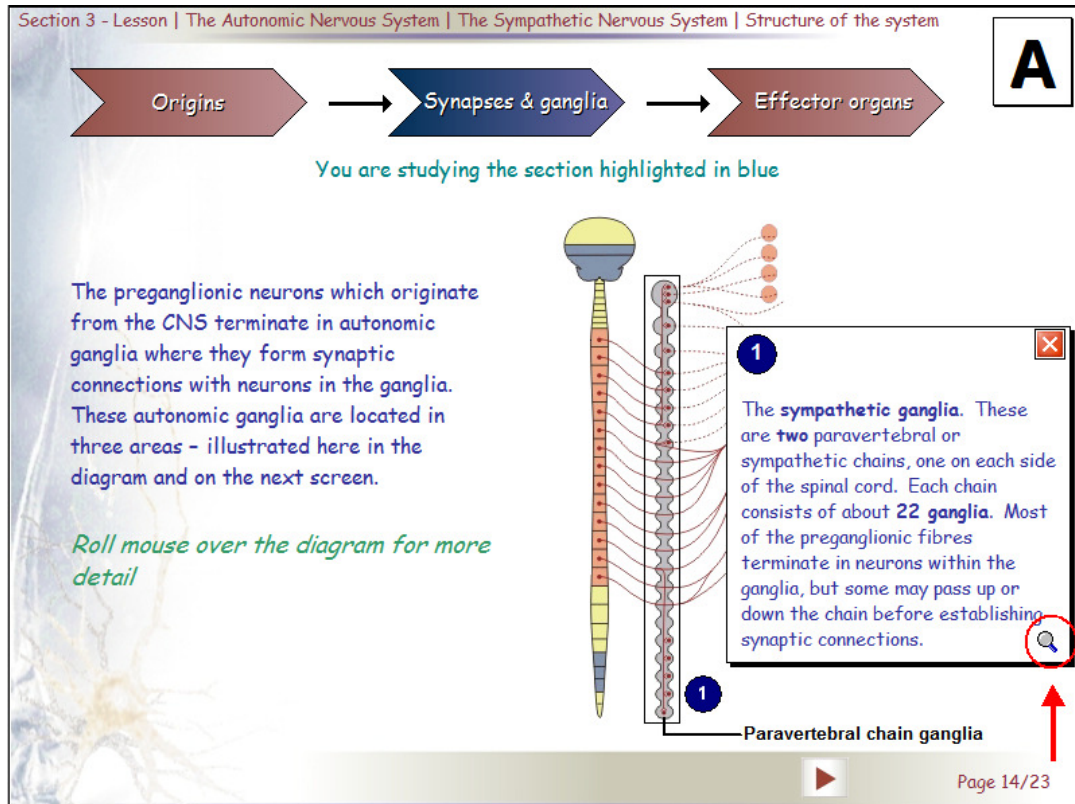


Figure 4.6: Access to screen 16 in the static images & text version

There were no instructions directing the participant to specifically click on this magnifying glass. This turned out to be a design flaw that resulted in 50% of the participants who used this version missing screen 16. The content for screen 16 can be viewed in Appendix I.

The number of times these screens were accessed was determined separately for each screen.

The frequencies for the number of times the screens were accessed are displayed in Table 4.5.

	Screens				
	Vers 1 (n=120)	Version 2 (n = 118)			
	12	13	14	15	16
No of times accessed	N	N	N	N	N
1	59	87	40	64	40
2	31	20	39	31	8
3	16	8	24	13	4
4	10	2	9	8	0
5	2	1	1	1	0
6	2	0	1	0	0
More than 6	0	0	4	1	0

Table 4.5: Frequency of access for screen 12 in the animation version and screen 13-16 in the static images & text version

In the animation version (n = 120) the number of participants who viewed screen 12 once or between 2 – 4 times was almost equal: 59 and 57 respectively. In the static images & text version (n = 118) this pattern was different. More than half of the participants only accessed screen 13 once, most of the participants viewed screen 14 between 2 and 4 times and the number accessing screen 15 either once or between 2–4 times was almost equal. The one advantage of the static images & text version was that the learner had more control over how they viewed the content. For example, if they wanted to view a small section of the animation (screen12) again they either had to run the full animation or use the control bar and scroll through the animation until they found the place they were looking for. I did not explore whether users found this action of 'scrolling' through an animation irritating or not, but as an experienced user myself I find this method cumbersome. In the static images & text version the user can click through the screens very quickly. It is much easier to scan the content and find the place you are looking for because there is no movement on the screen as is the case when viewing an animation. The design of the animation in this study also required the participant to actively start the animation. This and other types of user control and usability and the cognitive load associated with this type of navigation could be investigated further in future studies.

4.4.3 In summary: time spent on the program

This section first considered the time spent on the program in general. The participants spent between 15.03 and 76.76 minutes studying the content of the animation version and between 22.86 and 92.76 minutes studying the content of the static images & text version. The time spent on the program was divided into three groups, which were labeled Inadequate (IA), Adequate (AD) and More than Adequate (MA). The effect size was also determined for each group and for the categories 'Inadequate' and 'Adequate time' a very large effect size was returned. The effect size for the 'More than Adequate' time category was in the medium-to-large range. These results, indicating both statistical and practical significance, must still be interpreted cautiously as there **were** more screens in the static images & text version. The results only provide a big picture view of the time spent on the program. Of concern was the number of participants spending less than 30 min studying the animation version (17.5 % of the sample) and less than 40 min for static images & text version (39% of the sample).

Screen-wise comparisons, where the screens of interest were screen 12 in the animation version and screens 13-16 in the static images & text version, were conducted. The analysis indicated that the difference between the mean time spent on screen 12 and screens 13-16 was statistically significant. The effect size was $d = 0.710$, indicating practical significance. This finding was expected as there were more screens to access in the static images & text version.

I now move on to explore the role cognitive style plays in an authentic multimedia learning environment.

4.5 Exploring the role cognitive style plays in an authentic multimedia learning environment

This section discusses the cognitive styles of the participants and explores the use of the multimedia within the context of their particular styles. The analysis included looking at whether gender and culture are indicators of cognitive style.

The first sub-question asked in this study was 'What are the cognitive styles of the participants taking part in the study?' The two hypotheses, discussed in Chapter 3, are:

Null Hypothesis 1a

There will be no difference in the percentage of the sample having either an Analytic / Intermediate or Wholistic style on the WA style dimension.

Alternate Hypothesis 1a

More than 50% of the sample will have an Analytic or Intermediate style on the WA style dimension.

Null Hypothesis 1b

There will be no difference in the percentage of the sample having either a Verbaliser /Bimodal or Imager style on the VI style dimension.

Alternate Hypothesis 1b

More than 50% of the sample will have a Verbaliser or Bimodal style on the VI style dimension.

4.5.1 Cognitive style as measured using Riding's CSA

The cognitive style of each participant was measured using Riding's Cognitive Style Analysis (CSA). The CSA provides a score for each dimension in the cognitive style model. The ratios for each style dimension typically range from 0.4 through to 4.0 with a central value around 1.0 (Riding, 2005a).

In this study the Wholist-Analytic (WA) ratios of 235 participants (3 participants did not do the CSA) ranged from 0.70 to 4.47 (Mean = 1.57, SD = 0.5898). The Verbaliser-Imager (VI) ratios for these 235 participants ranged from 0.39 to 2.14, (Mean = 1.0868, SD = 0.1983). The WA ratios are above the 0.4 described by Riding for the lower end of the WA style dimension (Riding, 2005a) but just over the upper range of 4.0, indicating a shift to the Analytic end of the dimension for the sample in this study. The VI ratios for this study are just above the 0.4 at the one end of the continuum. The upper limit of the VI ratio for this study, 2.14, is well below the upper limit suggested by Riding.

The correlation between the Wholist-Analytic and Verbaliser-Imager style ratios, using Pearson's Correlation Coefficient, was low ($r = 0.0372$) and not significant ($p = 0.5706$). This is in line with correlations reported in the literature (Douglas & Riding, 1993; Rezaei & Katz, 2004; Riding & Grimley, 1999; Riding & Rayner, 1998; Riding & Sadler-Smith, 1992; Riding & Staley, 1998). The results indicate the independence of the two style dimensions, which is also described in the cognitive style model.

4.5.1.1 The Wholistic-Analytic dimension

The literature describes both two and three points along the Wholistic-Analytic dimension, used primarily for the purpose of grouping the participants. Using three categories, the profile of the sample is displayed in Table 4.6.

	WA ratio	Male	Female	TOTAL
Wholist	≤ 1.02	11 (14.7%)	25 (15.6%)	36
Intermediate	> 1.02 and ≤ 1.35	18 (24.0%)	47 (29.4%)	65
Analytic	> 1.35	46 (61.3%)	88 (55.0%)	134
TOTAL		75	160	235

Table 4.6: Profile of the sample for the WA dimension of the CSA using three style groups

The percentages in Table 4.6 are column percentages. Looking at the total number of participants in each style group we see that more than half of the sample are Analytic in style: fifty-seven percent ($n = 134$) of the sample had an Analytic style, 28% ($n=65$) had an Intermediate style and only 15% ($n=36$) had a Wholistic style.

Table 4.6 also provides a gender perspective of the style results. Sixty one percent (61%) of the male participants and 55% of the female participants were Analytic in style. Only 24% of the males were in the Intermediate style range for the WA dimension, while 29% of the females had an Intermediate style.

If only two categories of WA style are used, gender is not included in the analysis, and a ratio of 1.20 is used as the dividing point along the WA continuum, then the data indicates that 70% of the participants were Analytic in the way they process and organise information and 30% were Wholistic.

*The results from the analysis of the WA style dimension **support** hypothesis 1a that more than 50% of the sample will have an Analytic or Intermediate style on the WA style dimension.*

4.5.1.2 The Wholistic–Analytic dimension: looking at gender and culture

This section explores the WA style dimension and the relationship between WA style, gender and culture in more detail. As discussed in Chapter 2, many studies look at style and gender, but I could find no literature related to cognitive style and culture.

The mean WA style ratios for the sample, displayed according to gender, are presented in Table 4.7.

	Ratios	Male		Female	
		N	Mean (SD)	N	Mean (SD)
Wholist	≤ 1.02	11	0.8836 (0.1040)	25	0.8476 (0.0916)
Intermediate	>1.02 and ≤1.35	18	1.1878 (0.0944)	47	1.1949 (0.0892)
Analytic	> 1.35	46	1.9565 (0.6184)	88	1.9375 (0.4435)

Table 4.7: WA style ratios according to gender

A Chi-Square analysis, appropriate to test null hypotheses for categorical data, indicated that there was no significant difference between the gender groups for the Wholistic-Analytic style dimension, $\chi^2 = 0.9232$, $df = 2$, $p = 0.6303$. The phi coefficient, a measure of the strength of the relationship between gender and WA style is 0.0627. This is a very small effect indicating that in practice there is no relationship between gender and WA style.

The mean WA style ratios, displayed according to culture, are presented in Table 4.8.

	Black		Coloured		Indian		White	
	N	Mean	N	Mean	N	Mean	N	Mean
Wholist ≤ 1.02	5	0.8400 (0.1091)	1	1.0000 (n.a.)	2	0.7250 (0.0353)	28	0.8664 (0.0888)
Intermediate >1.02 & ≤1.35	15	1.2153 (0.0983)	1	1.2400 (n.a.)	7	1.1257 (0.0810)	42	1.1950 (0.0859)
Analytic >1.35	44	2.0107 (0.6193)	2	1.8050 (0.3182)	14	2.1186 (0.6027)	74	1.8751 (0.4058)

Table 4.8: WA style ratios according to culture

These four categories for culture were collapsed into two groups for further analysis: White participants and Non-white participants (included Black, Coloured and Indian participants).

A stepwise regression analysis was performed using the Wholistic-Analytic cognitive style as the dependent variable. This approach was followed in order to determine which of the factors influence cognitive style. The following variables, and all their interactions, were included in the regression analysis:

- Version of the multimedia
- Age
- Gender
- Cultural group
- Pretest scores

The result of the regression was $F(2, 231) = 6.10$, $p = .0026$, $R\text{-square} = 0.0502$ and $C(p) = 0.9483$. The value for $C(p)$ (known as Mallow's coefficient) provides an indication of the completeness of the model. The value of this coefficient should be as close as possible to the number of parameters in the regression. The variables retained in the stepwise regression were culture and the interaction between version and gender.

The stepwise regression equation is presented in Table 4.9.

Variable	Parameter Estimate	Standard Error	Type III Sum of squares	F-value	p-value
Intercept	2.1166	0.1628	56.0932	169.00	<.0001
Culture	-0.2521	0.0777	3.4932	10.52	0.0014
Version x Gender	-0.0545	0.03467	0.8221	2.48	0.1169

Table 4.9: Stepwise regression equation for WA style

From the data it can be determined that there was a statistically significant main effect for culture, $F(2, 231) = 10.52$, $p = 0.0014$, but not for the interaction between version and gender ($p = 0.1169$).

A confirmatory General linear model (GLM) was then run, based on the results from the stepwise regression. The GLM presents the standard error and p-difference tables, the results of which are not provided in the stepwise regression. The GLM allows comparison of the class variable effects (version, gender and culture) using pair wise least squares means comparisons. The model used for the GLM was: WA cognitive style = version, gender, culture, and the interaction between version and gender.

The GLM analysis returned a statistically significant finding, $F(3, 230) = 3.17$, $p = 0.0146$, $R^2 = 0.052$. The culture of the participants accounted for the significant result, $F(1, 230) = 9.12$, $p = 0.0028$, $R^2 = 0.052$. Further post hoc comparison, using Fischer's Least Squares Means, indicated that the non-White participants were statistically significantly more Analytic than the White participants. The

non-White participants had a Mean (\pm SE) WA ratio of 1.7268 (0.0639) and for the White participants it was 1.4909 (0.0497).

This effect size (d) was calculated using the following formula:

$$d = \frac{|\bar{x}_i - \bar{x}_j|}{\sqrt{MSE}}$$

where

\bar{x}_i and \bar{x}_j are the group means and MSE is the Mean Standard Error of the ANOVA (Ellis & Steyn, 2003).

The effect size obtained was 0.40. This effect size is bordering on the medium effect size range for this study. This is a visible effect. This conclusion is made cautiously as there are not enough similar comparisons (cognitive style and culture) in the cognitive style literature against which to benchmark this finding.

This concludes the section on the Wholistic-Analytic dimension of cognitive style. The next section considers the Verbaliser-Imager dimension of cognitive style.

4.5.1.3 The Verbaliser-Imager dimension

This section presents the analysis of the data to determine the Verbaliser-Imager profile of the sample used in this study.

The literature describes both two and three points along the Verbaliser-Imager dimension, used primarily for the purpose of grouping the participants. Using three categories the profile of the sample is displayed in Table 4.10.

	VI ratio	Male	Female	TOTAL
Verbaliser	≤ 0.98	19 (25.4%)	45 (28.1%)	64
Bimodal	> 0.98 and ≤ 1.09	25 (33.3%)	49 (30.6%)	74
Imager	> 1.09	31 (41.3%)	66 (41.3%)	97
TOTAL		75	160	235

Table 4.10: Profile of the sample for the Verbaliser-Imager Dimension of the CSA using three style groups

The percentages in Table 4.10 are column percentages, reflecting distribution for gender per style sub-group. The data in this table indicates that the sample appears to lean slightly more to the Imager

style. Forty one percent (n=97) of the sample had an Imager style, 31% (n=74) were Bimodal in style and only 28 % (n=64) were Verbaliser in style.

Table 4.10 also provides a gender perspective of the style results. Forty-one point three (41.3%) of both the male and female participants were Imagers. The distribution of the Bimodal style was also very similar for male and female participants (33.3% and 30.6% respectively). Only 25.4% of the males and 28.1% of the females were Verbalisers.

If only two categories of VI style are used, gender is not included in the analysis, and a ratio of 1.035 is used as the dividing point along the VI continuum, the data indicates that 40% of the participants are Verbalisers with regard to their inclination to represent information and 60% are Imagers in style. The results therefore indicate that the Imager style was strongly represented in this sample.

*The results from the analysis of the VI style dimension do **not** support the hypothesis that more than 50% of the sample will have a Verbaliser or Bimodal style in the VI style dimension.*

4.5.1.4 The Verbaliser-Imager dimension: looking at gender and culture

This section explores the Verbaliser-Imager (VI) style dimension and the relationship between VI style, gender and culture in more detail. The mean VI style ratios for the sample, displayed according to gender, are presented in Table 4.11.

	Ratios	Male		Female	
		N	Mean (SD)	N	Mean (SD)
Verbaliser	≤ 0.98	19	0.8800 (0.1193)	45	0.9002 (0.0943)
Bimodal	> 0.98 and ≤ 1.09	25	1.0392 (0.0294)	49	1.0445 (0.0288)
Imager	> 1.09	31	1.2842 (0.2448)	66	1.2301 (0.1572)

Table 4.11: VI style ratios according to gender

A Chi-Square analysis, appropriate to test null hypotheses for categorical data, indicated that there was no significant difference between the gender groups for the Verbaliser-Imager style dimension, $\chi^2 = 0.2652$, $df = 2$, $p = 0.8758$. The phi coefficient, a measure of the strength of the relationship between gender and VI style is 0.0336. This is a very small effect size indicating that in practice there is no relationship between gender and VI style.

The mean VI style ratios, displayed according to culture, are presented in Table 4.12.

	Black		Coloured		Indian		White	
	N	Mean	N	Mean	N	Mean	N	Mean
Verbaliser (≤ 0.98)	20	0.8920 (0.1322)	1	0.8200 (n.a)	8	0.9100 (0.0431)	35	0.8940 (0.0941)
Bimodal > 0.98 & ≤ 1.09	22	1.0545 (0.0292)	1	1.0500 (n.a.)	4	1.0500 (0.0264)	47	1.0359 (0.0276)
Imager (> 1.09)	22	1.2990 (0.2290)	2	1.1750 (0.0210)	11	1.1854 (0.0784)	62	1.2424 (0.1894)

Table 4.12: VI style ratios according to culture

These four categories for culture were collapsed into two groups for further analysis: White participants and Non-white participants (included Black, Coloured and Indian participants).

A stepwise regression analysis was performed using the Verbaliser-Imager cognitive style as the dependent variable. This approach was followed in order to determine which of the factors influence cognitive style. The following variables, and all their interactions, were included in the regression analysis:

- Version of the multimedia
- Age
- Gender
- Cultural group
- Pretest scores

The result of the regression was $F(1, 232) = 5.69$, $p = 0.0179$, R-square = 0.02 and $C(p) = -4.46$. Only one variable was retained in the stepwise regression and that was the pretest score. The stepwise regression equation is presented in Table 4.13.

Variable	Parameter Estimate	Standard Error	Type III Sum of squares	F-value	p-value
Intercept	0.9629	0.0533	12.5958	325.71	<.0001
Pretest	0.0106	0.0044	0.2200	5.69	0.0179

Table 4.13: Stepwise regression equation for VI style

No further analysis was conducted for the VI style dimension. The data does not indicate that there is any relationship between gender or culture and the VI dimension of style.

4.5.1.5 In summary: cognitive style profile of the participants in this study

Section 4.5.1 reported on the cognitive style profile of the sample used in this study. The relationship between cognitive style, gender and culture for both the WA and VI style dimension was explored. The profile of the sample in this study was that of a group that was predominantly Analytic (69% of the sample) and Imager in style (60% of the sample).

Regarding style and gender, the results of the analyses indicate that there was no statistically significant difference between the gender groups for both the Wholistic-Analytic style dimension, $\chi^2 = 0.0232$, $df = 2$, $p = 0.6303$, and the Verbaliser-Imager style dimension, $\chi^2 = 0.2652$, $df = 2$, $p = 0.8758$. The phi coefficient of 0.0627 for the WA style and 0.0336 for the VI style are very small effect sizes and suggest that in practice there is no relationship between gender and cognitive style.

Regarding the WA cognitive style dimension and culture, a stepwise regression determined that there was a statistically significant main effect for culture, $F(2, 231) = 10.52$, $p = 0.0014$. This was followed up with confirmatory general linear modeling (GLM), which returned a statistically significant finding, $F(3, 230) = 3.17$, $p = 0.0146$, $R^2 = 0.052$. The culture of the participants accounted for the significant result, $F(1, 230) = 9.12$, $p = 0.0028$, $R^2 = 0.052$. Further post hoc comparison, using Fischer's Least Squares Means, indicated that the non-White participants were significantly more Analytic than the White participants. This effect size (d) obtained was 0.40. This effect size is in the small to medium range, but is a visible effect, although the interpretation is made with caution.

A stepwise regression with the VI style as the dependent variable did not return a statistically significant finding for any of the demographic variables entered into the equation.

4.5.2 Cognitive style and time spent on the program

One of the early questions asked while planning this study was 'What do learners actually DO with multimedia programs?' One perspective of the answer is to look at the time learners spend studying content in a multimedia program. A refinement of this question is 'Is the time spent on the program related to cognitive style and/or load?' This section looks at the amount of time spent on the program in context of cognitive style.

The distribution of the WA and VI style across the two versions of the program is presented in Table 4.14.

	Wholistic	Analytic	Verbaliser	Imager
Animation version	37	83	50	70
Static images & text version	36	79	44	71
TOTAL	73	162	94	141

Table 4.14: Frequency of WA and VI styles for each version of the program

4.5.2.1 The Wholistic – Analytic dimension and time spent on the program

The hypothesis tested was:

Null Hypothesis 1c

There will be no difference between the Analytic and Wholistic learner in the time spent studying the content of the program.

Alternate Hypothesis 1c

The Analytic learner will spend more time studying the content of the program than the Wholistic learner.

The mean time spent on each version, for the WA style is presented in Table 4.15.

	Wholistic		Analytic	
	n	Mean (SD)	n	Mean (SD)
Animation version	37	40.505 (12.812)	83	43.172 (10.980)
Static images & text version	36	40.957 (13.748)	79	47.537 (13.795)

Table 4.15: Comparison of time spent on program for WA style and version

From Table 4.15 we observe that for both versions the Analytic learner spent more time on average studying the content. This provides tentative support for Hypothesis 1c. The Analytic learner will spend time processing the detail of the content, while the Wholistic learner will scan more quickly to get a big picture view.

A series of t tests for independent samples was performed to determine if the difference in the means were significant. Looking at the animation version a t test for significance indicated that the difference between the Wholistic and Analytic style was not statistically significant, $t = 1.17$, $df = 118$, $p = 0.246$. The effect size was $d = 0.20$, which is small and therefore also not practically significant.

The time the Analytic learner used to study the content of the static images & text version was statistically significantly more than the time the Wholistic learner used, $t = 2.37$, $df = 113$, $p = 0.0192$. The effect size was $d = 0.47$. This effect size is close to the medium effect size range and could be regarded as being a visibly longer time for Analytic.

Hypothesis 1c is therefore supported for the static images & text version of the program, but not for the animation version.

The data in Table 4.15 also allows a comparison of the time each style spent on the different versions. The Wholistic learner spent almost the same amount of time for each version (40.5 minutes on the animation version and 40.9 minutes on the static images & text version) in spite of the fact that there were more screens in the static images & text version. This difference in time was not statistically significant, $t = -0.15$, $df = 71$, $p = 0.8847$. The effect size was also very small and therefore not of any practical significance ($d = 0.03$).

The Analytic learner spent an average of 43.1 minutes on the animation version and 47.5 minutes on the static images & text version. Although this difference in time is statistically significant, $t = -2.23$, $df = 160$, $p = 0.0269$, this result must be interpreted cautiously as there were more screens in the static images & text version. The effect size obtained was also small ($d = 0.31$). Proportionally the Analytic learner spent slightly more time studying the content of the animation version.

The analysis then drilled down to look at the time spent by the Analytic and Wholistic learner at screen level. The screens of interest were screen 12 in the animation version and screens 13-16 in the static images & text version. The mean times for each screen are summarised in Table 4.16.

	Wholistic		Analytic	
	N	Mean (SD)	n	Mean (SD)
Screen 12: Animation version	37	5.4000 (3.4539)	83	5.4474 (3.1547)
Screens 13 – 16: Static images & text version	36	6.7194 (3.4538)	79	7.8211 (2.601)

Table 4.16: Comparison of time spent on selected screens for WA style

Once again Table 4.16 allows for comparison from two different perspectives. Firstly it is possible to look at the version (at screen level) and compare style differences. The second comparison looks at each style individually and compares the time spent across the two versions.

When comparing the mean time between the styles for each screen we see in Table 4.16 that there is virtually no difference between the Wholistic and Analytic participant in the time spent studying the animation. A t test did not return a statistically significant difference in these mean times, $t = -0.07$, $p = 0.9413$. The effect size was also very small $d = 0.02$. Looking at Table 4.16 we also see that both the Wholistic and Analytic learner spent more time on average studying the content of screen 13-16 in the static & images version. There is however a larger difference between the Wholistic and Analytic participants in the mean time spent studying the static images & text version, but it was also not statistically significant at an alpha of 0.05, $t = -1.89$, $p = 0.0607$. The effect size however does point to a visible effect ($d = 0.60$).

I then considered the time **each style group** spent on the different screens. In considering the Wholistic style, a t test for independent samples did not return a statistically significant value, $t = -1.63$, $df = 71$, $p = 0.107$, when comparing the mean time of screen 12 and screens 13-16. The effect size was $d = 0.38$, which suggests that the difference might be visible. The result for the Analytic style indicated that the difference in the mean time for screen 12 and screens 13-16 was statistically significant, $t = -5.21$, $df = 160$, $p < 0.0001$. The effect size was $d = 0.69$, indicating a visible effect.

At screen level the data suggests that hypothesis 1c is supported for the static images & text version of the program, but not for the animation version.

Using the time groups defined earlier (Inadequate, Adequate and More than Adequate) a 3 x 2 contingency table was used to analyse the data for time and style for the animation version. This analysis is presented in Table 4.17.

Animation version	Wholistic		Analytic	
	N	Col %	N	Col %
Inadequate time spent	10	27.03	11	13.25
Adequate time spent	19	51.35	53	63.86
More than Adequate time spent	8	21.62	19	22.89

Table 4.17: Comparison of time spent on the animation version for the WA style dimension

A Chi-Square analysis, appropriate to test a null hypothesis for categorical data, indicated that there was no significant difference in the time groups for the two styles, Wholistic and Analytic, $\chi^2 = 3.4597$, $df = 2$, $p = 0.1773$. The effect size was $w = 0.1698$. This is a small effect and cannot be regarded as practically significant.

A similar analysis was done for the static images & text version and the data is presented in Table 4.18.

Static images & text version	Wholistic		Analytic	
	N	Col %	N	Col %
Inadequate time spent	22	61.11	23	29.11
Adequate time spent	11	30.56	49	62.03
More than Adequate time spent	3	8.33	7	8.86

Table 4.18: Comparison of time spent on the static images & text version for the WA style dimension

A Chi-Square analysis, appropriate to test a null hypothesis for categorical data, indicated that there was a significant difference in the time groups for the two styles, Wholistic and Analytic, $\chi^2 = 11.1727$, $p = 0.0037$. The phi coefficient, a measure of the strength of the relationship between the time spent on the program and WA style is 0.3117. This effect size indicates that in practice this difference in time spent by the different style groups is visible. Sixty-two percent of the Analytic learners spent adequate time on the program, in contrast to only 30.56% of the Wholistic learners who spent adequate time on the program. Of concern here is the fact that 61% of the Wholistic learners spent inadequate time on the program compared to only 29.11% of the Analytic learners.

I now turn to the analysis of the Verbaliser-Imager dimension and the amount of time spent on the program by this group of participants.

4.5.2.2 The Verbaliser-Imager dimension and time

The hypothesis tested was:

Null Hypothesis 1d

There will be no difference between the Verbaliser and Imager learner in the time spent studying the content of the program.

Alternate Hypothesis 1d

The Verbaliser will spend less time studying the content of the program than the Imager.

The mean time spent on each version, for the VI style is presented in Table 4.19.

	Verbaliser		Imager	
	n	Mean (SD)	n	Mean (SD)
Version 1: Animation	50	40.936 (10.517)	70	43.359 (12.270)
Version 2: Static images & text	44	46.183 (12.912)	71	45.04 (14.229)

Table 4.19: Comparison of time spent on program for VI style and version

From Table 4.19 we observe that the participants with an Imager style spent more time on average studying the content of the animation version. Using the t test for independent samples, the analysis indicated that the difference between the styles for the animation version was not statistically significant, $t = -1.13$, $df = 118$, $p = 0.2605$. The effect size was $d = 0.20$, which is a small effect and can be regarded as having no practical significance.

The two style groups spent about the same time on average studying the content of the static images & text version. The difference was not statistically significant, $t = 0.42$, $df = 113$, $p = 0.6738$. The effect size was $d = 0.08$, which is very small and cannot be regarded as practically significant.

Hypothesis 1d, which states that the Verbaliser will spend less time studying the content of the program than the Imager is therefore **not supported** for two versions of the program.

The second hypothesis tested was:

Null Hypothesis 1e

There will be no difference in the amount of time the Verbaliser learner spends proportionally studying the two versions of the content.

Alternate Hypothesis 1e

The Verbaliser will spend proportionally more time studying the content of the static images & text version than the animation version.

When comparing the time the Verbaliser learner spent by version the data in Table 4.19 indicates that the Verbaliser learner spent more time on average studying the static images & text version ($M [\pm SD] = 46.183 [12.912]$) than studying the content in the animation version ($M [\pm SD] = 40.936 [10.517]$). Proportionally however, because of the difference in the number of screens across the two versions, the Verbaliser learner spent more time studying the content in the animation version than the static

images & text version, and therefore the t test for independent samples, which indicated that the difference between the time for the different versions was significant [$t = -2.08$, $df = 92$, $p = 0.0406$], might not be that relevant for this comparison. The effect size for the difference in the means was $d = 0.38$, which indicates that this result might be visible.

An analysis was done of the time spent by the Verbaliser learner and Imager learner at screen level. The screens of interest were screen 12 in the animation version and screens 13-16 in the static images & text version. The mean times for each screen are summarised in Table 4.20.

	Verbaliser		Imager	
	n	Mean (SD)	n	Mean (SD)
Screen 12 in animation version	50	5.5403 (3.716)	70	5.356 (2.869)
Screens 13 – 16. in Static images & text version	44	7.0394 (2.7435)	71	7.7469 (3.0188)

Table 4.20: Comparison of time spent on selected screens for VI style

Once again Table 4.20 allows comparisons from two different perspectives. Firstly it is possible to look at the version (at screen level) and compare style differences. The second comparison looks at each style individually and compares the time spent across the two versions

When comparing the mean time between the styles for each screen we see in Table 4.20 that there is a small difference between the Verbaliser and Imager participant in the mean time spent studying the animation. A t test did not return a statistically significant difference in these mean times, $t = 0.31$, $p = 0.7597$. The effect size was also very small $d = 0.05$. There is however a larger difference between the Verbaliser and Imager participant in the mean time spent studying the static images & text version, but it was still not statistically significant at an alpha of 0.05, $t = -1.26$, $p = 0.208$. The effect size is also small ($d = 0.23$) and not of any practical significance.

I then considered the time **each style group** spent on the different screens. In considering the Verbaliser style, a t test for independent samples returned a statistically significant value, $t = -2.20$, $df = 92$, $p = 0.0304$ when comparing the mean time of screen 12 and screens 13-16. The effect size was $d = 0.40$, which suggests that the difference is visible. The result for the Imager style indicated that the difference in the mean time for screen 12 and screens 13-16 was also statistically significant, $t = -4.82$, $df = 139$, $p < 0.0001$. The effect size was $d = 0.79$, indicating a practically significant difference.

Using the time groups defined earlier (Inadequate, Adequate and More than Adequate), a contingency table was used to analyse the data for time and style for the animation version. This analysis is presented in Table 4.21.

Version 1	Verbaliser			Imager		
	N	Row %	Col %	N	Row %	Col %
Inadequate time spent	11	52.38	22	10	47.62	14.29
Adequate time spent	29	40.28	58	43	59.72	61.43
More than Adequate time spent	10	37.04	20	17	62.96	24.29

Table 4.21: Comparison of time spent on the animation version for the VI style dimension

A Chi-Square analysis, appropriate to test a null hypothesis for categorical data, indicated that there was no significant difference in the time groups for the two styles, Verbaliser and Imager, $\chi^2 = 1.2871$, $df = 2$, $p = 0.0.5254$. The effect size was $w = 0.1036$, which is small and does not indicate practical significance.

A similar analysis was done for the static images & text version and is presented in Table 4.22.

Version 2	Verbaliser			Imager		
	N	Row %	Col %	N	Row %	Col %
Inadequate time spent	17	37.78	38.64	28	62.22	39.44
Adequate time spent	24	40	54.55	36	60	55.70
More than Adequate time spent	3	30	6.82	7	70	9.86

Table 4.22: Comparison of time spent on the static images & text version for the VI style dimension

A Chi-Square analysis, appropriate to test a null hypothesis for categorical data, indicated that there was no significant difference in the time groups for the two styles, Verbaliser and Imager, $\chi^2 = 0.3702$, $df = 2$, $p = 0.8310$. The effect size was $w = 0.0567$, which is very small and does not indicate practical significance.

4.5.2.3 In summary: cognitive style and time spent on program

This section considered the time spent on the program from the perspective of the different cognitive style dimensions. Looking at the animation version we note that the Analytic learner spent more time studying the program than the Wholistic learner but the difference was not statistically significant. The

effect size was $d = 0.20$, which is small and therefore also not practically significant. For the static images & text version the time the Analytic learner used to study the content was statistically significantly more than the time the Wholistic learner used. **Hypothesis 1c** was therefore supported for the static images & text version of the program ($p = 0.0192$), but not for the animation version ($p = 0.246$).

A comparison was also done for the amount of the time each style group spent on the different versions. The Wholistic learner spent almost the same amount of time for each version in spite of the fact that there were more screens in the static images & text version. This behaviour would seem to confirm the approach a Wholistic learner takes to learning, namely that of scanning in order to get a big picture. This reduced the time spent on the static images & text version, even though there were more screens in the static images & text version. Proportionally the Wholistic learner spent more time studying the content of the animation version. This is quite possible since the Wholistic learner had to watch the whole animation in order to understand it and could not scan the information. The Analytic learner spent less time studying the animation version than the static images & text version and the difference in time across the two versions was statistically significant ($p = 0.0269$). The effect size obtained was also small ($d = 0.31$). Proportionally the Analytic learner spent more time studying the content of the animation version.

Comparisons were then carried out at screen level, where once again the screens of interest were screen12 in the animation version and screens 13-16 in the static images & text version. When comparing the mean time between the styles for each screen there was virtually no difference between the Wholistic and Analytic participant in the time spent studying the animation. There was however a larger difference between the Wholistic and Analytic participant in the mean time spent studying the static images & text version, but it was also not statistically significant. The effect size however was in the medium-to-large range indicating practical significance. In considering the time **each style group** spent on the different screens it was found that the time spent by the Wholistic learner was not statistically different when comparing the mean time of screen 12 and screens 13-16. The effect size was $d = 0.38$, which suggests that the difference might be visible. The result for the Analytic style indicated that the difference in the mean time for screen 12 and screens 13-16 was statistically significant, with a large effect size with practical significance.

Sixty-two percent of the Analytic learners spent adequate time on the program, in contrast to only 30.56% of the Wholistic learners. Of concern here is the fact that 61% of the Wholistic learners spent inadequate time on the program compared to only 29.11% of the Analytic learners.

The last section considered the Verbaliser-Image dimension. Participants with an Imager style spent more time on average studying the content of the animation version but this difference was not statistically significant and the effect size was small. While not significant the Verbaliser did spend more time studying the content of the static images & text version, which had proportionally fewer

images than the animation version, and the Imager spent more time studying the content of the animation version, which had proportionally more images than the static images & text version.

An analysis was done of the time spent by the Verbaliser learner and Imager learner at screen level. The screens of interest were screen 12 in the animation version and screens 13-16 in the static images & text version. When comparing the mean time between the styles for each screen a small difference was found between the Verbaliser and Imager participant in the mean time spent studying the animation and it was not statistically significant. There was a larger difference between the Verbaliser and Imager participant in the mean time spent studying the static images & text version, but it was still not statistically significant. Effect sizes were small.

In considering the Verbaliser style, this participant spent more time on screens 13-16 than on screen 12 and the difference was statistically significant ($p = 0.0304$). In considering the Imager style, this participant spent more time on screens 13-16 than on screen 12 and the difference was statistically significant ($p < 0.0001$), with a large effect size pointing to practical significance.

The next section considers cognitive style and the use of the multimedia program.

4.5.3 Cognitive style and use of the multimedia program

Section 4.5.2 presented the findings related to the amount of time spent on the program by participants with different styles. The question asked was 'What do learners actually DO with multimedia programs?' This section analyses the data from another perspective, namely navigation patterns. The data collected also included the exact sequence of navigation through the program for each participant. These navigation patterns were diverse. This section will report on the following:

- A qualitative impression of navigation patterns
- The number of participants who went through the program once from start to finish
- Selected screens that were accessed more than twice and the cognitive style profile of the participants

4.5.3.1 A qualitative impression of navigation patterns and observation of participants during the computer session

The participants were allowed to make notes during the learning session. These notes were collected before they started the posttest. These notes were merely scanned in order to get a general impression of note-taking practices in the sample.

Scribble pages, the name given to this page hand-out for each participant, were collected from 244 participants. Fifty-two participants did not make any entries on this page. The pages received from the remaining 192 participants varied from very detailed, including the use of different colours of pen,

to almost nothing. Some participants attempted to write down everything for each screen, while others made detailed notes for selected screens. Some participants only made summaries, which were mostly key words and key ideas. Some of these pages were extremely neat and organised and others were almost undecipherable. There was a combination of textual and diagrammatic notes. The diagrams varied from basic drawings of the autonomic nervous system to the use of structures that resembled mind maps, with text and arrows.

A detailed analysis of these notes is beyond the scope of this study. Detailed analysis of these notes is an avenue for further research. Of particular interest would be to determine if the different style groups approached note-taking differently.

4.5.3.2 Participants who accessed the program once

Only 19 of the 238 participants worked through the program once, completing the program in a linear fashion. An analysis of the cognitive styles of these participants is presented in Table 4.23.

		Animation	Static images & text
WA style dimension	Wholistic	7	3
	Analytic	7	2
VI style dimension	Verbaliser	8	2
	Imager	6	3

Table 4.23: Cognitive style of participants who accessed the program once only

The pattern that emerges in Table 4.23 indicates that in each version the number of participants accessing each program only once were about the same for each of the style groups. These low frequencies make analysis difficult and the results must be interpreted cautiously. More participants accessed the animation version more than once.

A Chi-Square analysis indicated that there was no significant difference between the versions of the program for the Wholistic-Analytic style dimension, $\chi^2 = 0.1478$, $df = 1$, $p = 0.7007$. The phi coefficient, a measure of the strength of the relationship between version and WA style was -0.0882. This very small effect size indicates that in practice there is no relationship between version, WA style and the single access to the program.

A Chi-Square analysis indicated that there was no significant difference between the versions of the program for the Verbaliser-Imager style dimension, $\chi^2 = 0.4343$, $df = 1$, $p = 0.5099$. The phi coefficient, a measure of the strength of the relationship between gender and VI style is 0.1512. This small effect size indicates that in practice there is no relationship between version, VI style and the single access to the program.

4.5.3.3 Cognitive style profile of participants accessing selected screens more than twice

The majority of the participants entered selected screens more than twice. The screens of interest at this point are screens 2, 12 and 19 in the animation version, and screens 2, 13-16 and 23 in the static images & text version.

Screen 2, which was identical in both versions, had two functions. Firstly it provided the user with a big picture view of the entire program. Secondly, once the participant had worked through the program once they could use this screen as a menu to go back to whatever section of the program they wished to review again. This screen is illustrated in Figure 3.5 on page 148 of this thesis and in Appendix H.

I looked at the pattern of access for the animation version first. Looking at the WA style dimension, screen 2 was accessed between 2 and 16 times by the participants with access between 2 and 8 times accounting for the majority of the access. Looking at the VI style dimension screen 2 was also accessed between 2 and 16 times with access between 2 and 7 times accounting for the majority of the access.

I then looked at the pattern of access for screen 2 for the static image & text version. For both the WA and the VI style dimension, screen 2 was accessed between 5 and 23 times by the participants. Access between 5 and 11 times accounted for the majority of the access for both style dimensions.

The pattern that emerged was that screen 2 was used by some participants as a menu. This superficial analysis could not answer the question of whether the Wholistic learner used screen 2 to get the big picture view of the program. More detailed analysis of the path of navigation by each participant after they had accessed screen 2 is beyond the scope of this study, but could provide an avenue for future research.

A series of Chi-square analyses were conducted to determine if there was any relationship between the number of times screen 12 and 19 were accessed and the cognitive styles of the participants. Table 4.24 displays the results for the animation version for the two style dimensions, including the size effect (w).

	Wholistic-Analytic Dimension				Verbaliser-Imager Dimension			
	χ^2	df	p	w	χ^2	df	p	w
Screen 12	1.3636	3	0.7141	0.21	3.3767	3	0.3371	0.34
Screen 19	4.4211	3	0.2194	0.40	3.7238	3	0.2920	0.35

Table 4.24: Results of Chi-square analyses to determine relationship between style and access in the animation version

Table 4.25 displays the results for static images & text version for the two style dimensions, including the size effect (w). Screens 13-16 were analysed as a group and individually, with the exception of screen 16 due to the fact that there were so few entries into screen 16.

	Wholistic-Analytic Dimension				Verbaliser-Imager Dimension			
	χ^2	df	p	w	χ^2	df	p	w
Screen 13-16	12.5269	15	0.6388	0.3300	21.7566	15	0.1143	0.4350
Screen 13	0.9167	2	0.6323	0.2887	0.9167	2	0.6323	0.2887
Screen 14	8.2036	6	0.2236	0.4709	4.7591	6	0.5751	0.3586
Screen 15	1.1703	3	0.7601	0.2256	0.6192	3	0.8920	0.1641
Screen 23	3.0866	4	0.5434	0.3446	6.8014	4	0.1468	0.5115

Table 4.25: Results of Chi-square analyses to determine relationship between style and access in the static images & text version

None of the Chi-square analyses returned a significant finding. The values of many of the cells in each of the contingency tables set up for the Chi-square analyses presented in Tables 4.25 and 4.26 were smaller than 5. Chi-square analysis may therefore not be a valid test. In each case the Fischer's Exact Test was applied, but there were no significant findings either. It must be concluded that there seems to be no relationship between cognitive style and the number of times the screens were accessed. Other analyses, beyond the scope of this study might shed more light on the relationship between cognitive style and the pattern of navigation in multimedia learning.

4.6 Exploring the role cognitive load plays in an authentic multimedia learning environment

Four sub-questions were asked about the cognitive load of the two interventions used in this study.

The first sub-question explored was 'How do the participants rate the cognitive load of selected multimedia content?'

The 9-point rating scale developed by Paas (1993) was used in this study. Cronbach's coefficient alpha, using all the recorded measures of cognitive load, was used to determine the internal consistency of the scale. This coefficient was calculated for each version of the research intervention. The results obtained were $\alpha = 0.76$ for the animation version and $\alpha = 0.82$ for static images & text

version. These are acceptable values, indicating that this scale was found to be internally consistent for this study.

4.6.1 Self-report of cognitive load

4.6.1.1 The animated version

One hundred and twenty participants (n=120) were given the animation version. The self-reported cognitive load, on the scale of 1 – 9, was measured five times across the 19 screens of content. Table 4.26 summarises the content for the animation version that was studied prior to asking the participants to indicate the amount of mental effort they felt they had to invest in order to understand the content. The actual screens are illustrated in Appendices G - K.

* Self report cognitive load (SRCL)

Screens	SRCL	Question phrased
Screen 5 of 19: Different parts of the Autonomic Nervous System (ANS). Using text as a hyperlink, the user viewed three different images that each illustrated the different parts of the ANS. They could toggle between these different views to compare the difference visually.	SRCL1	Indicate how much mental effort you used to study the information on the screen you have just reviewed.
Screen 12 of 19: Animation of 1 min 45 secs in duration covering the structure and function of the sympathetic nervous system. User could pause, stop and restart the animation and view it as often as needed before moving on.	SRCL2	Indicate how much mental effort you used to study the information in the animation you have just reviewed.
Screen 13 of 19: Two concepts were described, using text and a static image for each. User clicked on a concept and viewed the content in a pop-up.	SRCL3	Indicate how much mental effort you used to study the information on the screen you have just reviewed.
Screen 17 & 18 of 19: The first screen was an interactive screen where users could select an organ and see a visual representation of the effect of the SNS and PNS on the organ. They could toggle between the two and compare the effect. The second screen summarised these effects in a table, using text.	SRCL4	Indicate how much mental effort you used to study the information on the two screens you have just reviewed.
Screen 19 of 19: Animation of 1 min 15 secs in duration explaining the neurotransmitters that function at the synapses in the ANS. User could pause, stop and restart the animation and view it as often as needed before moving on	SRCL5	Indicate how much mental effort you used to study the information on the screen you have just reviewed.

Table 4.26: Overview of content for the animation version on which self-report of cognitive load was based

With the exception of one measurement, the rating was based on the content viewed in the screen that preceded the question. The fourth time the participants had to rate mental effort invested they had to consider two screens. These screens were viewed one after the other and participants could go backwards and forwards between these screens as often as needed before moving on. The participants could **not** go back to view the screen in order to make the decision about how much mental effort they had invested. They only answered the cognitive load question once. If they returned to the screen later, which was possible using the back button in the program, they were not asked to self-rate the mental effort invested a second time.

The rationale for this design, which only measured the self-report of cognitive load once, was to control for the effect of prior knowledge and previous learning on the self-report of cognitive load. The study was designed to measure cognitive load the first time the content was viewed in the program, rather than investigate how cognitive load changed with increasing exposure to the content. However, since the learning environment was an authentic one it was deemed ethically appropriate to allow the learner to view the content as often as they needed prior to taking the posttest, and so they were allowed to go back to content.

Cognitive load was categorised into low (1 - 3.9), medium (4 – 6.9) and high (7 – 9).

The frequencies for the distribution of these cognitive load categories, for each time the cognitive load was measured (SRCL1 to SRCL5) are presented in Table 4.27.

	Self-report Cognitive Load				
	Screen	n	Low	Medium	High
SRCL1	5 of 19	117	14	91	12
SRCL2	12 of 19	114	12	71	31
SRCL3	13 of 19	114	36	70	8
SRCL4	17/18 of 19	113	12	77	24
SRCL5	19 of 19	112	14	78	20

Table 4.27: Frequencies of cognitive load reported as low, medium and high for the animation version

Table 4.27 indicates that a higher number of participants rated screens 12, 17 & 18 and 19 as requiring high mental effort. The animation on screen 12 had the highest percentage of high cognitive load ratings (27%), followed by screens 17 & 18 (21%) and screen 19 (17%) respectively. Screen 13 had the highest percentage (31.5%) of low cognitive load ratings compared to the other screens, and the lowest percentage (7%) of high cognitive load ratings.

The means and standard deviations for each instance of cognitive load measurement in the animation version, ranked from high to low, is presented in Table 4.28.

SRCL	Screen	n	Mean (SD)	Minimum	Maximum
SRCL2	12 of 19	114	5.76 (1.59)	1	9
SRCL4	17 & 18 of 19	113	5.56 (1.45)	2	9
SRCL5	19 of 19	112	5.25 (1.63)	1	9
SRCL1	5 of 19	117	4.96 (1.31)	1	8
SRCL3	13 of 19	114	4.22 (1.62)	1	9

Table 4.28: Mean cognitive load ratings for individual measurements of cognitive load in the animated version

Using the same categories of low (1 - 3.9), medium (4 – 6.9) and high (7 – 9) cognitive load, it would appear from the data in Table 4.28 that the mean ratings are all within the range for medium cognitive load, with the mean value for SRCL2 (the animation on screen 12) approaching the higher end of the medium scale. The highest cognitive load was found for screen 12, the animation, and the lowest load for screen 13, a screen that used active text and pop-ups. The pop-ups each displayed one static image and text. The design of screen 13 is illustrated in Appendix H.

This data suggests that animation requires more cognitive resources from working memory than do static images and text, resulting in a higher self-report rating of mental effort invested. The influence of presentation formats on cognitive load will be explored in more detail in the next section.

4.6.1.2 The influence of presentation formats on cognitive load in the animation version

The fourth sub-question asked about cognitive load was ‘To what extent do the presentation formats influence cognitive load?’

A series of univariate analyses for pair-wise dependent samples was conducted. The results of the Sign test (M) indicated whether the difference in the measured cognitive load of the two screens being compared was significant. This test is appropriate for dependent samples where the data has been collected using a rating scale (Garson, 2006). The results from the two-tailed test, together with the associated *p* values, are reported. The Bonferroni correction (McMillan & Schumacher, 2001) was applied to these *p* values.

The results are displayed in Table 4.29. The screen listed in the column marked 1 (one of the presentation formats) had a higher cognitive load than the screen listed in the column marked 2 (the second presentation format).

Screen-wise comparison of cognitive load			
Column 1	Column 2	Sign test statistic (M)	p
Screen	Screen		
12	17/18	0.5	> 0.95
12	19	14.5	0.0024*
12	5	-23	0.0008*
12	13	32.5	0.0008*
17/18	19	12.5	0.0176*
17/18	5	-22	0.0008*
17/18	13	-30.5	0.0008*
19	5	-7	> 0.05
19	13	-21	0.0008*
5	13	21.5	0.0008*

* Alpha $p < 0.05$

Table 4.29: A comparison of the cognitive load for selected screen pairs in the animation version

From Table 4.29 we see that there are statistically significant differences for all the comparisons excepting the following screen pairs:

- Screen 12 and 17/18
- Screen 19 and 5

Effect sizes were also calculated using the mean values of the cognitive load for each screen.

The screen pairs that had both statistical and visible to large effects (effect size greater than 0.40) are presented in Table 4.30.

Comparison of Mean Values				Effect Size
Column 1		Column 2		
Screen	Mean (SD)	Screen	Mean (SD)	
12	5.76 (1.59)	5	4.96 (1.31)	0.50
12	5.76 (1.59)	13	4.22 (1.62)	0.95
17/18	5.56 (1.45)	5	4.96 (1.31)	0.41
17/18	5.56 (1.45)	13	4.22 (1.62)	0.82
19	5.25 (1.63)	13	4.22 (1.62)	0.63
5	4.96 (1.31)	13	4.22 (1.62)	0.56

Table 4.30: Effect sizes for comparisons between screen in the animation version

There was a statistically significant difference in the cognitive load of screen 12 and 19 ($p = 0.024$), but the effect size was only ($d = 0.31$). In looking at the design of each of the animations however it is possible that this difference could still be visible. The design issues will be discussed in Chapter 5.

4.6.1.3 The static images & text version

One hundred and eighteen participants ($n=118$) used the static images & text version to learn the content. The self-reported cognitive load, on the scale of 1 – 9, was measured six times across the 23 screens of content. Table 4.31 summarises the content that was studied prior to asking the participants to indicate the amount of mental effort they felt they had to invest in order to understand the content. The actual screens are illustrated in Appendices G – K.

* Self report cognitive load (SRCL)

Screens	SRCL	Question phrased
Screen 5 of 23: Different parts of the Autonomic Nervous System (ANS). See Table 4.26.	SRCL1	Indicate how much mental effort you used to study the information on the screen you have just reviewed.
Screen 13 of 23: Content about the origins of the preganglionic nerves, presented with text, a static image and a hot spot on this image. When mouse was rolled over the hot spot the pop-up displayed a close up cross-section view of the spinal cord.	SRCL3	Indicate how much mental effort you used to study the information on the screen you have just reviewed.
Screen 14, 15 & 16 of 23: Screen 14 used text and an image to explain content. There were two mouse-overs for the image. Rolling the mouse-over the image displayed a pop-up that provided more information. One of the pop-ups had a link to another screen (screen 16). The screen zoomed in to the larger image and a very short animation, using text labels, explained the path of the preganglionic fibre. Screen 15 explained the content using text and two static images.	SRCL4	Indicate how much mental effort you used to study the information on the three screens covering the synapses, which you have just reviewed.
Screen 17, 18 & 19 of 23: Screen 17 introduced the topic of innervation using text only. Screen 18 was exactly the same as screen 17 in Version 1 (see Table 4.26). The second screen summarised these effects using text links and pop-ups (See Appendix J).	SRCL5	Indicate how much mental effort you used to study the information about the innervation of the target organs.
Screen 20 of 23: Two concepts were described, using text and a static image for each. User clicked on concept and viewed content underneath the link. The user could toggle between the two links and observe the subtle changes in the image.	SRCL6	Indicate how much mental effort you used to study the information on the screen you have just reviewed.
Screen 23 of 23: Animation of 1 min 15 secs. Same as Version 1. See Table 4.26.	SRCL7	Indicate how much mental effort you used to study the information on the screen you have just reviewed.

Table 4.31: Overview of content in the static image & text version on which self-report of cognitive load was based

With the exception of two measurements the rating was based on the content viewed in the screen that preceded the question. The third time the participants rated mental effort invested they had to consider three screens: 14 - 16. These screens, together with screen 13, replaced screen 12 of the animation version. The screens were presented in a linear order, but the participants could go backwards and forwards between these screens as often as needed before moving on. Several participants missed the link to screen 16. The fourth time the participants had to rate cognitive load they also had to consider three screens.

In all instances the participant could not go back to view the screen in order to make the decision about how much mental effort they had invested. They also only answered the question once. If they returned to the screen later, which was possible using the back button in the program, they were not asked to self-rate the mental effort invested a second time. The rationale for this design has already been explained in Section 4.6.1.1 on page 207 of this chapter.

The frequencies for the distribution of the cognitive load is presented in Table 4.32.

	Self-report Cognitive Load				
	Screen	n	Low	Medium	High
SRCL1	5 of 23	117	10	85	22
SRCL3	13 of 23	116	22	83	11
SRCL4	14/15/16 of 23	117	12	78	27
SRCL5	17/18/19 of 23	113	7	74	32
SRCL6	20 of 23	115	8	81	26
SRCL7	23 of 23	117	2	80	35

Table 4.32: Frequencies of cognitive load reported as low, medium and high for the static images & text version

Table 4.32 indicates that a higher proportion of participants rated screens 23 and 17 - 19 as requiring high mental effort. The animation on screen 23 had the highest percentage of high cognitive load ratings (30%), followed by screens 17 - 18 (28%) and screens 14 - 16 (23%) and 20 (23%) respectively. Only 1.7% of the participants rated screen 23 (the animation) in the low cognitive load category. Screen 13 had the highest percentage (19%) of low cognitive load ratings when compared to the other screens, and the lowest percentage (9.5%) of high cognitive load ratings.

The means and standard deviations for each instance of cognitive load measured in the static images & text version, ranked from high to low, is presented in Table 4.33.

SRCL	Screen	n	Mean (SD)	Minimum	Maximum
SRCL7	23 of 23	117	6.14 (1.20)	3	9
SRCL5	17/18/19 of 23	113	5.73 (1.43)	1	9
SRCL4	14/15/16 of 23	117	5.50 (1.48)	2	9
SRCL6	20 of 23	115	5.53 (1.39)	1	9
SRCL1	5 of 23	117	5.39 (1.40)	2	9
SRCL3	13 of 23	116	4.80 (1.51)	1	8

Table 4.33: Mean cognitive load ratings for individual measurements of cognitive load for the static images & text version

Using the same categories of low (1 - 3.9), medium (4 – 6.9) and high (7 – 9) cognitive load it would appear from the data in Table 4.33 that the mean ratings of all are within the range for medium cognitive load, with the value for SRCL7 (the animation on screen 23) approaching the higher end of the scale. The highest cognitive load was found for screen 23 ($M [\pm SD] = 6.14 [\pm 1.20]$), the animation, and the lowest load for screen 13 ($M [\pm SD] = 4.8 [\pm 0.5]$), a screen that used text and a static image with a hot spot. When the hot spot was activated another image was displayed on the screen.

The influence of the presentation formats on cognitive load in the static images & text version will be explored in more detail in the next section.

4.6.1.4 The influence of presentation formats on cognitive load in the static images & text version

The fourth sub-question asked about cognitive load was ‘To what extent do the presentation formats influence cognitive load?’ This section considers the static images & text version.

A series of Univariate analyses for pair-wise dependent samples was conducted. The results of the Sign test (M) indicated whether the difference in the measured cognitive load of the two screens being compared was significant. This test is appropriate for dependent samples where the data has been collected using a rating scale (Garson, 2006). The results from the two-tailed test, together with the associated p values, are reported. The Bonferroni correction (McMillan & Schumacher, 2001) was applied to these p -values.

The results are displayed in Table 4.34. The screen listed in the column marked 1 (one of the presentation formats) had a higher cognitive load than the screen listed in the column marked 2 (the second presentation format).

Screen-wise comparison of cognitive load			
Column 1	Column 2	Sign test statistic (M)	p
Screen	Screen		
23	17/18/19	-11.5	0.067
23	20	-20.5	0.001*
23	14/15/16	-15.5	0.002*
23	5	-27.5	0.001*
23	13	-37	0.001*
17/18/19	5	-10.5	0.257
17/18/19	20	6	> 0.95
17/18/19	14/15/16	-10	0.169
17/18/19	13	-22.5	0.001*
20	5	-5	> 0.95
20	14/15/16	2	> 0.95
20	13	-22	0.001*
14/15/16	5	-3	> 0.95
14/15/16	13	-24	0.001*
5	13	16	0.003*

* Alpha $p < 0.05$

Table 4.34: A comparison of the cognitive load for selected screen pairs in the static text & images version

From Table 4.34 we see that there are statistically significant differences for a number of the screen pair comparisons.

Effect sizes were also calculated using the mean values of the cognitive load for each screen. The screen pairs that had both statistical and visible to large effects (effect size close to or greater than 0.40) are presented in Table 4.35.

Comparison of Mean Values				Effect Size
Column 1		Column 2		
Screen	Mean (SD)	Screen	Mean (SD)	
23	6.14 (1.20)	20	5.53 (1.39)	0.43
23	6.14 (1.20)	14-16	5.50 (1.48)	0.43
23	6.14 (1.20)	5	5.39 (1.40)	0.53
23	6.14 (1.20)	13	4.80 (1.51)	0.88
17-19	5.73 (1.43)	13	4.80 (1.51)	0.61
20	5.53 (1.39)	13	4.80 (1.51)	0.48
14-16	5.50 (1.48)	13	4.80 (1.51)	0.46
5	5.39 (1.40)	13	4.80 (1.51)	0.39

* Alpha $p < 0.05$

Table 4.35: Effect sizes for comparisons between screen in the static images & text version

The largest effect size was obtained between screen 23 and 13 ($d = 0.88$). These results once again suggest that animation places a heavier load on the resources needed for cognitive processing than do text and static images & text.

There was no significant difference between the cognitive load of screens 17 – 19 and screens 14 – 16, where the amount of content was approximately the same across the two screen sets. This suggests that amount of content might also influence load. Another method of measuring cognitive load is to use the direct cognitive load measurement technique rather than self-report ratings, as the direct method measures the load at the time it occurs rather than after the event (Brünken, Plass & Leutner, 2003).

In both sets of screens the user was required to interact with images, text and buttons in order to study the content. In both sets of screens there was some form of animation, albeit very simple and very short. Smith (2007) measured the cognitive load of screens 14, 15, 16 and 18 using the direct measurement technique. This data was collected for her study using the same experimental as I did, with the same participants, at exactly the same time. Smith (2007) however did not compare the cognitive load of individual screens within the same version of the program or report on any effect sizes. I am using this data with permission (Appendix C) as part of the exploration into how cognitive load is affected by content and design.

The mean cognitive load values for each of these screens, obtained using the direct measurement technique, is presented in Table 4.36.

Screen	No of Obs	N	Mean (\pm SD)	Std Error	Minimum	Maximum
14	889	118	5.4563 (\pm 3.2911)	0.3029	0.2500	10
15	554	118	6.2432 (\pm 3.5856)	0.3300	0.3333	10
16	256	51	6.6049 (\pm 2.5330)	0.3547	0.6666	10
18	773	118	5.5594 (\pm 3.2760)	0.3015	0.6000	10

Table 4.36: Cognitive load using direct measurement technique

General Linear Modeling (GLM) was conducted using repeated measures analyses in a within subjects design. This test is appropriate for continuous data, like the data for the direct measurement of cognitive load. Data from the same sample makes the repeated measures analysis the most appropriate test to use. The analysis returned a statistically significant result, $F(2,100) = 14.74$, $p < 0.0001$. Further analysis was done to determine which screen-wise comparisons contributed to this finding. The results are presented in Table 4.37.

Screen	with screen	F-value	P value
14	15	1.76	0.1911
14	16	29.77	< 0.0001 *
15	16	11.60	0.0013 *

* Alpha $p < 0,05$

Table 4.37: Results of GLM Repeated Measures Analysis for the cognitive load comparisons

The results indicate that there was a statistically significant cognitive load between screen 14 and 16 and between screen 15 and 16. In each case screen 16 had the higher load. Screen 16 had the short non-narrated animation. These results once again provide evidence that animation poses a higher burden on the resources of the working memory than other strategies such as the use of text and static images.

The effect sizes for the pair-wise comparisons of the mean cognitive load of screens 14,15,16 and 18 in the static images & text version are presented in Table 4.38.

Comparison of Mean Values				
Column 1		Column 2		Effect Size
Screen	Mean (SD)	Screen	Mean (SD)	
14	5.4563 (3.2911)	15	6.2432 (3.5856)	0.22
14	5.4563 (3.2911)	16	6.6049 (2.533)	0.35
14	5.4563 (3.2911)	18	5.5594 (3.276)	0.03
15	6.2432 (3.5856)	16	6.6049 (2.533)	0.10
15	6.2432 (3.5856)	18	5.5594 (3.276)	0.19
16	6.6049 (2.533)	18	5.5594 (3.276)	0.32

Table 4.38: Effect sizes of comparison of mean cognitive load measured with the direct method

The effect sizes obtained indicate that most of the differences in the set of screens reported on in Table 4.38 are not practically significant. Only two effect sizes are larger than 0.30, but only by a very small margin. The comparison of screen 14 and 16 returned an effect size of 0.35, and that of screen 16 and 18 returned an effect size of 0.32. In spite of the fact that the comparison between screen 15 and 16 returned a statistically significant result ($p = 0.0013$) the effect size obtained was very small ($d = 0.10$) indicating that there is no practical significance in this results. Although these are still small effect sizes it must be pointed out that in both comparisons the screen with the animation (screen 16) had the higher cognitive load.

4.6.1.5 Comparing the animation and static images & text versions

The preceding sections (Section 4.6.1.1 through to Section 4.6.1.4) looked at the cognitive load for each version separately. The study, however, was also designed to explore cognitive load using two different formats and so the analysis of cognitive load must compare the findings across the two versions.

The three hypotheses tested were:

Null Hypothesis 2a

There will be no difference between the animation and static images & text version in the cognitive load of the program as a whole.

Alternate Hypothesis 2a

In considering the program as a whole the animation version will have a higher cognitive load than the static images & text version.

Null Hypothesis 2b

There will be no difference in the cognitive load of the screen using animation and the alternative version using static images & text.

Alternate Hypothesis 2b

The screen using animation will have a higher cognitive load than the alternative version using static images & text.

Null Hypothesis 2c

At screen level there will be no difference in the cognitive load across the versions where the content and presentation format are the same.

Alternate Hypothesis 2c

At screen level, the cognitive load will be the same in each version where the content and presentation format are the same.

This comparison used the total cognitive load values for each participant when calculating the cognitive load of the program as a whole, and the method of calculating cognitive load is provided below.

Self-report of cognitive load (SRCLV) for each participant:

$$SRCLV = \text{AVERAGE}(SR_1, SR_2 \dots SR_n)^2$$

Self report of cognitive load for each version:

$$SR1^3 = \text{AVERAGE}(SRCLV_1, SRCLV_2, \dots) \text{ and } SR2 = \text{AVERAGE}(SRCLV_1, SRCLV_2, \dots)$$

The total cognitive load for each version was calculated as described.

² SR_n refers to the number of screens where cognitive load was measured

³ SR1: Cognitive load of Version 1, SR2: Cognitive load of Version 2

The Mean (Confidence Limit (CL)) and Standard deviation for the cognitive load of the two versions is presented in Table 4.39.

Version	n	Mean (\pm CL)	SD	Min	Max
Animation	120	5.1501 (\pm 0.1994)	1.093	1.8	8.6
Static images and text	118	5.5147 (\pm 0.2337)	1.034	2	8.3

Table 4.39: Mean cognitive load per version using the subjective rating technique

At an alpha of 0.05, a t test to determine if these two means were statistically significantly different returned a t-statistic of -2.63, $p = 0.0091$. The results indicate that when considering the cognitive load of the program as a whole, the static images & text version had a statistically significant higher cognitive load than the animation version. The effect size was $d = 0.33$, which is a visible but small effect. Conclusions about practical significance must be interpreted cautiously.

*The results from the analysis of the total cognitive load of each version **do not support** hypothesis 2a that proposed that as a whole the animation version would have a higher cognitive load than the static images & text version.*

This was an unexpected finding, following the results presented in the previous three sections.

The next step in the analysis was to consider a series of screen-wise comparisons across the two versions. These sets of screens isolate the instructional strategies and presentation formats: animation versus static images & text, and the use of pop-ups.

The Means, Confidence limits, Standard deviations and Standard errors for the screen-wise comparisons of cognitive load between the two versions is presented in Table 4.40.

Strategy	Vers	Screen	n	Subjective rating of cognitive load							
				Mean	Confidence Limit	SD	Std Error	Otained <i>t</i>	DF	<i>p</i> value	Effect size
Animation versus static image and text	Animation	12	114	5.763	±0.295	1.592	0.149	3.30	230	0.0011	0.40
	Static	13 - 16	118	5.131	±0.240	1.318	0.121				
Same content, same presentation format	Animation	19	112	5.250	±0.306	1.636	0.155	- 4.73	227	< 0.001	0.55
	Static	23	117	6.145	±0.220	1.205	0.111				
Same content, presentation format different	Animation	5	117	4.966	±0.240	1.313	0.121	-2.40	232	0.017	0.33
	Static	5	117	5.393	±0.258	1.408	0.130				
Same content, presentation format different	Animation	13	114	4.228	±0.301	1.624	0.152	-6.56	227	< 0.001	0.81
	Static	20	115	5.539	±0.257	1.391	0.130				
Same content, presentation format different	Animation	17-18	113	5.566	±0.271	1.457	0.137	-0.87	232	0.383	0.12
	Static	17-19	113	5.734	±0.268	1.440	0.135				

Table 4.40: Screen-wise comparison of cognitive load for selected presentation formats

At an alpha of 0.05, a t test to determine if these means were significantly different, returned significant differences for four of the five comparisons. The relevant t-statistics, *p* values and effect sizes are presented in Table 4.40.

The first comparison was between screen 12, the long animation in the animation version, and screens 13-16, which presented the same content using static images & text. The cognitive load for the animation version was significantly higher than the cognitive load for the static images & text version, $t = 3.30$, $df = 230$, $p = 0.0011$. The effect size of 0.40 indicates a visible effect.

*The results from the analysis therefore **supports** hypothesis 2b which stated that the screen using animation will have a higher cognitive load than the alternative screens using static images & text.*

The next group of comparisons was between screens that had the same content and the same presentation format. The first comparison was between screen 19 from the animation version and screen 23 from the static images & text version. The design of these screens has been described in Sections 4.5.1.1 to 4.5.1.4.

Hypothesis 2c stated that at screen level there will be no difference in the cognitive load across the versions where the content and presentation are the same. The difference for both groups of screen comparisons was significant, as can be seen in Table 4.40, with the difference between screen 19 and screen 23 highly significant ($p < 0.001$). The effect size was 0.55, also indicating a visible effect. The comparison of screen 5 returned a *p* value of 0.017 and an effect size of 0.33. This small effect size suggests that the difference in cognitive load does not have much practical significance. It is unlikely that the finding with regard to Screen 5 can be explained by participant fatigue, as this was the fifth screen in the program for both versions. The reason for this difference will be explored from a cognitive style perspective later in this chapter.

*The results from the analysis therefore **do not support** hypothesis 2c, which proposed that at screen level there will be no difference in the cognitive load across the versions where the content and presentation format are the same.*

Another reason for the finding might relate to the fact that subjective ratings are always exactly that: subjective. A look at the direct measurement of cognitive load might return different findings.

The mean cognitive load for screens 5, 19 and 23, obtained using the direct measurement technique, and measured by Smith (2007) is presented in Table 4.41.

Screen	Version	N	Mean	SD	Effect size
Animation – Screen 5	1	119	6.1840	3.4439	0.20
Static – Screen 5	2	117	5.4903	3.1933	
Animation- Screen 19	1	119	9.3825	0.7893	0.29
Static – Screen 23	2	117	9.1551	1.3003	

Table 4.41: Cognitive load of selected screens using the direct method of measurement

Both comparisons returned small effect sizes. The effect size of 0.20 for the comparison of the means for screen 5 can be considered not to have any practical significance, and suggests that hypothesis 2c could not be rejected. The effect size for the comparison of the means for screens 19 and 23 is slightly larger but still small. The fatigue effect might still be appropriate.

4.6.1.6 In summary

In this section the cognitive load of each version of the program was explored. A series of screen-wise comparisons of cognitive load, both within the same version of the program and across the two different versions of the program was conducted. The method of measurement was the subjective rating, using the 9-point scale of Paas. Significance testing was conducted using t tests. Effect sizes were calculated for all the comparisons. Where unexpected findings were obtained the direct measurement of the cognitive load for the particular screen or screens was considered, using the data collected by Smith (2007) during the same experiment.

When the program was considered as a whole the total cognitive load of the animation version was $M (\pm SD) = 5.1501 (\pm 0.1994)$ and that of the static images & text version $M (\pm SD) = 5.5147 (\pm 0.2337)$. This difference was statistically significant ($t = -2.63, p = 0.0091$) with an effect size of 0.33, which is a visible effect.

The cognitive load was divided into three categories: low, medium and high. For both versions of the program the total cognitive load of each measurement was always within the medium range, which was set as greater than 4 and smaller than or equal to 6.9. **A consistent finding was that the screens with animation had the highest cognitive load compared to screens with static images & text.**

Comparison of the cognitive load of screens within the animation version indicated that the screen with the animation had a statistically significantly higher load than the three screens with which it was

compared. The effect size for two of the four screen-wise comparisons was larger than 0.50, indicating visible effects. Comparison of the cognitive load of screens within the static images & text version indicated that the screen with the animation had a statistically significantly higher load than the four screens with which it was compared. The effect size for four of the five screen-wise comparisons was larger than 0.45, indicating visible effects.

The next section considers the relationship between cognitive load and cognitive style. This relationship is examined without considering the performance of the learner in the posttest.

4.6.2 The relationship between cognitive load and cognitive style

This section looks at the relationship between cognitive load and cognitive style. Five hypotheses were tested:

Null Hypothesis 3a

There will be no difference between the Wholistic and Analytic learner for the cognitive load of the animation version.

Alternate Hypothesis 3a

The cognitive load of the animation version will be lower for the Wholistic learner than for the Analytic learner.

Null Hypothesis 3b

There will be no difference between the Wholistic and Analytic learner for the cognitive load of the static images & text version, used as an alternative for the animation version.

Alternate Hypothesis 3b

The cognitive load of the static images & text version, used as an alternative for the animation version, will be lower for the Analytic learner than for the Wholistic learner.

Null Hypothesis 3c

There will be no difference between the Verbaliser and Imager learner for the cognitive load of the animation version.

Alternate Hypothesis 3c

The cognitive load of the animation version will be lower for the learner with Imager style than for the learner with Verbaliser style.

Null Hypothesis 3d

The amount of time spent on the program will make no difference to the rating of cognitive load by the Analytic learner.

Alternate Hypothesis 3d

Analytic learners who spent inadequate time on the program will rate the cognitive load lower than other Analytic learners.

Null Hypothesis 3e

The amount of time spent on the program will make no difference to the rating of cognitive load by the Verbaliser and Imager learner.

Alternate Hypothesis 3e

Verbalisers and Imagers who spent inadequate time on the program will rate the cognitive load more highly.

The analyses conducted and reported in this section were done primarily using t tests to establish statistical significance and effect sizes for the comparison of means to establish practical significance.

4.6.2.1 Cognitive load and the WA style

The cognitive load of the program for the participants who were found to have Wholistic and Analytic styles is presented in Table 4.42.

	Wholistic style		Analytic style		Effect size
	n	Mean (SD)	n	Mean (SD)	<i>d</i>
Cognitive load	73	5.2356 (1.1035)	162	5.3737 (1.1035)	0.13

Table 4.42: Cognitive load for WA style dimension

From Table 4.42 we can see that the difference in these mean values is small. The effect size for the comparison of mean values is also small, $d = 0.13$, indicating that there is no practical significance between the two groups of participants for the WA dimension of style when version was left out of the analysis.

Further analysis was done to explore the relationship between load and WA style. The cognitive load obtained for each of the WA styles was compared for each version. The results of this comparison are presented in Table 4.43.

Version	Cognitive load for WA style dimension				
	Wholistic style		Analytic style		Effect size
	N	Mean (SE)	n	Mean (SE)	<i>d</i>
Animation	37	5.0611 (1.0743)	83	5.1892 (1.1063)	0.12
Static images & text	36	5.4102 (1.1197)	79	5.5650 (1.0147)	0.14

Table 4.43: Cognitive load and the WA style dimension by version

Looking at the means in Table 4.43 we see that the differences in the cognitive load between the Wholistic and Analytic style was small for both the animation and the static images & text version.

Hypothesis 3a predicts that the cognitive load of the animation version will be lower for the Wholistic learner than for the Analytic learner. In Table 4.43 we see that the Wholistic learner did report a lower

cognitive load than the Analytic learner for the animation version. A t test did not return a statistically significant difference between the two groups, $t = -0.58$, $df = 116$, $p = 0.5602$. The effect size is very small, $d = 0.12$ and does not provide support for practical significance.

*The hypothesis that the cognitive load of the animation version will be lower for the Wholistic learner than for the Analytic learner is therefore **not supported**.*

Hypothesis 3b predicts that the cognitive load of the static images & text version, used as an alternative for the animation version, will be lower for the Analytic learner than for the Wholistic learner. In Table 4.43 we see that the Analytic learner had a higher cognitive load than the Wholistic learner for the static images & text version. The difference between the two groups was not statistically significant, $t = -0.73$, $df = 113$, $p = 0.44643$. The effect size is also very small, $d = 0.14$ and does not provide support for practical significance.

*The hypothesis that the cognitive load of the static images & text version, used as an alternative for the animation version, will be lower for the Analytic learner than for the Wholistic learner is therefore **not supported**.*

I then looked at this relationship between cognitive load and WA style from another perspective and compared the cognitive load for the Wholistic learner by version and then the cognitive load for the Analytic learner by version. The results are presented in Table 4.44.

Version	Cognitive load for WA style dimension				
	Animation		Static images & text		Effect size
	N	Mean (SE)	n	Mean (SE)	<i>d</i>
Wholistic	37	5.0611 (1.0743)	36	5.4102 (1.1197)	0.31
Analytic	83	5.1892 (1.1063)	79	5.5650 (1.0147)	0.34

Table 4.44: Cognitive load of the WA styles for each version

Looking at Table 4.44 we see that the Wholistic learner experienced a higher cognitive load in the static images & text version. A t test did not return a statistically significant difference between the two versions for the Wholistic learner, $t = -1.35$, $df = 70$, $p = 0.184$. The effect size, $d = 0.31$, was in the small-to-medium range.

In Table 4.44 we see that the Analytic learner also experienced a higher cognitive load in the static images & text version. A t test **did** return a statistically significant difference between the two versions for the Analytic learner, $t = -2.24$, $df = 159$, $p = 0.0262$. The effect size, $d = 0.34$, was in the small-to-medium range.

Cognitive load was also divided into three categories: low, medium and high. The mean values of the cognitive load for each of these categories for the WA style dimension for is presented in Table 4.45.

Cognitive load and the WA style dimension					
	Wholistic		Analytic		Effect size
	N	M (SD)	N	M (SD)	
Low (≤ 3.9)	8	3.0479 (0.6791)	8	3.0125 (0.7568)	0.05
Medium (>3.9 and ≤ 6.9)	60	5.3781 (0.6266)	142	5.3212 (0.7065)	0.08
High (> 6.9)	4	7.4750 (0.4113)	11	7.7667 (0.5254)	0.56

Table 4.45: Cognitive load levels and the WA style dimension

The effect sizes were calculated for the comparison between the Wholistic and Analytic learner at each level. The data in Table 4.45 indicates that the effect sizes for the low and medium cognitive load groups for this style dimension were very small, 0.05 and 0.08 respectively, and therefore do not have practical significance. The effect size for the high cognitive load group (where the cognitive load was > 6.9) was 0.56, which is a medium effect and can be regarded as a visible effect, in spite of the fact that a t test of the means in the high cognitive load group did not return a statistically significant result, $t = -1.00$, $df = 13$, $p = 0.3373$. The Analytic learner reported a higher cognitive load in this group. Since I could find no studies with which to compare these results, and version was not considered in this analysis the recommendation is that this effect be investigated further.

There is the possibility that time would also influence the relationship between cognitive load, cognitive style and performance. Those Analytic learners who spent inadequate time on the program (for reasons that I did not examine) might have not been able to process as elaborately as they would have liked, and would therefore rate the cognitive load as being lower.

Inadequate time was defined on page 181 of this thesis. It was different for each version of the program due to the fact that the two versions were different in length. The same category of time was used in the analyses to test hypothesis 3d.

The cognitive load ratings for the Analytic learner, for each category of time spent is displayed in Table 4.46.

Cognitive load rating for the Analytic Learner						
	N	Inadequate	N	Adequate	N	More than Adequate
Animation	11	5.0909 (1.2308)	53	5.2956 (1.0581)	18	4.9361 (1.1859)
Static images & text	23	5.0725 (0.7244)	49	5.7442 (0.9968)	7	5.9286 (1.4715)

Table 4.46: Cognitive load for the Analytic learner grouped by time spent on program

In Table 4.46 we see that the Analytic learner who spent 'Inadequate' time on the program had a lower cognitive load rating than the Analytic learner who spent 'Adequate' time on the program, for both the animation and the static images & text version. An interesting finding is that the Analytic learner who spent 'More than adequate time' studying the animation version reported the lowest cognitive load for all the 'time groups' in the animation version, while for the static images & text version the cognitive load continues to increase as the Analytic learner spent more time on the program. A series of t tests for independent samples were performed to determine whether the cognitive load ratings for the different time groups were statistically significant.

The results of these t tests, and the associated effect sizes are displayed in Table 4.47.

In Table 4.47 we see that none of the comparisons of cognitive load were statistically significant for the animation version. In looking at the static images & text version we see a different picture. The difference in the cognitive load for Analytic learners who spent 'Inadequate' and 'Adequate time' was statistically significant, $t = -2.89$, $df = 70$. $p = 0.0051$, as was the difference for Analytic learners who spent 'Inadequate' and 'More than Adequate time', $t = -2.12$, $df = 28$. $p = 0.0431$. Medium-to-large effect sizes were returned, indicating practical significance.

The hypothesis that the Analytic learners who spent inadequate time on the program will rate the cognitive load lower than other Analytic learners was supported for the static images & text version, but not for the animation version.

Animation Version									
Time	n	Cognitive load	Time	n	Cognitive load	df	t value	p value	Effect size
Inadequate	11	5.0909 (1.2308)	Adequate	53	5.2956 (1.0581)	62	-0.57	0.5721	0.17
Inadequate	11	5.0909 (1.2308)	More than Adequate	18	4.9361 (1.1859)	27	0.34	0.7392	0.13
Adequate	53	5.2956 (1.0581)	More than Adequate	18	4.9361 (1.1859)	69	1.21	0.2312	0.30
Static images & text version									
Time	n	Cognitive load	Time	n	Cognitive load	df	t value	p value	Effect size
Inadequate	23	5.0725 (0.7244)	Adequate	49	5.7442 (0.9968)	70	-2.89	0.0051*	0.67
Inadequate	23	5.0725 (0.7244)	More than Adequate	7	5.9286 (1.4715)	28	-2.12	0.0431*	0.58
Adequate	49	5.7442 (0.9968)	More than Adequate	7	5.9286 (1.4715)	54	-0.43	0.6686	0.13

* Alpha at $p < 0.05$

Table 4.47: T-test for significance for the comparisons of cognitive load per time group for the Analytic learner

4.6.2.2 Cognitive load and VI style

The cognitive load of the program for the participants who were found to have Verbaliser and Imager styles, using Ridings CSA, is presented in Table 4.48.

	Verbaliser style		Imager style		Effect size
	n	Mean (SD)	n	Mean (SD)	<i>d</i>
Cognitive load	94	5.3070 (0.9073)	141	5.3472 (1.1914)	0.03

Table 4.48: Cognitive load and VI style

From Table 4.48 we see that the difference in these mean values is very small. The effect size for the comparison of these mean values is also very small, $d = 0.03$, indicating that there is no practical significance between the two groups of participants for the VI dimension of style, when version is not considered.

Further analysis was done to explore the relationship between load and VI style. The cognitive load obtained for each of the VI styles was compared for each version. The results of this comparison are presented in Table 4.49.

Version	Cognitive load for VI style dimension				
	Verbaliser style		Imager style		Effect size
	n	Mean (SD)	n	Mean (SD)	
Animation	50	5.1847 (0.9663)	70	5.1248 (1.1850)	0.05
Static images & text	44	5.4458 (0.8241)	71	5.5603 (1.1663)	0.10

Table 4.49: Cognitive load and the VI style dimension by version

Table 4.49 compares the cognitive loads for each VI style group by version. The results of the analysis show that for each version, the differences in the cognitive load means for the two VI style groups being compared were small. Looking at the means in Table 4.47 we see that the difference in the cognitive load mean of the VI styles in the animation version was very small, with a very small effect size, $d = 0.05$. The difference in the cognitive load mean of the VI styles for the static images & text version was slightly larger, but the effect size was also very small, $d = 0.10$. This indicates that there is no practical significance in the differences between these two groups, for both versions.

Hypothesis 3c predicted that the cognitive load of the animation version will be lower for the learner with Imager style than for the learner with Verbaliser style. In Table 4.49 we see that the Imager learner did report a lower cognitive load than the Verbaliser learner for the animation version, but a t test did not return a statistically significant difference between the two groups, $t = 0.29$, $df = 116$, $p = 0.7701$.

*The hypothesis that the cognitive load of the animation version will be lower for the learner with Imager style than for the learner with Verbaliser style is therefore **not supported**.*

I then looked at this relationship between cognitive load and VI style from another perspective and compared the cognitive load for the Verbaliser learner by version and then the cognitive load for the Imager learner by version. The results are presented in Table 4.50.

Version	Cognitive load for VI style dimension				
	Animation		Static images & text		Effect size
	n	Mean (SE)	n	Mean (SE)	<i>d</i>
Verbaliser	50	5.1847 (0.9663)	44	5.4458 (0.8241)	0.27
Imager	70	5.1248 (1.1850)	71	5.5603 (1.1663)	0.37

Table 4.50: Cognitive load of the VI styles for each version

Looking at Table 4.50 we see that both the Verbalisers and Imagers experienced a higher cognitive load in the static images & text version. A t test did not return a statistically significant difference between for the Verbalisers, $t = -1.49$, $df = 92$, $p = 0.1649$. The effect size, $d = 0.27$, was small and suggest that this difference does not have practical significance. The t test did return a statistically significant difference for the Imagers, $t = -2.18$, $df = 137$, $p = 0.0307$. The effect size, $d = 0.37$, in the small to medium range.

The mean values of cognitive load for the three categories of load are presented in Table 4.51 for the VI style dimension.

	Cognitive load and the VI style dimension				
	Verbaliser		Imager		Effect size
	n	M(SD)	n	M (SD)	
Low (≤ 3.9)	4	3.0000 (0.5416)	12	3.0403 (0.7602)	0.05
Medium (>3.9 & ≤ 6.9)	87	5.3310 (0.6589)	115	5.3435 (0.7029)	0.02
High (>6.9)	3	7.6833 (0.1607)	12	7.6903 (0.5613)	0.01

Table 4.51: Cognitive load levels and the VI style dimension

The effect sizes were calculated for the comparison between the Verbaliser and Imager learner at each level. The data in Table 4.51 indicates that the effect sizes for all three groups were very small, and therefore do not have practical significance. No further analysis was done.

It was argued in Chapter 3 that the format of the program should suit both the Verbaliser and Imager. VI style should therefore not influence learning performance, unless the cognitive load is high and inadequate time was spent in the program. Hypothesis 3e predicts that Verbalisers and Imagers who spent inadequate time on the program will rate the cognitive load more highly.

The cognitive load ratings for the Verbaliser and Imager learners, for each category of time spent in the animation version is displayed in Table 4.52.

	Cognitive load rating for the VI style dimension in the Animation version					
	N	Inadequate	N	Adequate	N	More than Adequate
Verbaliser	11	4.4364 (1.2769)	29	5.5322 (0.6888)	10	5.000 (0.8273)
Imager	10	5.22 (0.9211)	42	5.1151 (1.1179)	16	5.0906 (1.52295)

Table 4.52: Cognitive load for the VI style grouped by time spent on the animation version

Verbalisers, who are thought to share some of the characteristics common to Analytic learners, have a pattern of load for the animation version that is similar to the pattern of the Analytic learner for the animation version (see Table 4.46): low where 'Inadequate' time is spent, higher for the level 'Adequate' time and lower again for the 'More than Adequate' time group. The only difference is that the lowest rating is found in the 'Inadequate' time group. The pattern that emerges for the Imager style in the animation version is quite different. The highest cognitive load was reported by the

Imagers who spent 'Inadequate' time on the animation version and the cognitive load decreased as more time was spent on the program.

A series of t tests for independent samples was performed to determine whether the cognitive load ratings for the different time groups were statistically significant. The results of these t tests, and the associated effect sizes are displayed in Table 4.53.

Verbalisers: Animation Version									
Time	n	Cognitive load	Time	n	Cognitive load	df	t value	p value	Effect size
Inadequate	11	4.4364 (1.2769)	Adequate	29	5.5322 (0.6888)	38	-3.51	0.0012	0.86
Inadequate	11	4.4364 (1.2769)	More than Adequate	10	5.000 (0.8273)	19	-1.19	0.2501	0.44
Adequate	29	5.5322 (0.6888)	More than Adequate	10	5.000 (0.8273)	37	2.00	0.0527	0.42
Imagers: Animation Version									
Time	n	Cognitive load	Time	n	Cognitive load	df	t value	p value	Effect size
Inadequate	10	5.22 (0.9211)	Adequate	42	5.1151 (1.1179)	50	0.27	0.7846	0.09
Inadequate	10	5.22 (0.9211)	More than Adequate	16	5.0906 (1.52295)	24	0.24	0.8120	0.08
Adequate	42	5.1151 (1.1179)	More than Adequate	16	5.0906 (1.52295)	56	0.07	0.9468	0.02

Table 4.53: T-test for significance for the comparisons of cognitive load per time group for the Verbaliser learner

In Table 4.53, looking at the results for the Verbalisers, we see that two of the three comparisons of cognitive load were statistically significant for the animation version, and medium to large effect sizes were returned for all three comparisons ($d = 0.85$, $d = 0.44$ and $d = 0.42$ respectively), even though the comparison between cognitive load for the 'Adequate' and 'More than Adequate' time group was not statistically significant (although it closely approached statistical significance). These results have both statistical and one has practical significance.

In looking at the Imagers we see a different picture. None of the comparisons were statistically significant and effect sizes were all very small.

Hypothesis 3e is therefore not supported in the animation version of the program, although the findings were in the projected direction for the Imager who spent inadequate time on the program.

The cognitive load ratings for the Verbaliser and Imager learners, for each category of time spent in the static images & text version is displayed in Table 4.54.

Cognitive load rating for the VI style dimension in the Static images & text version						
	n	Inadequate	n	Adequate	n	More than Adequate
Verbaliser	17	5.3235 (0.64140)	24	5.5326 (0.9756)	3	5.4444 (0.3849)
Imager	28	5.1714 (1.1367)	36	5.7681 (1.0715)	7	6.0476 (1.2009)

Table 4.54: Cognitive load for the VI style grouped by time spent on the static images & text version

The lowest cognitive load was reported for the Verbalisers who spent 'Inadequate' time on the program. This pattern is similar to that of the Verbaliser who used the animation version (see Table 4.52). The pattern that emerges for the Imager style in the static images & text version is the **opposite** of the pattern found for the Imagers who used the animation version. In the static images & text version the **lowest** cognitive load was reported by those who spent 'Inadequate' time on the program and the cognitive load **increased** as more time was spent on the program (see Table 4.52).

A series of t tests for independent samples was performed to determine whether the cognitive load ratings for the different time groups were statistically significant. The results of these t tests, and the associated effect sizes are displayed in Table 4.55.

Verbalisers: Static images & text version									
Time	n	Cognitive load	Time	N	Cognitive load	df	t value	p value	Effect size
Inadequate	17	5.3235 (0.6414)	Adequate	24	5.5326 (0.9756)	39	-0.77	0.4448	0.21
Inadequate	17	5.3235 (0.6414)	More than Adequate	3	5.4444 (0.3849)	18	-0.31	0.7584	0.19
Adequate	24	5.5326 (0.9756)	More than Adequate	3	5.4444 (0.3849)	25	0.15	0.8797	0.23
Imagers: Static images & text version									
Time	n	Cognitive load	Time	n	Cognitive load	df	t value	p value	Effect size
Inadequate	28	5.1714 (1.1367)	Adequate	36	5.7681 (1.0715)	62	-2.15	0.0353	0.52
Inadequate	28	5.1714 (1.1367)	More than Adequate	7	6.0476 (1.2009)	33	-1.73	0.0936	0.73
Adequate	36	5.7681 (1.0715)	More than Adequate	7	6.0476 (1.2009)	41	-0.60	0.5545	0.23

Table 4.55: T-test for significance for the comparisons of cognitive load per time group for the Imager learner

In Table 4.55, looking at the results for the Verbalisers, we see that none of the comparisons were statistically significant and effect sizes were all very small. For the Imager style the results are different. The comparison between the 'Inadequate' time and the 'Adequate' group was statistically significant, $t = -2.15$, $df = 62$, $p = 0.0353$ and the effect size ($d = 0.52$) was in the medium-to-large range indicating practical significance. The comparison between the 'Inadequate' time and the 'More than Adequate' group was not statistically significant, $t = -1.73$, $df = 33$, $p = 0.0936$, but the effect size ($d = 0.73$) was large indicating practical significance.

*Hypothesis 3e is therefore **not** supported in the static images & text version of the program.*

4.6.2.3 In summary: cognitive load and cognitive style

The analysis considered each style group separately as these styles are independent of each other. Five hypotheses were tested.

A comparison of the cognitive load obtained for each of the WA styles for each version revealed that there was no statistical or practical significance in the difference in the rating of the two style groups in the animation version. It was therefore not possible to reject the null hypothesis that there will be no difference between the Wholistic and Analytic learner for the cognitive load of the animation version. In considering the static images & text version it was also not possible to reject the null hypothesis which stated that there will be no difference between the Wholistic and Analytic learner for the cognitive load of the static images & text version, used as an alternative for the animation version. The expectation was that the cognitive load would be lower for the Analytic learner. In fact the Analytic learner had a higher cognitive load than the Wholistic learner for the static images & text version, but the difference between the two groups was not statistically or practically significant.

The analysis also compared the cognitive load for the Wholistic and Analytic learner by version. The Wholistic learner experienced a higher cognitive load in the static images & text version. The analysis did not return a statistically significant difference between the two groups, but the effect size ($d = 0.31$) suggested that this findings could have practical significance. The Analytic learner also experienced a higher cognitive load in the static images & text version and there was a statistically significant difference between the two groups. The effect size was in the small to medium range which suggests that this finding could be visible.

The cognitive load was divided into three categories: low, medium and high. The mean values of the cognitive load for each of these categories was compared for the two styles in the WA style dimension. The effect sizes for the comparison of cognitive load between the Wholistic and Analytic styles were very small, 0.05 and 0.08 for the low and medium cognitive load groups respectively. The

effect size for the comparison of the high cognitive load group (where the cognitive load was > 6.9) was 0.56, which is a medium effect and can be regarded to be visible. This analysis was followed up with a t test to determine statistical significance. The analysis did not return a statistically significant result.

Similar analyses were conducted for the VI style dimension. A comparison of the cognitive load obtained for each of the VI styles for each version revealed that there was no statistical or practical significance in the difference in the rating of the two style groups in the animation version. It was therefore not possible to reject that null hypothesis that there will be no difference between the Verbaliser and Imager learner for the cognitive load of the animation version.

The analysis also compared the cognitive load for the Verbaliser and Imager learner by version. Both the Verbalisers and Imagers experienced a higher cognitive load in the static images & text version. This result was not statistically or practically significant for the Verbalisers. The result was statistically significant for the Imagers and the effect size, which was in the small to medium range, suggests it might be visible.

The cognitive load was divided into three categories: low, medium and high. The mean values of the cognitive load for each of these categories was compared for the two styles in the VI style dimension. The effect sizes for the comparison of cognitive load between the Verbaliser and Imager styles were very small for all three the cognitive load groups respectively. No further analysis was conducted to determine statistical significance.

The final analysis sought to understand whether the time spent on the program made a difference to the cognitive load for the learners with different style. The analysis was considered by version of the program. Only the Analytic style was considered from the WA dimension. An interesting finding was that the Analytic learner who spent 'More than Adequate' time studying the animation version reported the lowest cognitive load for all the 'time groups' in the animation version, while for the static images & text version the cognitive load continued to increase as the Analytic learner spent more time on the program. Following significance testing it was concluded that the hypothesis that the Analytic learners who spent inadequate time on the program will rate the cognitive load lower than other Analytic learners was supported for the static images & text version, but not for the animation version.

Both the Verbaliser and Imager styles were considered when analysing the results for the VI style. In the animation version a pattern similar to the Analytic learner emerged for the Verbaliser learner and two of the three comparisons of cognitive load were statistically significant with effect sizes that indicate practical significance for all three comparisons. In looking at the Imagers we see a very interesting pattern. In the animation version the highest cognitive load was reported by the Imagers who spent 'Inadequate' time on the animation version and the cognitive load decreased as more time was spent on the program, while in the static images and text version the lowest cognitive load was

reported by the Imagers who spent 'Inadequate' time on the animation version and the cognitive load increased as more time was spent on the program. The only statistically significant finding was between the cognitive load for the 'Inadequate' and 'Adequate' time group.

The next section will consider the influence of the both the cognitive style and demographic variables on the subjective rating of cognitive load.

4.6.3 Subjective rating of cognitive load and other variables

In the preceding sections the influence of time on the subjective rating of cognitive load for the different style groups was explored. This section considers the influence of several other variables on the subjective rating of cognitive load. A stepwise regression, followed by general linear modeling was used to conduct the analyses. The model for this stepwise regression analysis used cognitive load as the dependent variable.

The following variables, and all their interactions, were included in the regression analysis:

- Version of the program
- Gender
- Culture
- Had they studied the topic previously or not
- Rating of their knowledge
- English as home language
- WA style
- VI style
- Total time spent on the program

The result of the regression was $F(3, 226) = 8.33$, $p < 0.0001$, $R\text{-square} = 0.0995$ and $C(p) = -3.5055$. The R-square is in the medium effect size range and is a visible effect. The variables retained in the stepwise regression were home language and the interactions between version and total time spent on the program and between culture and the rating of knowledge about the topic.

The stepwise regression equation is presented in Table 4.56.

Variable	Parameter Estimate	Standard Error	Type III Sum of squares	F-value	p-value
Intercept	5.0932	0.3634	212.0975	196.46	< 0.0001
Home Language	212.0975	0.1573	3.3433	3.10	0.0798
Version x total time spent on program	0.0060	0.0022	8.0969	7.50	0.0067
Culture x rating of knowledge	-0.2047	0.0660	10.3901	9.62	0.0022

Table 4.56: Stepwise regression equation for self-report rating of cognitive load

From the data it can be determined that there was no statistically significant main effects, although the main effect for home language approached significance, $F(3, 226) = 3.10$, $p = 0.0798$. There were two statistically significant interactions: between version and the total time spent on the program, $F(3, 226) = 7.30$, $p = 0.0067$ and between culture and the rating of knowledge about the topic, $F(3, 226) = 9.62$, $p = 0.0022$.

A confirmatory General linear model (GLM) was then run, based on the results from the stepwise regression. The GLM presents the standard error and p-difference tables, the results of which are not provided in the stepwise regression. The GLM allows comparison of the class variable effects (home language, version, culture and rating of knowledge about the topic) using pair wise least squares means comparisons.

The model used for the GLM was: Self-report rating of cognitive load = version, culture, rating of knowledge about the topic, home language, the interaction between culture and rating of knowledge about the topic, total time spent on the program and the interaction between version and the total time spent on the program.

The GLM analysis returned a statistically significant finding, $F(7, 224) = 3.67$, $p = 0.0009$, $R^2 = 0.1029$.

The rating of knowledge about the topic accounted for the significant result, $F(1, 224) = 3.98$, $p = 0.0474$. Home language, $F(1, 224) = 2.79$, $p = 0.0964$, and the total time spent on the lesson, $F(1, 224) = 2.82$, $p = 0.0943$ approached significance. Further post hoc comparison, using Fischer's Least Squares Means was conducted. The effect size was $d = 0.47$, indicating this is a visible effect.

The Least Squares Means for the cognitive load ratings for the two groups of rating of knowledge about the topic are displayed in Table 4.57.

	N	Cognitive load LS Mean	Standard error	F statistic	p value	Effect size
Know nothing about topic	19	5.7828	0.2442	3.98	0.0474	0.47
Have basic & intermediate knowledge about topic	214	5.2748	0.0858			

Table 4.57: GLM analysis results for cognitive load and rating of knowledge about the topic

The Least Squares Means for the cognitive load ratings for the two different home language groups are displayed in Table 4.58. The effect size is $d = 0.27$, a small effect.

	N	Cognitive load LS Mean	Standard error	F statistic	p value	Effect Size
English is first home language	62	5.3947	0.1688	2.82	0.0943	0.27
English not first home language	171	5.6629	0.1377			

Table 4.58: GLM analysis results for cognitive load and home language

Although this result was not statistically significant the results support the personal observation referred to on page 177 of this thesis that learners whose first language is not English find Physiology more difficult.

4.6.4 Cognitive load and time spent on the program

In Section 4.5.2 an analysis was conducted on the amount of time the participants spent on the program and the question 'Is the time spent on the program related to cognitive style?' was asked. This section looks at whether the cognitive load of the program influenced the :

- total amount of time spent on the program for each version, and
- number of times a specific screen was accessed.

The cognitive load was measured after the first entry to specific screens, and then not again. The analysis considered the self-reported cognitive load for selected screens where the participant

accessed the screen more than twice. The screens reviewed were screens 12 and 19 from the animation version and screens 13-16 and 23 from the static images & text version.

The mean (SD) time spent on the animation version was 42.3496 (11.5879) minutes and for the static images & text version it was 45.6333 (14.0197) minutes. The effect size for the difference between these means is $d = 0.23$, which is small, indicating no practical significance.

4.6.4.1 Cognitive load and time spent on the animation version

This section looks at the cognitive load and time spent on the animation version. In section 4.5.2.1 (page 181) the time spent on the program was divided into three groups: Inadequate (IA), Adequate (AD) and More than Adequate (MA). The cognitive load for each of these time groups, by version is displayed in Table 4.59.

Cognitive load for time spent on program by version					
	Animation Version		Static images & text version		Effect size (d)
	n	Mean (\pm SD)	n	Mean (\pm SD)	
Inadequate	21	4.8095 (1.1653)	46	5.2457 (0.9694)	0.45
Adequate	72	5.2854 (0.9820)	62	5.6575 (1.0197)	0.36
More than Adequate	27	5.0558 (1.2853)	10	5.8667 (1.2368)	0.66

Table 4.59: Cognitive load for time spent on program by version

4.6.4.2 Cognitive load and time spent on the static version

We see in Table 4.59, when comparing versions, that the static images & text version had a higher cognitive load for all three the levels of time. All the mean values of cognitive load displayed in the table are in the medium range. Effect sizes indicate visible effect for all the comparisons of the means. The lower load for the 'Inadequate' time group was unexpected, and raises the question as to whether or not time is an indicator of cognitive load, since there was a lower load for the level of time that was categorised as 'Inadequate' and higher load for the level of time categorised as 'More than Adequate'. The reasons for spending so little time on the program were not followed up with the participants. It is therefore not possible to make any further assumptions about this data.

4.6.4.3 Cognitive load and multiple access to selected screens

This section considers the cognitive load and the number of times the participant accessed a particular screen. The analysis considered the cognitive load of selected screens where the

participants entered the screen **more than twice**. Table 4.60 displays the mean cognitive load for screen 12 in the animation version, where it was accessed more than twice.

Cognitive Load of Screen 12 in Animation version				
No of times entered screen	N	Mean (SD)	Minimum	Maximum
3	16	5.3389 (0.7815)	4	6.8
4	10	5.3167 (1.1832)	2.4	6.6
5	2	5.3000 (0.4243)	5	5.6
6	2	5.6000 (0.8485)	5	6.2

Table 4.60: Cognitive of screen 12 where there were multiple entries to the screen

Screen 12 was accessed between 2 and 6 times. The value of n for each of the items in Table 4.60 was small when compared to the value of n of the animation ($n = 120$). In Table 4.60 we see that the cognitive load of screen 12 where there were 3 and 4 entries were very similar: 5.3389 and 5.3167 respectively. The effect size for the comparison of these two means was very small $d = 0.02$, indicating no practical significance. The comparison of the means where there were 3 and 6 entries returned an effect size of $d = 0.31$, suggesting small to medium effect size. No t tests were conducted to determine statistical significance due to the very small sample size. It is unlikely that the result will be statistically significant.

Table 4.61 displays the mean cognitive load for screen 19 in the animation version, where it was accessed more than twice.

Cognitive Load of Screen 19 in Animation version				
No of times	N	Mean (SD)	Minimum	Maximum
3	21	5.5143 (1.2596)	3.25	8.6
4	5	5.096 (0.2428)	4.8	5.33
5	1	5.2000 (n.a.)	5.2	5.3
7	1	6.0000 (n.a.)	6	6

Table 4.61: Cognitive of screen 19 where there were multiple entries to the screen

Screen 19 was accessed between 2 and 7 times. The value of n for each of the categories in Table 4.61 is small when compared to the value of n of the animation ($n = 120$). In Table 4.61 we see that

the cognitive load of screen 19 differed for each of the entry groups. There does not seem to be a discernable trend, as an average cognitive load of 5.5143 was obtained when there were 3 entries, but the average load was reduced to 5.096 where the number of entries was 4. The average cognitive load where there 3 entries was also higher than the average cognitive load where the number of entries was 5. Only one participant entered screen 19 five times. The one participant who entered screen 19 seven times reported a cognitive load of 6.000. The effect size for the comparison of 3 and 7 entries was $d = 0.39$, but an interpretation of a visible difference is only applicable to the one participant who accessed screen 19 seven times.

The analysis now turns to selected screens in the static images & text version. Table 4.62 displays the mean cognitive load for screen 23 in the static images & text version, where it was accessed more than twice. This screen was exactly the same as screen 19 in the animation version.

Cognitive Load of Screen 23 in static images & text version				
No of times entered screen	N	Mean (SD)	Minimum	Maximum
3	16	5.3389 (0.7815)	4	6.8
4	10	5.3167 (1.1832)	2.4	6.6
5	2	5.3000 (0.4243)	5	5.6
6	2	5.6000 (0.8485)	5	

Table 4.62: Cognitive load of screen 23 where there were multiple entries to the screen

Screen 23 was accessed between 2 and 6 times. The value of n for each of the categories in Table 4.62 is small when compared to the value of n of the animation ($n = 120$). In Table 4.62 we see that like the cognitive load of screen 19 (Table 4.61) there seems to be no discernable trend. The average cognitive load for the group where screen 23 was entered 5 time is lower than for the group where the screen was entered 3 times. The highest average cognitive load was reported where the screen was entered 6 times, but only 2 participants entered screen 23 six times. The effect size for the comparison of 6 and 7 entries was $d = 0.35$, but an interpretation must be made very cautiously due to the small sample size.

The number of times the screen group 13-16 was analysed per screen, rather than as a group, since the cognitive load was measured after each of these screens. Table 4.63 (on the next page) displays the mean cognitive load for each group of entries for screens 13, 14,15 and 16.

Mean Cognitive load for number of times screen accessed																
	3		4		5		6		7		8		10		12	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Screen 13	3	5.9250 (1.0196)	2	5.9167 (0.8250)	1	6.6666 (n.a.)										
Screen 14	24	5.9069 (1.2746)	9	5.1815 (0.7678)	1	5.5000 (n.a.)	1	4.6667 (n.a.)	2	4.1500 0.4950	1	6.6666 (n.a.)			1	6.3333 (n.a.)
Screen 15	13	5.4718 (1.4234)	8	4.7666 (0.7972)	1	6.6666 (n.a.)							1	6.3333 (n.a.)		
Screen 16	4	5.0750 (1.2324)														

Table 4.63: Cognitive load of screens 13, 14, 15 and 16 where there were multiple entries to each of these screens

The data displayed in Table 4.63 is the first attempt to separate the cognitive load values of this group of screens, which up to this point has always been reported for the group of screens. Looking at Table 4.63 we see that the cognitive load of screen 13 is in the medium-high range. We also see that screen 14 was accessed the most by the participants, with one participant accessing it as many as 12 times. The cognitive load, as measured using the subjective rating, varied between 4.1500 and 6.6666 for this group that accessed screen 14 more than twice. In Section 4.6.1.4 in Table 4.36 on page 217 of this chapter the cognitive load for screen 14, using the direct method of measurement was reported. The cognitive load of screen 14 was not higher than that of screen 15 or 16 and yet more participants accessed this screen (screen 14) more than twice than those who accessed screen 15 and 16 more than twice, where the cognitive load, by direct measurement, was higher than that of screen 14.

4.6.4.4 In summary: other variables influencing cognitive load

Section 4.6.3 considered other variables that could have had an influence on cognitive load, including time. Section 4.6.4 considered cognitive load and the amount of time spent on the program in more depth.

The influence of the version of the program, gender, culture, whether or not participants had studied the topic previously, the rating of their knowledge about the topic, English as home language, WA style, VI style and total time spent on the program on the subjective rating of cognitive load was determined using a stepwise regression. The regression analysis was followed up with a confirmatory GLM. The one variable that appeared to have an influence on the cognitive load was the rating of knowledge about the topic. The finding was in the expected direction in that participants who indicated they had basic or intermediate knowledge about the topic had a lower cognitive load than participants who indicated they had known nothing about the topic. Home language and the total time spent on the lesson approached significance.

The analysis in Section 4.6.4 explored whether the cognitive load of the program influenced the total amount of time spent on the program for each version, and the number of times a specific screen was accessed. A comparison of the mean time spent on each version, using only effect sizes, indicated that the difference in time was not practically significant. Each version was then considered separately, using the three levels of time that had already been defined. The static images & text version had a higher cognitive load for all three the levels of time and effect sizes indicated practical significance ($d > 0.30$) for all the comparisons of the means. The lower load for the 'Inadequate' time group and the higher load for the level of time categorised as 'More than Adequate' was unexpected.

Screen 12 in the animation version was accessed between 2 and 6 times. The comparison of the mean cognitive load where there were 3 and 6 entries returned an effect size of $d = 0.31$. Screen 19 in the animation version was accessed between 2 and 7 times and the mean cognitive load varied

between 5.096 and 6.000. The size of this sub-sample was very small, making analysis complex, if not impossible.

Screen 23 in the static images & text version was accessed between 2 and 6 times. Like the cognitive load of screen 19 (which was the same screen) there seemed to be no discernable trend. The number of times the screen group 13-16 in the static images & text version was analysed per screen, rather than as a group, since the cognitive load was measured after each of these screens. The cognitive load of screen 13 was in the medium-high range. Screen 14 was accessed the most by the participants, with one participant accessing it as many as 12 times. The cognitive load varied between 4.1500 and 6.6666 for this group that accessed screen 14 more than twice.

4.7 The correlation between self-report of cognitive load and the direct measure of cognitive load

The second sub-question asked about cognitive load in this study was ‘What is the correlation between the participant’s self-report of cognitive load and the direct measure of the cognitive load?’

4.7.1 Comparing correlation of the measurement techniques by version

The hypotheses tested were:

Null Hypothesis 2d

There will be no difference in the cognitive load for each version when two different methods are used to measure cognitive load.

Alternate Hypothesis 2d

The two different methods used to measure cognitive load will return results that are the same for each version.

Null Hypothesis 2e

There will be no correlation between the self-report method and direct measurement method for determining cognitive load.

Alternate Hypothesis 2e

There will be a positive correlation between the self-report method and direct measurement method for determining cognitive load.

The cognitive load was measured by Smith (2007) using the direct method. A total cognitive load for the program was calculated for each participant.

The Mean and standard deviation for the cognitive load for each of the two versions, using the direct measurement, is presented in Table 4.64.

Version	n	Mean (SD)
Animation	1198	6.6408 (3.3348)
Static images & text	1463	5.6841 (3.4885)

Table 4.64: Mean cognitive load using direct measurement (Smith, 2007)

Note that the frequencies used to obtain the means are considerably higher when compared to the subjective rating technique (see Section 4.6.1.5 of this chapter, page 218 and 219). This is due to the methodology of the direct measurement technique, where cognitive load was measured at multiple points on each screen, rather than just once.

At an alpha of 0.05, an analysis of variance, using the GLM procedure, revealed a significant difference between the two versions, $F(1, 2659) = 52.39$, $p < 0.001$, $R^2 = 0.0193$ (Smith, 2007, page 71). The results indicate that the animation version had a statistically significantly higher cognitive load than the static images & text version. The R^2 , an indicator of the effect size is small and of no practical significance. This result must be interpreted with caution, because the large sample size gives a p-value that was statistically highly significant.

With this data available it was possible to compare the cognitive load obtained for each version using two different techniques. This comparison will focus on the total cognitive load values for each version as a whole. I also explored the correlation between the cognitive load values obtained using these two different techniques.

The Mean and Standard deviation (SD) for the different measurement techniques applied to each of the two versions is presented in Table 4.65.

Version	Direct Measurement		Subject rating	
	N	Mean (SD)	n	Mean (SD)
Animation	1198	6.640 (3.3347)	118	5.150 (1.0936)
Static images & text	1463	5.684 (3.4385)	118	5.514 (1.0349)

Table 4.65: Means for the cognitive load, using different measurement techniques

The cognitive load value obtained for the animation version, using the direct measurement method, was based on 1198 values and the cognitive load value obtained for the static images & text version

was based on 1463 values. In contrast the values obtained using the subjective rating method were based on 118 values for each version (one for each participant). To test for statistical significance it was necessary to pair off the 1198/1463 with the 118/118 observations. A mean cognitive load value was calculated for each participant for the direct measurement method. This value was then paired, per participant, with the value obtained using the subjective rating method. The difference between the mean values was determined and a univariate analysis performed on the difference variable. The Sign test (M) indicated that there was a statistically significant difference in the cognitive load values obtained for the animation version ($M = 24$, $p < 0.0001$), with a visible effect size ($d = 0.45$), but not for the static images & text version ($M = -4$, $p = 0.5195$), where the effect size is also small ($d = 0.05$).

Hypothesis 2d that the two methods used to measure cognitive load will return results that are the same for each version was therefore rejected for the animation version but not for the static images & text version.

4.7.2 Comparing correlation of the measurement techniques without considering version

The Mean (\pm SD) total cognitive load, without considering version, using the subjective rating method of measurement was 5.3324 (1.0780). The Mean (\pm SD) total cognitive load, without considering version, using the direct method of measurement was 6.1781 (2.4995). Pearson product moment correlation analysis between the direct method and subjective rating method cognitive load measures was carried out. A very small positive correlation between these two measures, with a small effect, was found, but it was not significant, $r = 0.076$, $p = 0.2462$.

Hypothesis 2e that there will be a positive correlation between the self-report method and direct measurement method for determining cognitive load was therefore not supported.

4.7.3 In summary

This section compared the direct measurement measure and subjective rating measure of cognitive load obtained for the program as a whole and for each version of the program.

Using the direct method Smith (2007) found a significantly higher cognitive load for the animation version, in contrast to this study where there was a significantly higher cognitive load for the static images & text version.

A comparison of the two different cognitive load measures by version indicated that the difference in the cognitive load was statistically significantly different for the animation version but not for the static images & text version.

The Pearson product moment correlation analysis revealed a very small positive correlation between the two cognitive load measures that was not significant.

The analysis now turns to consider the results of the posttest and the relationship between cognitive load, cognitive style and learning outcomes.

4.8 The interaction between cognitive style, cognitive load and learning performance in an authentic multimedia learning environment

The final sub-question of this study addresses the issue of learner performance: How do participants with different cognitive styles perform when using the same content with different cognitive load? A related question, not articulated in the list of sub-questions, is whether any learning took place at all. This section presents the findings related to the research question under the following headings:

- Results of the pre- and posttest, independent of cognitive style and cognitive load
- Cognitive style and learning performance
- The subjective rating of cognitive load and learning performance
- Cognitive style, cognitive load and learning performance

4.8.1 Results of the pre- and posttest, independent of cognitive style and cognitive load

4.8.1.1 Pretest results

The pretest, which tested knowledge relevant to the content of the multimedia, was the third and final method used to determine the prior knowledge of the sample. It was also the only objective method used. There were nine questions in the pretest (see Appendix D). The test assessed prior knowledge at the first two levels of Bloom's taxonomy (Clark, 2007) namely recall and comprehension. The test was marked electronically and the format of the questions included multiple choice, multiple response, drag and drop and short answer questions. A selection of the questions required the participant to view a static image prior to answering the question.

An example of such a question is illustrated in Figure 4.7.

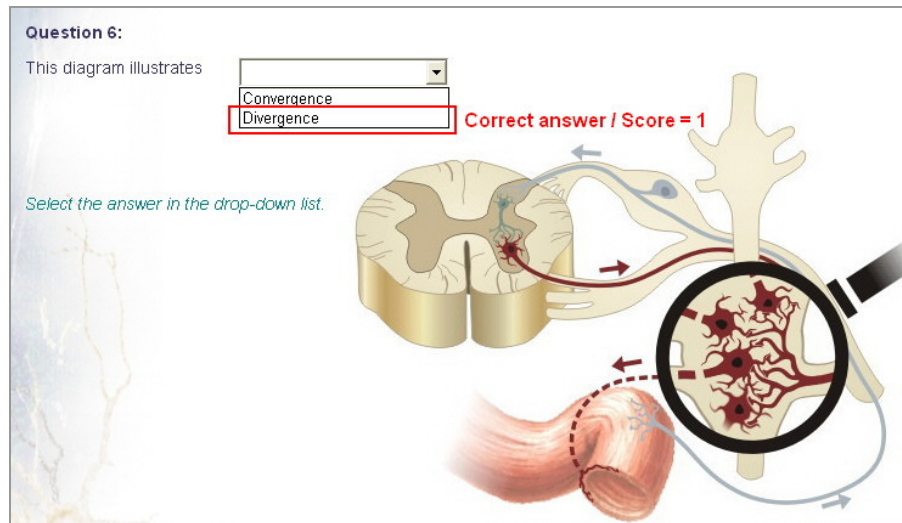


Figure 4.7: Example of question in pretest

The maximum possible score was 22 points. Two-hundred and thirty seven participants completed the pretest. The pretest scores for the sample ranged from 5 – 19, with a Mean (\pm SD) of 11.7257 (\pm 2.9308).

Figure 4.8 illustrates the distribution of the pretest results.

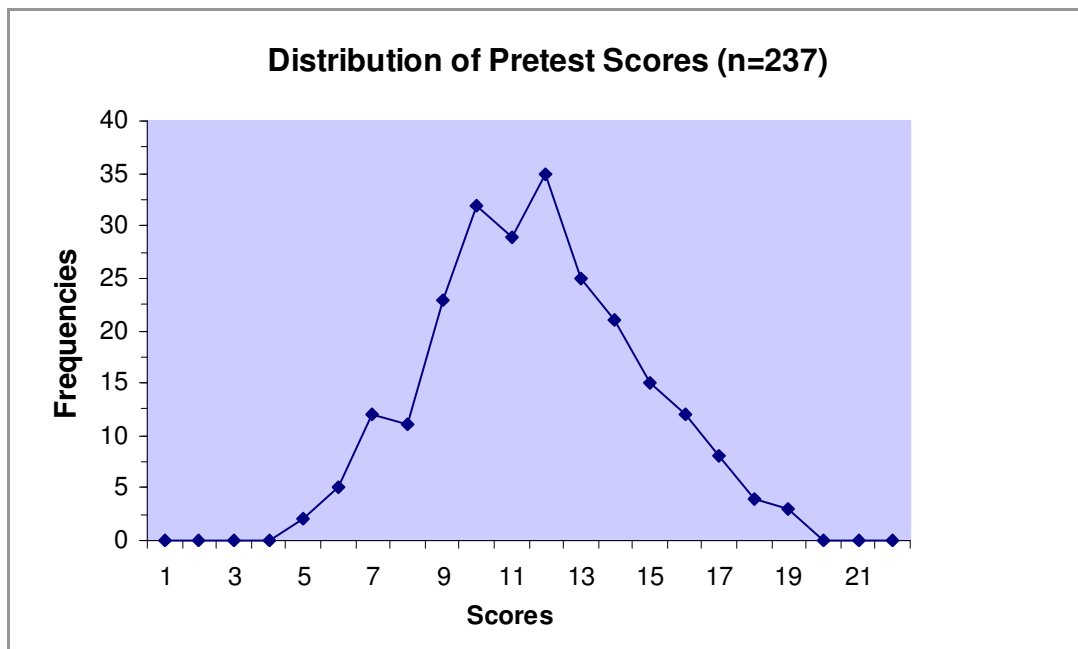


Figure 4.8: Distribution of pretest scores

Figure 4.8 illustrates a relatively symmetric distribution of the pretest scores for this sample. The questions, together with the correct answers (highlighted in bold for the multiple choice questions),

the frequency of respondents who answered the question correctly and incorrectly and the percentage of the sample who answered the question correctly are displayed in Table 4.66. The value is highlighted in bold where the percentage of **correct** answers is less than 50% of the sample.

From Table 4.66 we can see that the following questions were answered correctly by the majority of the sample (more than 75%):

- **Question 3.5:** Identifying the target organ in a diagram (86%)
- **Question 8.1:** What effect does increased sympathetic stimulation have on the heart(94%)?
- **Question 8.2:** What effect does increased sympathetic stimulation have on the pupils(90%)?

Question		Pretest results			
Question	Options (Correct answer in bold)	Correct (n)	Incorrect (n)	% Correct	
1	Which area of the brain is most directly involved in the reflex control of the autonomic system?	Cerebellum / Hypothalamus / Medulla Oblongata / Cerebral cortex	64	171	27%
2.1	Which organs are innervated mainly by the sympathetic nervous system?	Salivary glands / Stomach glands / Sweat glands /	115	82	48%
2.2		Blood vessels of the skin / Pancreas glands	169	68	71%
3.1	Drag and drop question - label a diagram with 6 labels	Afferent sympathetic neuron	122	114	51%
3.2		Preganglionic neuron	77	159	32%
3.3		Dorsal root ganglion	133	102	56%
3.4		Postganglionic neuron	87	149	37%
3.5		Target organ	205	31	86%
3.6		Sympathetic ganglion	131	104	55%
4	Which group of receptors are stimulated when the bladder is full?	Baroreceptors / Stretch receptors / Volume receptors	53	184	22%
5	Parasympathetic ganglia are located...	in a chain parallel to the spinal cord. / in the dorsal roots of spinal nerves. / next to or within the organs innervated. / in the brain.	60	177	25%
6	This diagram illustrates	convergence / divergence	148	89	62%

Question			Pretest results		
	Question	Options (Correct answer in bold)	Correct (n)	Incorrect (n)	% Correct
7.1	Which neurotransmitters are released by the neurons at the synapses in this diagram?	Ach	166	69	70%
7.2		Ach	153	83	64%
7.3		Ach	72	161	30%
7.4		Nor	142	90	60%
7.5		Ach	80	154	34%
8.1	Heart	Increases heart rate / decreases heart rate	223	12	94%
8.2	Pupils	Constrict / Dilate	214	20	90%
8.3	Bladder	Wall relaxes - bladder fills / Wall contracts - bladder empties	148	86	62%
8.4	Sweat glands	None / Increased secretion / Decreased secretion	137	97	58%
9	This diagram illustrates the	Autonomic Nervous System / Enteric Nervous System / Parasympathetic Nervous System / Sympathetic Nervous System	40	193	17%

Table 4.66: Pretest questions and answers and the number of participants who answered correctly and incorrectly

Question 3 was a diagram that needed to be labeled. Question 3.5 was the one structure on the diagram that the majority of the sample identified correctly. Their knowledge about the remaining structures was in fact poor, as can be seen in Table 4.66, considering that this content would have been covered in Anatomy as well.

Question 8.1 and 8.2 are common signs of sympathetic stimulation, part of the so-called 'fight or flight' response that is often described in the lay literature. The good performance for these two questions is therefore not surprising.

The questions most poorly answered (less than 30% of the participants answered the question correctly) were:

- **Question 1:** Which area of the brain is most directly involved in the reflex control of the autonomic system?
- **Question 4:** Which group of receptors are stimulated when the bladder is full?
- **Question 5:** Parasympathetic ganglia are located...(had to identify the location)
- **Question 7.3:** Which neurotransmitters are released by the neurons at the synapses in this diagram?
- **Question 9:** This diagram illustrates the... (had to identify the macro-structure of the sympathetic nervous system).

How did prior knowledge, determined by the subjective self-report methods compare with the pretest results? The pretest results were categorised into three performance groups: low (0 – 7), average (8 – 14) and high (15 – 22).

Table 4.67 displays the distribution of the pretest scores for these three performance groupings, further grouped according to their response to the question 'Have you studied this topic previously?'

Have studied content previously	Pretest results (n=236)					
	Low (0-7)		Average (8-14)		High (15-22)	
	n	%	n	%	n	%
Yes (n=167)	13	7.78%	125	74.85%	29	17.37%
No (n=69)	6	8.70%	51	73.91%	12	17.39%

Table 4.67: Learning performance and self-report of previous exposure to topic

The expectation was that proportionally more participants who had **not** studied the topic previously would fall in the 'Low' performance group. Similarly, proportionally more participants who **had** studied the topic previously would fall in the 'High' performance group. The results in Table 4.67 indicate that this was not the case. The distributions for both groups (those who had not studied the topic previously and those who had) were approximately the same for each performance group. The greater majority of participants also scored in the 'Average' range of performance. These results suggest that when setting up a research design that uses prior knowledge as a variable for creating different groups, actual pretest results are a better criterion to use than self-report measures.

Participants were also asked to rate the level of their knowledge about the topic. Four options were provided: I think I know and understand:

- absolutely nothing about the topic.
- the basic concepts of the topic.
- the concepts beyond a basic level of understanding.
- the topic at an expert level.

None of the participants thought they had expert knowledge about the topic.

Table 4.68 displays the distribution of the pretest scores for the three remaining responses, further grouped according to their response to the request to rate their level of knowledge about the topic.

Rating of knowledge	Pretest results (n=233)					
	Low (0-7)		Average (8-14)		High (15-22)	
	n	%	n	%	n	%
Nothing (n=19)	3	15.79%	16	84.21%	0	0%
Basic (n=203)	14	6.89%	151	74.39%	38	18.72%
Intermediate (n=11)	0	0%	8	72.73%	3	27.27%

Table 4.68: Learning performance and self-report of level of knowledge of topic

The expectation was that proportionally more participants who indicated that they knew nothing about the topic would fall in the 'Low' performance group. Similarly, proportionally more participants who indicated that they had more than a basic understanding of the content (labeled as intermediate in Table 4.63) would fall in the 'High' performance group. The results in Table 4.68 support this expectation. The results also indicate that, once again, the majority of the participants in each of the 'Rating of knowledge' groups scored in the average range of performance.

A Chi-Square analysis indicated that there was no significant difference between the levels of performance in the pretest and the rating of knowledge about the topic, $\chi^2 = 7.0316$, $df = 4$, $p = 0.1432$. The phi coefficient, a measure of the strength of the relationship the levels of performance in the pretest and the rating of knowledge about the topic is 0.1737. This is a small effect indicating that in practice there is no relationship between the levels of performance in the pretest and the rating of knowledge about the topic.

Table 4.69 displays the mean scores for the pretest grouped according to their response to the request to rate their level of knowledge about the topic.

Rating of knowledge	Pretest results (n=233)					
	Low (0-7)		Average (8-14)		High (15-22)	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Nothing (n=19)	3	6.3333 (0.5773)	16	10.3750 (1.8211)	0	n.a
Basic (n=203)	14	6.5714 (0.7559)	151	11.3642 (1.7300)	38	16.1578 (1.2197)
Intermediate (n=11)	0	n.a.	8	10.0000 (1.4142)	3	16.3333 (0.5773)

Table 4.69: Pretest results and self-rating of prior knowledge and understanding of topic

The majority of the sample scored in the 'Average' range of performance, where the Mean score for the pretest ranged between 10.000 (45%) and 11.3642 (51.6%). Participants who rated their knowledge as 'knowing and understanding more than the basic concepts' only managed to obtain an average score of 45% for the pretest. This is a relatively low score, which suggests that self-report ratings of knowledge are not necessarily reliable predictors of prior knowledge.

The final analysis to be discussed with regard to the pretest results is the findings for the pretest by version of content.

The percentages of participants who scored in each of the three levels of performance in the pretest, grouped by the version they subsequently used, are displayed in Table 4.70.

Version	Pretest results (n=237)					
	Low (0-7)		Average (8-14)		High (15-22)	
	n	%	n	%	n	%
Animation	5	26.32%	89	50.57%	25	59.52%
Static images & text	14	73.68%	87	49.43%	17	40.48%

Table 4.70: Pretest by version

A Chi-Square analysis indicated that there was no significant difference between the performance of the two groups of participants in the pretest, $\chi^2 = 5.8056$, $df = 2$, $p = 0.0549$. The phi coefficient, a measure of the strength of the relationship is 0.1565. This small effect size indicates that in practice there is also no real relationship between pretest results and grouping by version. The groups could therefore be regarded as equal in terms of prior knowledge, as measured by the pretest.

4.8.1.2 Posttest results

The computer-based test

The posttest was taken as soon as the participants had finished working through the multimedia. It had two sections. A computer-based test and two open-ended questions that were paper-and-pencil based. The computer-based test, which was marked electronically, was identical to the pretest with two exceptions:

- The order in which the questions were asked was changed.
- Question 9 required the user to identify the macro structure in the image. In the pretest the image illustrated the sympathetic nervous system and in the posttest an image of the parasympathetic system was displayed. The wording of the question was exactly the same, and the correct answer was selected from a drop-down list, which was identical in both the pre- and posttest.

The two open-ended questions assessed at the application level of Blooms taxonomy.

The maximum possible score was 22 points for the computer-based test. Each question in the open-ended question counted 10 points. The analysis of these two questions will be considered separately.

Two-hundred and thirty four participants completed the electronic posttest. The total posttest scores for the participants ranged from 8 – 22, with a Mean (\pm SD) of 17.200 and a SD (\pm 2.449). Figure 4.9 illustrates the distribution of the posttest results.

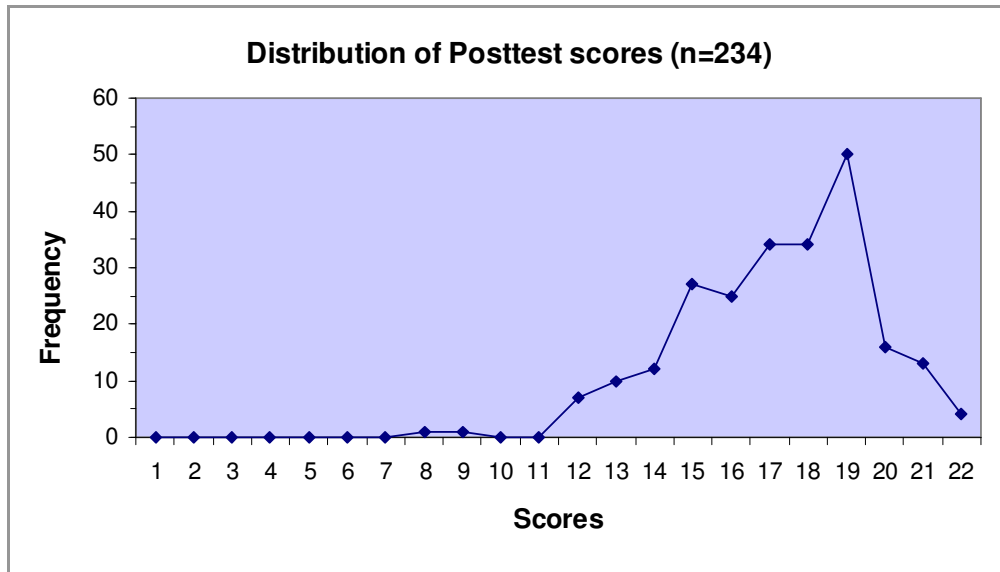


Figure 4.9: Distribution of posttest scores

The distribution of the scores for both the pre- and posttest is displayed in Figure 4.10. This display enables better comparison of the pre- and posttest scores.

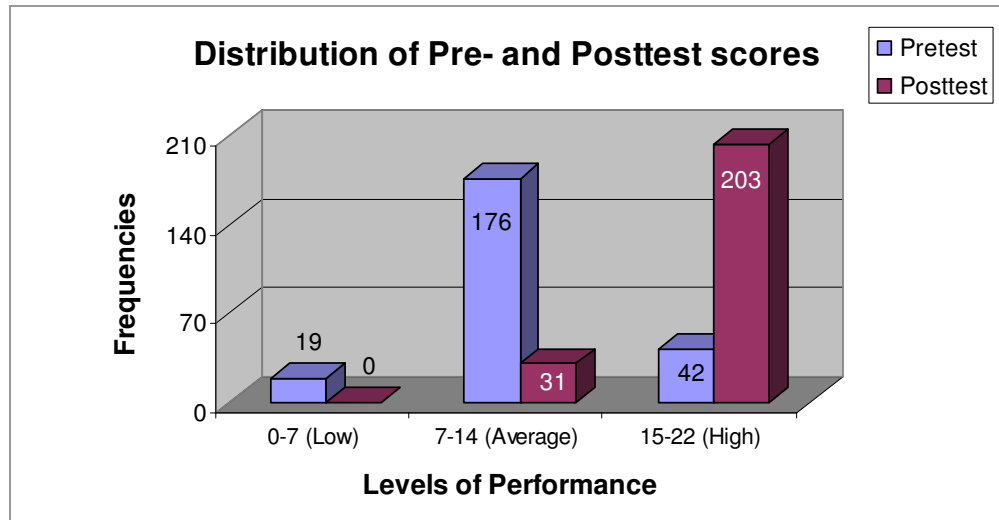


Figure 4.10: Distribution of pre- and posttest scores: Low, Average and High Performance

Figure 4.10 illustrates that in the pretest the scores of the majority of the sample were in the 'Average' level of performance, while in the posttest there was a marked shift in the distribution of the scores, with the majority of the sample scoring in the 'High' level of performance.

The posttest questions, together with the correct answers (highlighted in bold for the multiple choice questions), the frequency of respondents who answered the question correctly and incorrectly and

the percentage of the sample who answered the question **incorrectly** is displayed in Table 4.71 (see page 261-262). Where this percentage is less than 50% of the sample the value is highlighted in bold.

From Table 4.71 we can see that the following questions were answered incorrectly by the majority of the sample:

- **Question 1:** Identifying the area of the brain most directly involved in the reflex control of the autonomic system (79.5%).
- **Question 5:** The location of the parasympathetic ganglia.(69.9%)
- **Question 7.3:** The release of the specific neurotransmitters at the synapses between postganglionic neuron and the target organ (61.6%).

Both question 1 and 5 were straightforward multiple choice questions testing knowledge.

- The content for question 1 was presented in screen 15 in the animation version and in screen 21 in the static images & text version.
- The content for question 5 was presented in screen 5 for both versions.
- The content for question 7.3 was presented using a narrated animation. This was screen 19 in the animation version and screen 23 in the static images & text version.

		Question		Posttest results		
	Question	Options (correct answer in bold type)	Correct (n)	Incorrect (n)	% Incorrect	
1	Which area of the brain is most directly involved in the reflex control of the autonomic system?	Cerebellum / Hypothalamus / Medulla Oblongata / Cerebral cortex	48	186	79.5%	
2.1	Which organs are innervated mainly by the sympathetic nervous system?	Salivary glands / Stomach glands / Sweat glands /	186	48	20.5%	
2.2		Blood vessels of the skin / Pancreas glands	212	22	9.4%	
3.1	Drag and drop question - label a diagram with 6 labels	Afferent sympathetic neuron	203	30	12.9%	
3.2		Preganglionic neuron	204	30	12.8%	
3.3		Dorsal root ganglion	215	19	8.1%	
3.4		Postganglionic neuron	200	34	14.5%	
3.5		Target organ	225	9	3.8%	
3.6		Sympathetic ganglion	220	14	6%	
4	Which group of receptors are stimulated when the bladder is full?	Baroreceptors / Stretch receptors / Volume receptors	199	34	14.5%	
5	Parasympathetic ganglia are located...	in a chain parallel to the spinal cord. / in the dorsal roots of spinal nerves. / next to or within the organs innervated. / in the brain.	70	163	69.9%	
6	This diagram illustrates	convergence / divergence	202	32	13.7%	

Question			Posttest results		
	Question	Options (correct answer in bold type)	Correct (n)	Incorrect (n)	% Incorrect
7.1	Which neurotransmitters are released by the neurons at the synapses in this diagram?	Ach	209	25	10.7%
7.2		Ach	211	23	9.8%
7.3		Ach	89	143	61.6%
7.4		Nor	207	26	11.1%
7.5		Ach	138	95	40.7%
8.1	Heart	Increases heart rate / decreases heart rate	230	4	1.7%
8.2	Pupils	Constrict / Dilate	222	11	4.7%
8.3	Bladder	Wall relaxes - bladder fills / Wall contracts - bladder empties	216	18	7.7%
8.4	Sweat glands	None / Increased secretion / Decreased secretion	156	78	33.3%
9	This diagram illustrates the	Autonomic Nervous System / Enteric Nervous System / Parasympathetic Nervous System / Sympathetic Nervous System	168	71	30.3%

Table 4.71: Posttest questions and answers and the number of participants who answered correctly and incorrectly

The next analysis discussed is the results for the posttest by version of content.

The Mean (SD) posttest score for the participants who studied the animation version was 17.5294 (\pm 2.2953) and the Mean (SD) posttest score for the participants who studied the static images & text version was 16.8609 (\pm 2.5646). A *t* test for statistical significance indicated that this difference was statistically significant, $t = 2.10$, $df = 232$, $p = 0.0366$. The effect size for the difference between these means was 0.26, indicating an effect in the small to medium range and of no practical significance.

The percentages of participants who scored in each of the three levels of performance in the posttest, grouped by the version they studied, are displayed in Table 4.72.

Version	Posttest results (n=234)					
	Low (0-7)		Average (8-14)		High (15-22)	
	n	%	n	%	n	%
Animation	0	n.a	12	38.71%	107	52.71%
Static images & text	0	n.a	19	61.29%	96	47.29%

Table 4.72: Posttest by version

A Chi-Square analysis indicated that there was no significant difference between the groups reflected in Table 4.72 in the computer-based posttest, $\chi^2 = 2.1098$, $df = 1$, $p = 0.1464$. The phi coefficient, a measure of the strength of the relationship is -0.0949. This effect size indicates that in practice there is also no practically significant relationship between the posttest results, when using the three levels of performance, and grouping by version.

The open-ended questions

The participants answered two open-ended questions after completing the computer-based test. These questions tested at the application level of Bloom's taxonomy (Clark, 2007). The questions are presented in Appendix E. The maximum score possible for each question was 10 points. A memorandum was used to guide the marking. These questions were marked by myself.

The scores for the first question ($n = 217$), independent of version, ranged from 0 – 10, with a Mean (\pm SD) of 5.7673 (\pm 1.5921). The scores for the second question ($n = 217$), independent of version, ranged from 0 – 7, with a very low Mean (\pm SD) of 1.1613 (\pm 1.1291).

The analysis was extended to consider the performance for these two questions by version in order to determine whether the version studied had any influence on the results. The mean scores, by version, for Question 1 of the open-ended section of the posttest are displayed in Table 4.73.

Results for Question 1 of the opened ended posttest question						
	n	Mean (SD)	t	df	p	Effect size
Animation	109	5.7477 (1.7208)	-0.18	215	0.8561	0.02
Static images & text	108	5.787 (1.4585)				

Table 4.73: Mean score for Question 1 of the open-ended questions for each version of the program

The mean scores for the two groups were almost identical. A t test to determine statistical significance for the difference between the mean scores for Question 1 was done, although it was not expected that there would be a significant difference between the two groups. This was confirmed, $t = -0.18$, $df = 215$, $p = 0.8561$. The very small effect size of $d = 0.02$ points to no practical significance either. Version of the program did not appear to influence performance in Question 1 in any way.

The mean scores, by version, for Question 2 of the open-ended section of the posttest are displayed in Table 4.74.

Results for Question 2 of the opened ended posttest question						
	n	Mean (SD)	t	df	p	Effect size
Animation	109	1.3119 (1.2149)	1.99	215	0.0481	0.30
Static images & text	108	1.0093 (1.0185)				

Table 4.74: Mean score for Question 2 of the open-ended questions for each version of the program

The participants who studied the animation version had the higher mean for this question. A t test to determine statistical significance between the mean scores for Question 2 was done. The analysis returned a marginally statistically significant result, $t = 1.99$, $df = 215$, $p = 0.0481$. The effect size of $d = 0.30$, suggests that this difference has no practical significance. These results must be interpreted

very cautiously, due to the very low mean scores and the possibility that there were many confounding variables that influenced this outcome.

4.8.1.3 Learning gains

Figure 4.11 displays a detailed comparison of the pre- and posttest results, drilling down to look at the frequencies for each score in the range.

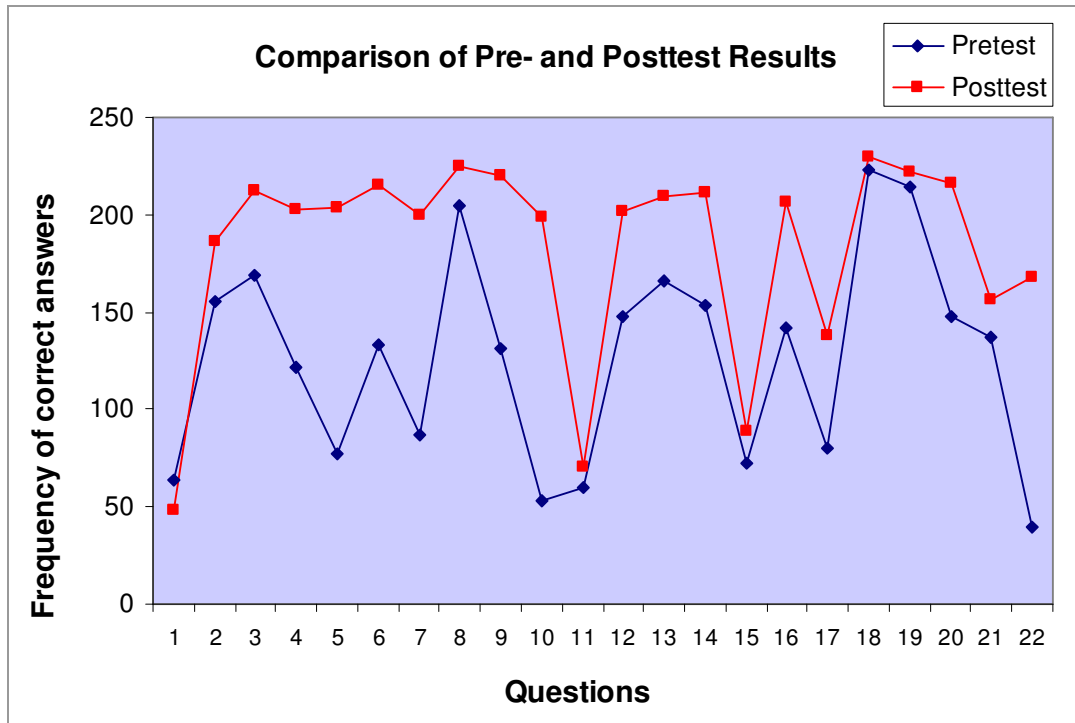


Figure 4.11: Comparison of pre- and posttest scores for each score in the range

In Figure 4.11 we can see that there was an increase in the number of correct answers in all the questions except Question 1, where in fact there were fewer correct answers in the posttest than the pretest. There were relatively small increases in the number of correct answers for Question 5, 7.3 and 8.4. These items have already been identified as problematic in Section 4.8.1.2. It would therefore appear that some learning took place. Was this learning gain significant or not?

The Mean (\pm SD) for the pretest and posttest, excluding the results of the two open-ended questions, was 11.7257 (2.9308) and 17.2008 (2.4491) respectively. There was a significant increase in the posttest scores. Table 4.75 displays the results of the pre- and posttest by version.

Learning Gain: Knowledge & Recall				
	Version 1: Animation		Version 2: Static images & text	
	n	Mean (SD)	n	Mean (SD)
Pretest	119	12.0840 (2.8152)	118	11.3644 (3.0118)
Posttest	119	17.5294 (2.2953)	115	16.8609 (2.5646)

Table 4.75: Mean scores for the pre- and posttests by version

A univariate analysis of the difference between the pre- and posttest results was conducted. The results of the Sign test (M) indicated that the difference was statistically significant, $M = 109$, $p < 0.001$. The learning gain in this study was significant, irrespective of version used.

4.8.2 Cognitive style and learning performance

This section discusses the learning performance of the different style groups in the sample. Since the results of the pretest indicated that there was no difference between the groups only the posttest results will be discussed.

4.8.2.1 Wholistic-Analytic style and learning performance

Using two categories for the WA style dimension it was found that 69% of the sample had an Analytic style and 31% were Wholistic in style.

General Linear Modeling analysis was conducted to determine if cognitive style had any influence on learning performance. The model used for the GLM was : Posttest = Wholistic-Analytic style, pretest and the interaction between WA style and the pretest. The GLM analysis returned a significant finding, $F(3, 227) = 23.74$, $p < 0.0001$, $R^2 = 0.2388$, but it was the pretest that accounted for this significant result, $F(1, 227) = 67.13$, $p < 0.0001$. There were no main effects found between WA style and learning performance, $F(1, 227) = 1.99$, $p = 0.1592$ and no interaction effects between WA style and the pretest, $F(1, 227) = 1.87$, $p = 0.1732$.

The least squares means for the posttest scores and standard errors for the two styles, the F statistic and p value are displayed in Table 4.76

	N	LS Mean	Standard error	F statistic	p value
Wholistic style	73	17.1094	0.2546	1.99	0.1592
Analytic style	162	17.2305	0.1712		

Table 4.76: GLM analysis results for WA style and learning performance

The Wholistic-Analytic style therefore did not appear to have any influence on leaning performance in the posttest.

4.8.2.2 Verbaliser-Imager style and learning performance

Using two categories for the VI style dimension it was found that 40% of the sample had a Verbaliser style and 60 % were Imagers in style.

General Linear Modeling analysis was conducted, with the posttest as the dependent variable and the VI style dimension as the independent class variable. The pretest results were entered as a covariate. The GLM analysis returned a significant finding, $F(3, 227) = 23.58$, $p < 0.0001$, $R^2 = 0.2376$, but it was once again the pretest that accounted for this significant result, $F(1, 227) = 69.69$, $p < 0.0001$.

The least squares means for the posttest scores and standard errors for the two styles, the F statistic and p value are displayed in Table 4.77.

	N	LS Mean	Standard error	F statistic	p value
Verbaliser style	94	17.3136	0.2271	23.58	0.6012
Imager style	141	17.1603	0.1849		

Table 4.77: GLM analysis results for VI style and learning performance

The Verbaliser-Imager style therefore did not appear to have any influence on leaning performance in the posttest.

In both Tables 4.76 and 4.77, which display the least squares mean scores for the posttest for the WA and VI style groups respectively, it is evident that the style groups performed almost equally well.

For both style groups the results indicate that cognitive style did not have an impact on the learning performance for this group of participants.

4.8.3 The subjective rating of cognitive load and learning performance

This section presents the results of the analyses conducted to determine whether cognitive load influenced learning performance in any way.

General Linear Modeling analysis was conducted, with the posttest as the dependent variable and the cognitive load and version as the predictor variables. The pretest results were entered as a covariate.

The GLM analysis returned a significant finding, $F(3, 229) = 24.05$, $p < 0.0001$, $R^2 = 0.2395$, but it was once again the pretest that accounted for this significant result, $F(1, 228) = 60.57$, $p = p < 0.0001$. There were no main effects found between version and learning performance, $F(1, 229) = 1.58$, $p = 0.2098$ or between cognitive load and learning performance, $F(1, 229) = 1.29$, $p = 0.2576$. Further post hoc comparison, using Fischer's Least Squares Means, demonstrated that the difference in the posttest means for the two versions were not significant when cognitive load was included in the analysis.

The least squares means for the posttest scores, standard errors, F statistic and p value are displayed in Table 4.78.

	N	LS Mean	Standard error	F statistic	p value
Animation version	120	17.3714	0.1999	1.58	0.2098
Static images & text version	118	17.0101	0.2025		

Table 4.78: GLM analysis results for version and performance

These results indicated that cognitive load, as measured using the subjective rating technique did not appear to influence the learning performance in any way.

4.8.4 Cognitive style, cognitive load and learning performance

The final sub-question of this study asks ‘How is learning performance influenced when content with different cognitive load is studied by learners with different cognitive styles?’

In Chapter 3 it was argued that if cognitive load is high, then cognitive style might influence performance, although this could be moderated by the amount of time spent on the program. The following was predicted:

Null Hypothesis 4a

There will be no difference in the posttest results of the Analytic learner who spent inadequate time on the program and who rated the cognitive load as low and the other Analytic learners.

Alternate Hypothesis 4a

The Analytic learner who spends inadequate time of the program, and who rated the cognitive load as low, will perform more poorly on the posttest.

It was also predicted that the interaction of time and cognitive load will influence the performance of the learners with VI style.

Null Hypothesis 4b

There will be no difference in the posttest results of the Verbaliser and Imager learner who spent inadequate time on the program and who rated the cognitive load as high and the other Verbaliser and Imager learners.

Alternate Hypothesis 4b

The Verbaliser and Imager learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest.

These hypotheses are of particular interest since there were proportionally a high number of participants who spent inadequate time on the program (discussed in Section 4.4.1, page 180).

4.8.4.1 Cognitive load, the Analytic style and learning performance

The first analysis considers the Analytic Style. The Mean and standard deviation for the posttest for this style, grouped by time spent on the program (Inadequate, Adequate or More than Adequate) and rating of cognitive load (Low, Medium or High) are displayed in Table 4.79 (next page).

The values of n for each group is very small, since the analysis only used a sub-set of the data. Significance testing is not reliable when the sample is so small. The results displayed here can only point to a trend and no interpretation or conclusions can be made. Hypothesis 4a predicted that the Analytic learner who spends inadequate time on the program, and who rated the cognitive load as low, will perform more poorly on the posttest.

I will consider the animation version first. In Table 4.79 we see that the participant (n = 1) in this group (inadequate time, low cognitive load) in fact had the highest score (posttest score of 21) for the animation version as a whole. The participant (n = 1) who scored the lowest (posttest score of 17) in the animation version spent more than adequate time on the program but rated the program in the high cognitive load range.

Posttest results for the Analytic Style												
	Animation Version						Static images & text version					
	n	Inadequate time	n	Adequate time	n	More than Adequate time	n	Inadequate time	n	Adequate time	n	More than Adequate time
Low cognitive load (≤ 3.9)	1	21 (n.a)	2	17.5 (2.1213)	3	19 (2.000)	1	18	1	20	0	n.a
Medium cognitive load (>3.9 and ≤ 6.9)	10	18.4 (1.7127)	48	17.083 (2.4996)	14	17.929 (1.94)	22	16.773 (2.3285)	42	16.952 (2.3783)	3	17.333 (1.5275)
High cognitive load (>6.9)	0	n.a	3	17.667 (02.0817)	1	17 (n.a)	0	n.a	5	15 (3.000)	2	16.6 (2.1213)

Table 4.79: Posttest results for the Analytic learner, by time and load

Turning to the static images & text version the Analytic learner (n = 1) who spent inadequate time on the program and who rated the cognitive load as low obtained 18 for the posttest. This was not the lowest score, but neither was it the highest. The highest score (posttest score of 20) was for the participant (n = 1) who spent adequate time on the lesson and whose rating for cognitive load was in the low range. The lowest average score was for the participants (n = 5) who spent adequate time on the program, but whose cognitive load rating was in the high range (> 6.9).

There is a trend that suggests that the higher the load the poorer the learning performance.

4.8.4.2 Cognitive load, the Verbaliser-Imager style and learning performance

The Means and standard deviations for the posttest for this style dimension, grouped by time spent on the program (Inadequate, Adequate or More than Adequate) and rating of cognitive load (Low, Medium or High) are displayed for the animation version in Table 4.80 on page 273 and for the static images & text version in Table 4.81 on page 275.

The values of n for each group in the respective versions are very small, since the analyses only used a sub-set of the data. Significance testing is not reliable when the sample is so small. The results displayed here can only point to a trend and no interpretation or conclusions can be made. Hypothesis 4b predicted that the Verbaliser and Imager learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest.

I will consider the animation version first. In Table 4.80 we see that for both the Verbaliser and Imager style there were no participants this group (inadequate time, high cognitive load). In the Verbaliser group the participant (n = 1) with the highest score (posttest score of 21) spent more than adequate time on the program and rated the cognitive load in the low range. This was the highest score for the VI style in the animation version. The participants (n = 28) with the lowest average score (Mean posttest score of 16.857) spent adequate time on the program and rated the cognitive load in the medium range. In the Imager group the participant (n = 1) with the highest score (posttest score of 20) spent inadequate time on the program and rated the cognitive load in the low range. The participants (n = 2) with the lowest average score (Mean posttest score of 15.5) spent more than adequate time on the program and rated the cognitive load in the high range. This was the lowest mean score for the VI style in the animation version.

Once again we see the trend that suggests that the higher the load the poorer the learning performance.

Posttest results Animation version												
Verbalisers						Imagers						
	n	Inadequate time	n	Adequate time	n	More than Adequate time	n	Inadequate time	n	Adequate time	n	More than Adequate time
Low cognitive load (≤ 3.9)	3	19.333 (1.5275)	0	n.a	1	21 (n.a)	1	20 (n.a)	3	18 (1.7321)	3	17 (2.000)
Medium cognitive load (>3.9 and ≤ 6.9)	8	18.125 (2.7999)	28	16.857 (2.8767)	9	18 (1.7321)	9	18.111 (1.2693)	37	17.27 (2.194)	11	18 (1.8974)
High cognitive load (> 6.9)	0	n.a	0	n.a	0	n.a	0	n.a	2	18 (2.8284)	2	15.5 (2.1213)

Table 4.80: Posttest results for the animation version learner for the VI style by time and load

Turning to the static images & text version we see in Table 4.81 we see that for both the Verbaliser and Imager style there were no participants this group (inadequate time, high cognitive load). In the Verbaliser group the participants ($n = 2$) with the highest mean score (posttest score of 18) spent more than adequate time on the program and rated the cognitive load in the medium range. The participants ($n = 22$) with the lowest scores (Mean posttest score of 16.455) spent adequate time on the program and rated the cognitive load in the medium range. In the Imager group the participants ($n = 3$) with the highest mean scores (posttest score of 19) spent inadequate time on the program and rated the cognitive load in the low range. This was the highest score for the VI style in the static images & text version. The participant ($n = 1$) with the lowest scores (posttest score of 16) spent inadequate time on the program and rated the cognitive load in the high range. This was the lowest mean score for the VI style in the static images & text version.

Once again we see a trend that suggests that the higher the load the poorer the learning performance.

This trend that suggests that the higher the load the poorer the learning performance was confirmed in subsequent analysis. A stepwise regression was performed for the dependent variable posttest outcome. The following variables and their interactions, were included in the regression analysis:

- Version of the multimedia
- Cognitive style (both dimensions)
- Subjective rating of cognitive load
- Total time spent on the lesson

The result of the regression was $F(2, 227) = 4.67$, $p = 0.0103$, $R\text{-square} = 0.0395$, a small effect, and $C(p) = 4.0197$. The variables retained in the stepwise regression were version and subjective rating of cognitive load.

Posttest results Static images & text version												
	Verbalisers						Imagers					
	n	Inadequate	n	Adequate	n	More than Adequate	n	Inadequate	n	Adequate	n	More than Adequate
Low (≤ 3.9)	0	n.a	0	n.a	0	n.a	3	19 (1.000)	2	18.5 (2.1213)	0	n.a
Medium (>3.9 and ≤ 6.9)	17	16.765 (2.6582)	22	16.455 (2.5209)	2	18 (2.8284)	24	16.167 (2.9439)	28	17.5 (2.1688)	4	17.5 (3.4157)
High (> 6.9)	0	n.a	0	n.a	0	n.a	1	16 (n.a)	5	16.4 (2.881)	2	16.5 (2.1213)

Table 4.81: Posttest results for the static images & text version learner for the VI style by time and load

The stepwise regression equation is presented in Table 4.82

Variable	Parameter Estimate	Standard Error	Type III Sum of squares	F-value	p-value
Intercept	19.7723	0.8709	3021.7804	515.44	< 0.0001
Version	-0.5420	0.3244	16.3605	2.79	0.0962
Cognitive load	-0.3327	0.1494	29.0866	4.96	0.0269

Table 4.82 Stepwise regression equation for posttest results

From the data it can be determined that there was a statistically significant main effect for cognitive load $F(2, 227) = 4.96$, $p = 0.0269$, and a marginally significant main effect for version, $F(2, 227) = 2.79$, $p = 0.0962$.

A confirmatory General linear model (GLM) was then run, based on the results from the stepwise regression. The model used for the GLM was: Posttest outcome = version, cognitive load and the interaction between version and cognitive load.

The GLM analysis returned a statistically significant finding, $F(3, 229) = 3.06$, $p = 0.0290$, $R^2 = 0.0386$. The subjective rating of cognitive load accounted for the significant result, $F(1, 229) = 0.03$, $p = 0.0278$. Further post hoc comparison, using Fischer's Least Squares Means, indicated that the posttest score was higher in the animation version, $M (\pm SE) = 17.4599 (\pm 0.2260)$ compared to the static images & text version, $M (\pm SE) = 16.9292 (\pm 0.2295)$. The effect size was small, $d = 0.26$, indicating that this might not be a visible effect.

Hypothesis 4a that the Analytic learner who spends inadequate time on the program, and who rated the cognitive load as low, will perform more poorly on the posttest is not supported.

Hypothesis 4b that the Verbaliser and Imager learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest is not supported.

4.9 Summary of Chapter 4

Chapter presented the results of the data analysis, using the research question and hypotheses as a framework. Both descriptive and inferential statistics were used to conduct the analysis. A detailed demographic profile of the participants was presented at the beginning of the chapter. There were more female participants in the study. The group was made up of predominantly white participants. The age range was between 17 and 37, with the majority of the participants in the 18 – 21 year old range.

The analysis sought to find answers for the following themes.

- The time spent on the program as a whole and on individual screens.
- The cognitive style profile of the participants and the relationship between style and time spent on the program.
- The extent to which cognitive style influences the use of the program. In this section the number of times selected screens were accessed was analysed in depth.
- The cognitive load of each research intervention as a whole.
- The cognitive load of selected screens were determined and compared in order to determine how the presentation format influenced the cognitive load.
- Cognitive load and the relationship with the amount of time spent on the program.
- The relationship between cognitive load and cognitive style.
- The cognitive load obtained using two different measurement techniques was investigated and the correlation between the two methods explored.
- The learning performance of the participants in both a pretest and a posttest and the learning gain.
- The influence of both cognitive load and cognitive style on learning performance.
- The relationship between cognitive load, cognitive style and learning performance.

The results will now be summarised, interpreted and discussed in the Chapter 5, the next and final chapter.



Discussion and Recommendations

5

5.1 Introduction

This final chapter is divided into 6 sections. The focus of each section is presented in Table 5.1.

Section	Focus
5.2	A summary of the research
5.3 – 5.6	A reflection of the lessons learned from the study under three headings: <ul style="list-style-type: none"> • Reflecting on the methodology. • Summary of the findings and reflections on how this study supports or contradicts existing empirical research, a substantive reflection. • Reflection on the contribution of this study to the field of cognitive style, cognitive load and multimedia learning, a scientific reflection.
5.7	Recommendations for instructional design practice and further research.
5.8	Conclusion

Table 5.1: Overview of Chapter 5

5.2 Summary of the research

The framework of this study is outlined in Chapter 1. Chapters 1 and 3 provide concise definitions of the constructs investigated in this study.

5.2.1 Purpose of the study

The purpose of this study was to explore the role that cognitive load and cognitive style play in the successful achievement of learning outcomes when narrated animation and static images are used in multimedia learning formats in an authentic learning environment. The study also investigated the relationship between cognitive load, which is influenced by both the nature of the content and the specific design strategies used, and the cognitive style of the individual who uses different multimedia formats.

Cognitive load, a multidimensional construct, was defined as the burden imposed on the cognitive system when performing a particular task, as measured using a subjective rating of mental effort.

A clear distinction was made between cognitive and learning style. The position taken was that these are two independent constructs. The focus of this study was on **cognitive style**, and the widely used definition of cognitive style as ‘an individual’s preferred and habitual approach to both organising and representing information’ (Riding & Rayner, 1998) was applied to this study.

Multimedia learning, also a multidimensional construct, was explained from three perspectives:

- various systems can be used to deliver the instruction, including paper, the computer or a cell phone;
- instruction is presented in one or more modes or presentation formats, for example text, illustrations and animation;
- at least one sensory channel through which information is processed, for example visual, auditory, tactile or any combination of these, is used.

These three perspectives define the domain of multimedia learning.

5.2.2 Background to the study

One of the aims of instructional design practice is to create an environment that facilitates and enables learning. A few key questions, arising out of my working environment, served as the drivers for the first steps of this study. These questions were initially very broad, and focused on issues such as:

- the balance between design and development effort, and between time and learning gain,
- the use of multimedia resources by the learner in the actual learning environment; and
- techniques and methods to improve the instructional design of multimedia instruction.

5.2.3 The literature review

The issues listed in Section 5.2.2 provided the direction for the literature review, which became increasingly narrow in focus. The initial literature review helped build the immediate case for the research questions and rationale of the study. This case was presented in Chapter 1. The aim of the review was to:

- ground the study within an appropriate theoretical framework, which was used to interpret the findings of the study,
- synthesize existing empirical evidence in order to establish the scope of the field, looking both for issues that have been addressed and issues that have been neglected in the literature,
- use this synthesis of existing findings as a benchmark against which to compare the findings of this study.

The outcome of this literature review is presented in Chapter 2. The theoretical frameworks selected for the study included the Cognitive Theory of Multimedia Learning, Cognitive Load Theory (CLT) and Riding's Cognitive Style Model. There were 3 themes in the literature review: cognitive style and multimedia learning, cognitive load and multimedia learning and multimedia learning in the health science education context. The cognitive load research literature directed the instructional design

perspective of this study, while the cognitive style research literature directed the focus on the individual learner who engaged with instructional material in a multimedia learning environment.

Each of these themes were explored from various perspectives, including contributions to the relevant theory, measurement of the construct or concept, research initiatives, current trends and directions within the theme, specific research into how the theme is thought to influence learning and achievement and the development of heuristics for instructional design practice.

5.2.4 Finding the research question

One major research question, with five sub-questions, was formulated for the study. These questions have already been presented in both Section 1.6 (page 14) and Section 3.4 (page 127) of this dissertation.

5.2.5 The rationale of the study

The case for the study was built by exploring the theoretical, empirical, methodological, media and contextual dimensions in the cognitive load, cognitive style and multimedia learning literature.

Recent empirical research in the field of cognitive style and multimedia either failed to address the fact that the outcomes of the research might be due to cognitive overload (Ghinea & Chen, 2003) or only hinted that there might be a relationship between cognitive style and cognitive load (Graff, 2003b). The results of Riding, Grimley, Dahraei and Banner's study (2003) seemed to indicate that effective working memory capacity had a major influence on the performance of learners with specific styles. The relationship between these constructs was an avenue of research that did not seem to have been explored in great detail.

As early as 1994 the literature on cognitive load theory stated that there are three factors that contribute to cognitive load: task characteristics, learner characteristics and the interactions between these two (Paas & van Merriënboer, 1994a). One of the learner characteristics listed includes cognitive style. Cognitive load research has explored the influence of prior knowledge and learner experience on cognitive load in considerable depth (Kalyuga, 2006; Kalyuga, Ayres, Chandler & Sweller, 2003; Kalyuga, Chandler & Sweller, 2001), but the field has been strangely silent on the influence of cognitive style.

The content domain of a large majority of the studies included mathematics, science, technical subjects such as electrical circuits, computer applications or statistics. These were conducted primarily within primary, secondary and vocational education contexts. Any one of a number of contexts could have been selected for this study, and many different contexts have been used in cognitive style, cognitive load and multimedia research. The issues of both cognitive style and cognitive load are under-researched in health science education. There is also a lack of research in

authentic learning environments and so this study was conducted in an authentic environment, using a topic that is part of the physiology curriculum for second year medical and dental students.

Methodological limitations in previous research included:

- studies that did not measure achievement directly (Riding, Grimley, Dahraei & Banner, 2003),
- small samples of under 100 participants (Ayres, 2006a; Dutke & Rinck, 2006; Mayer, Sobko & Mautone, 2003; Riding & Grimley, 1999; van der Meij, & de Jong, 2006),
- cautions that not all findings from experimental laboratory studies were easily generalisable to the classroom setting (Tabbers, Martens & van Merriënboer, 2004),
- failure to measure the cognitive load of the intervention (Chandler & Sweller, 1991; Ghinea & Chen, 2003; Mayer, Moreno, Boire & Vagge, 1999; Moreno, 2006), and
- giving the participants material to learn that was generally not relevant to their own coursework (Mayer, Fennell, Farmer & Campbell, 2004; Moreno, 2004).

Why consider animation and static images above other media options? The use of static images and text in instructional resources has received considerable attention in the research community since the early 1980s (Carney & Levin, 2002, Mayer, 2003; Mayer & Gallini, 1990; Mayer, Mautone, & Prothero, 2002; McKay, 1999; Moreno & Valdez, 2005; Verdi & Kulhavy, 2002). Improvement in technology has seen the increased use of sound, video, animation and 3D presentation formats in instructional materials. The impact of these newer media formats on cognitive processes are being researched with the same rigour and vigour that have been applied to researching the use of text and images in traditional classroom-based, face-to-face learning environments. My study contributes to the body of research investigating the newer media (animation in particular), with the added dimension of looking at how learners with different cognitive styles use and experience a strategy such as animation.

The empirical, theoretical and media imperatives suggest that there might be a relationship between cognitive load and cognitive style, but the existence and nature of such a relationship has not been explored in any depth.

5.2.6 The research methodology

5.2.6.1 The research approach

This was a quantitative study. An experimental and correlation design was used, in that the study aimed, on the one hand, to determine whether a particular intervention (multimedia learning with animation and images) made a difference for the learning outcomes of a group of participants, and on the other hand, it investigated whether the relationship between two factors (cognitive load and cognitive style) could have impacted the learning outcomes.

5.2.6.2 The research design

The variables under investigation were cognitive style (independent variable), cognitive load (independent variable), presentation format (independent variable) and achievement of learning outcomes (dependent variable).

A between-subjects design was used. In this design, subjects who had different cognitive styles (independent variable) were exposed to a different version (independent variable), and each version had a different cognitive load (independent variable). These three independent variables and their effect on the learning outcomes (dependent variable) were considered in the analysis of the results. The design also had the characteristics of a within-subjects (repeated measures) design. The repeated measure is performance in a knowledge and comprehension test, due to the fact that the same participants completed a pre- and posttest after random assignment to the research interventions.

The participants were randomly assigned to one of two presentation formats of the multimedia program, the narrated animation version (Version 1) or the static images & text version (Version 2). Both groups took Riding's Cognitive Styles Analysis test. Both groups took a pretest to determine prior knowledge at the beginning of the experiment. The pretest tested for recall and comprehension of knowledge. On completion of the instruction both groups took a posttest. The posttest assessed for recall and comprehension of knowledge, and included an additional section that tested for application of knowledge.

5.2.6.3 The research sample

The multimedia program designed to teach this topic is relevant for all health science students who study the topic of the Autonomic Nervous System (the population) for the first time (novice learners). In this study it was used by the students at the University of Pretoria (the target population). The second year medical, dental and physiotherapy students (the sample) were all studying the topic of the multimedia, the Autonomic Nervous System (ANS), at a time that coincided with this research program.

Initial sampling used convenience selection, which was deemed appropriate for this study for the following reasons:

1. The potential participants were available.
2. The group identified for selection had characteristics common to both the wider research population and the target population.
3. The number of students in the group allowed for an adequate sample size.

Once the sample had been identified, participants were randomly assigned to one of the treatment interventions. This is in line with the more rigorous approach required in experimental research.

5.2.6.4 The research data

Two separate datasets were collected for cognitive style. The two datasets were ratios that indicated the position of an individual on each of the two dimensions of Riding's Cognitive Style Model, namely the Wholist-Analytic (WA) and the Verbaliser-Imager (VI) dimension.

The methods of measuring cognitive load have been described in Chapter 2. This study used the self-report rating method. Smith (2007), working with the author of this study, measured the cognitive load using the direct measure dual-task methodology. Some of the data from Smith's study was used to answer the following question

What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?

Self-report rating of cognitive load required the participant to indicate, on a nine-point scale, the mental effort they invested in understanding the content. Using the mental effort scores from the screens where this self-report was administered, a total mental effort score for the entire program was calculated for each participant. The mean of the individual scores for each intervention provided a total score for the cognitive load of the intervention.

Learner performance was measured using a pre- and posttest design. The computer-based test items were the same for the pretest and posttest, except that the order in which the questions were presented differed for the two tests. A score was calculated by the computer for each participant for both the pre- and posttest. A score of 1 was allocated for each correct answer, giving a total score of 22 for both the pre- and posttest. The posttest also included a pencil and paper test, which tested application of knowledge. There were 2 questions in this section of the posttest. The maximum score for each question was 10.

5.2.6.5 The format of the research intervention

Table 5.2 provides a summary of the structure of the two formats of the program.

	Animation version	Static images & text version
Formats used to present same content	Animation on one screen	Static images and text using six screens
Number of screens in program	19	23
Total number of animations in program	2	3

Table 5.2: A summary of the major design similarities and differences for the programs

The design of the multimedia program was informed by previous cognitive load and multimedia learning research. A detailed summary of the instructional strategies and instructional design was presented in Chapter 3. The actual screens for the intervention are illustrated in Appendices G – K.

5.2.6.6 The research hypotheses

Research hypotheses were formulated, both as null and alternate hypotheses, for each of the research questions. These will be presented in the discussion in Section 5.3.2 of this chapter.

5.2.7 Conducting the study

This study included two pilot studies and a main study. The window of opportunity for conducting both the pilot and main study was very small, due to the fact that the intervention was part of the normal study programme for the participants. Initially only one pilot study was planned but a second pilot study was necessary in order to sort out some technological problems with data collection.

The number of participants in the class and the final sample for the first pilot study and the main study is presented in Table 5.3.

Study	Class / Potential population	Class	Sample	Completed
Pilot 1	Physiotherapists	38	38	35
Pilot 2	Volunteers from the residences on the Faculty of Education campus at the University of Pretoria	250	40	33
Main study	2 nd year medical and dental students	262	245	238

Table 5.3: Profile of the research sample

The detail of how each study was conducted was presented in Section 3.18 and 3.19 of Chapter 3.

5.2.8 Analysis of the results

Two hundred and forty five data logs were retrieved for analysis. Once the data had been cleaned and prepared for analysis the final sample included 238 participants.

Data was analysed using the SAS[®] software system. Descriptive statistics were reported as frequencies, means and standard deviations ($M \pm SD$) or standard error ($M \pm SE$). Analyses of data were performed using regression analysis, the general linear model (GLM), t tests, Chi-square analysis and Pearson's correlation. Confidence intervals of 95% were reported. Differences were considered significant for p values ≤ 0.05 . Effect sizes were measured and interpreted using the guidelines provided by Cohen (1988).

Chapter 4 presented the results of the data analysis, using the research questions and hypotheses as a framework for the chapter. These results and the interpretation thereof are summarised and discussed in Section 5.5 of this chapter.

The next three sections discuss and reflect on this study from three perspectives:

- A methodological discussion and reflection, which considers the research methodology and design.
- A substantive discussion and reflection, which compares the findings of this study with other related and similar studies.

- A scientific discussion and reflection, which discusses the contribution of this study to the existing body of knowledge in the three themes or research streams covered in this study. Areas for future research are recommended.

Together, these three perspectives aim to inform the reader about the lessons learnt from this study.

5.3 Methodological reflection

This section is a reflection on the research methodology, considering both the strengths and weaknesses of the final design. Using a framework similar to that of Chapter 3, this section reflects on the considerations and trade-offs that had to be made for each step of the research process. The design and the rationale for each decision made is explained in great detail in Chapter 3. This section reflects on whether the decision made was the correct one and considers the strengths and limitations of each component as it was implemented in the final design of this study.

5.3.1 The research approach and design

This was a quantitative study, which met the most important characteristics of quantitative research, explained in more detail in Chapter 3, Section 3.5. A quantitative approach aligns this study with the trends of the cognitive load, cognitive style and multimedia research streams, which use predominantly quantitative approaches to research. There are two strengths in working within the boundaries of existing trends:

- New research provides the necessary replication studies for in-depth investigation of specific research questions, which in turn contributes to theory building.
- Researchers are able to make meaningful comparisons between studies.

A potential weakness of always working within a quantitative paradigm is that issues which may arise from within a qualitative paradigm are often neglected. Qualitative designs also need careful planning in order to get to the rich data that is collected during qualitative research. During the analysis of the data collected in my study it became evident that a qualitative approach was needed in order to explore certain issues in depth, but the research design had not planned for a qualitative approach and that avenue of investigation was lost for this study.

The research design implemented was an experimental design using an authentic setting rather than a more controlled laboratory setting. Although the initial sample was a convenience sample, all subsequent allocation to the research intervention was done randomly, thereby ensuring that the design adhered to one of the most important requirements for an experimental design (Creswell, 2005; McMillan & Schumacher, 2001).

Chapter 3 (Section 3.6.1 page 129) made a case for using a between-subjects design, but there were also elements of a within-subjects (repeated measures) design, where the repeated measure

included comparing the cognitive load of a specific screen with several other screens in the same version of the program. Nothing in the subsequent implementation of this design indicated that the decisions made were incorrect and the motivation for using this particular experimental design still stands.

The strength of using an authentic context was that the content was relevant to the sample. I do believe that participants are then more willing to invest the mental effort necessary in order to master the new learning. The fact that the content of the research intervention was part of their course also motivated them to use the multimedia program to learn. The participant groups really 'knuckled down' and got on with the task of learning the content after the briefing session. My observation of the groups during the laboratory sessions was that the majority of the participants took the session seriously. There were no disruptive participants and it was never necessary to remind them to stay on task and stay focused. Even after a long session in front of the computer many participants willingly completed the offline posttest that tested application of knowledge. Many of the participants provided substantial answers to the two questions. When the time spent on each version was analysed it was therefore very surprising to see the number of participants who spent less than adequate time, defined specifically for each version, using the multimedia.

The weakness of the authentic setting was the inability to predict all the variables that could influence the study. Even if it is possible to predict these variables, it is not possible to control for every variable in any experimental setting, and especially in an authentic setting. It is therefore possible that variables not identified and controlled for could have influenced the results of the study. For example, due to the fact that the campus did not have an experimental research laboratory for multimedia learning, the study had to be conducted in the laboratories used by the participants every day. Not only was the content authentic, but so was the physical environment. Factors that are difficult, if not impossible, to control in an environment like this include ambient noise, possible distractions from other participants, lighting of the room and the amount of workspace available next to the computer.

I had considered using the laboratory setting where there was more control, even though there were calls in the literature for research to be conducted in more authentic learning environments.

The issues important for the design of this study are listed in Table 5.4. The ease with which each of these issues could be included in the design, given the time frames, availability of the sample and ability to randomly assign participants to research intervention, amongst others, are indicated for each of the options that were considered.

Issue	Authentic environment	Laboratory setting
Content relevant to sample	✓	✓
Content from the curriculum of the sample	✓	✓
Groups larger than 30	✓	?
Ability to find enough volunteers	✓	?
Randomisation	✓	✓
Control of variables	?	✓

Table 5.4: Reflecting on the research design

From Table 5.4 it appears that the issues considered were feasible in an authentic setting and so the final decision was made to go ahead with the research design using this authentic setting.

Two alternative approaches to the design of this study that should be considered in any replication studies include the

- point at which cognitive style is measured, and
- work needed to ensure that the cognitive load of the two research interventions differ significantly.

This study measured cognitive style at the start of the laboratory session. This contributed in part to the length of the session. A better design would have been to measure the cognitive style of the participants before the main session, if only to reduce the length of the experimental session. The end-result of this long session was that participants appeared to rush through the two open-ended paper-and-pencil questions that formed part of the posttest, particularly the final question. The low scores obtained by the majority of the participants for this second question could have been due to the fact that they simply did not have enough time to do justice to the question.

Another approach to consider is to use cognitive style as a starting point for randomisation. In other words, take the participants from one style dimension and randomly assign them to one of the two interventions, then take the participants from the second style dimension and do the random assignment. However, since the cognitive style model uses two independent dimensions, and each participant can be placed at some point on the continuum of both dimensions of style, a decision will have to be made about which of the two dimensions to use when grouping the participants.

The original intention was to go into the main study already knowing what the cognitive load of the particular version of the program was. The problems experienced with regard to the measurement of cognitive load in the pilot study have been described in detail in Chapter 3, Section 3.18. As a result the cognitive load was only measured for the first time during the main study. While there **was** a difference in cognitive load between the two versions and it was statistically significant, the effect size was still relatively small ($d = 0.33$). In the planning of future research designs the researcher should consider working on the cognitive load of each version until it is possible to take two interventions with clearly different loads, at a practical level of significance, into a second study.

5.3.2 The research sample

One of the strengths of this study was the large sample used – 238 participants. Once the participants had been randomly allocated to one of two versions they were randomly allocated again to one of two lessons for each version, and the size of each group was still between 50 – 60 participants.

Two of the issues considered in Table 5.4 regarding the planning, the design and sample were the ability to find volunteers and the ability to get research groups that were larger than 30 participants per group.

The participants in this study had an extremely heavy study programme. This in fact applies to all medical students in learning programmes across the globe. Any study that was not immediately relevant to their coursework would have to be conducted outside 'normal' working hours, which in real terms means after hours (after 17:00 during the week or over weekends). This time is usually spent studying or doing clinical work and so the chance of recruiting enough volunteers to ensure that the groups for the interventions were large enough was regarded as a high risk.

Homogeneity of the sample was also important. The cognitive load needed to be measured using learners who were novices with regard to their knowledge and understanding of the topic. Controlling homogeneity of the sample therefore required the sample to be selected from the same year group. Calling for volunteers from across the entire learning program would have interfered with this homogeneity and could have introduced other variables that might have been very difficult to control. The study might have produced an entirely different result.

Two methods for finding and retaining volunteers for a study is to offer them class credit (Dutke & Rinck, 2006; Renkl, Atkinson & Große, 2004; Rieber, Tzeng & Tribble, K. 2004; Wallen, Plass & Brünken, 2005) or pay each participant a small honorarium for their participation (Bodemer, Ploetzner, Feuerlein & Spada, 2004; Dutke & Rinck, 2006; Lee, 2007; Schwan, & Riempp, 2004). There were two limitations to this approach for this study. Firstly I was not in a position to offer extra credit for participation and neither was I able to secure enough funding for this study to pay each participant such an honorarium. Secondly, many learners in South Africa pay for their own tertiary

education and therefore do a variety of small jobs (they must be small to fit into their study schedule) for extra money. The end result of paying an honorarium might have resulted in a very skewed sample for the study. The risk of using a skewed sample that was not truly representative of the larger population of medical students was regarded as high and one that could be avoided. This risk provided further justification for using content relevant to the sample in an authentic setting.

One of the limitations of a large sample includes the logistics of getting rich qualitative data to support the quantitative results. Even grouping 238 participants into focus groups of 10 – 15 participants per group would have been an enormous undertaking requiring a great deal of time. This was logistically impossible in this study, given the restrictions placed on my access to the sample. One of the weaknesses identified in the literature review was the small sample size in many of the studies reviewed. I aimed specifically to address this weakness in my study. The trade-off was the opportunity to collect qualitative data within the time frame for the research. The complexity of the design also called for a large sample in order to have large enough groups that could be included in the data analysis.

5.3.3 The research instruments and data

All the instruments, with the exception of the open-ended questions in the posttest, were presented to the participants in electronic format and embedded into the research intervention at the appropriate points. This format worked well and all the data was eventually written out exactly as designed. The instruments appeared easy to use and the basic format was not changed after the pilot study. The problems in the pilot study were due to hardware problems with the computers used during the pilot study and were resolved before the main study took place.

The electronic file produced for each participant was very long - up to 32 pages for each participant. The reason for this was that the same experimental session and sample was used by Smith (2007). The design of the multimedia program, with the instruments embedded in the program, was almost complete when it was decided that Smith would investigate the cognitive load using the direct measurement technique. The mass of data that was generated eventually had to be divided into three different data sets. This proved to be an enormous undertaking, requiring multiple rounds of checking to ensure that no mistakes were made. In hindsight, a better way of getting the data would have been to write it out to several files instead of one large file. Smaller files would have speeded up the process of getting the data cleaned and ready for analysis.

In three of the eleven times participants were required to self-rate the mental effort they invested they were required to consider the content of either two or three screens. This in itself might have imposed unnecessary cognitive load on some participants, even though they had been briefed about the nature of this question. They were not able to go back to previous screens, the reason for this having been explained in Chapter 4. The one change I will recommend in future research is that the self-report rating of cognitive load only require the learner to consider one screen per rating.

Table 5.5 summarises the alignment of the instruments used in this study with the criteria for good instruments. The criteria used are those described by Creswell (2005).

	Self-report rating of Cognitive Load	Cognitive Styles Analysis	Pre- and posttests
When was the instrument first developed?	1993	1991	Developed specifically for study
Has the instrument been updated?	Some researchers have used a 7-point scale rather than a 9 point scale	Yes	Not applicable
Does the study use the latest version of the instrument?	✓	✓	
Is the instrument widely cited by other authors?	✓	✓	
Are reviews available for the instrument?	✓	✓	
Is there information about the reliability and validity of scores from past uses of the instrument?	✓	✓	
Does the procedure for recording data fit the research questions in this study?	✓	✓	
Does the instrument contain accepted scales of measurement?	✓	✓	

Table 5.5: Alignment of the instruments used in the study with criteria for good instruments

5.3.4 The research intervention

The authentic nature of the study and the requirements from the Department of Physiology (that the content replace one lecture of 40 minutes in duration) determined the scope of the program. An additional period was made available to fit in the Cognitive Styles Analysis, the electronic questionnaires and the posttest. The animation version of the program (19 screens in length) was developed first. Once this version had been developed, the animation (screen 12) was replaced with other screens for the static images & text version. Other design changes were also made in order to make maximum use of the opportunity to explore the influence of different presentation formats on cognitive load and learning performance.

Participants were given unlimited time to complete the program. This aligned with the authentic nature of the study, but made control of some of the variables and the subsequent comparison of the data more complex. An example of this was the fact that participants were allowed to go back to previous screens, but the cognitive load was only measured once. When the time on screen was considered all the entries were used in the time calculation. The total time spent on a screen was then used to see if there was any relationship between time spent on the screen and the self-rating of cognitive load. Participants knew that they could go back to the previous screens. If they had been

informed that they only had one chance to study the content they might have either spent more time on a particular screen or invested more mental effort (or both) and subsequently rated the cognitive load differently.

If I had to re-design this study I would make the following changes:

- Allow only forward progression through the program in order to collect the data, after which the participant can be allowed to go back to review selected screens.
- Limit the time to be spent on the program, or at least limit the time on screens where data was collected to facilitate better comparison.
- Consider moving the study from an authentic setting to a more controlled environment in a special experimental laboratory in order to better control the variables.

The actual implementation of the study went surprisingly well considering there were 70 students in the laboratory at one time. I used two laboratories simultaneously and had three research assistants in each laboratory to assist me. The participant groups were extremely co-operative. In spite of this I would recommend that smaller groups work together. If time had allowed I would certainly have considered running 3 – 4 sessions rather than 2 sessions. The session was also long and participant fatigue could have played a role in the outcomes.

5.4 Substantive reflection

This study addressed the following research question:

***The research question:** What is the relationship between cognitive style and cognitive load as factors in the achievement of learning outcomes when learning with different multimedia formats of the same content?*

This section will use each of the sub-questions and the appropriate hypotheses as a framework for the discussion and reflection.

5.4.1 The role of cognitive style in an authentic multimedia learning environment

The question:

What are the cognitive styles of the participants taking part in the study?

5.4.1.1 The cognitive style profile of the participants in the study

The literature review, discussed in Chapter 2, did not provide enough empirical evidence of a specific cognitive profile style for health science education students. There is also limited evidence in the literature that cognitive style influences a person's choice of occupation (Murphy & Casey, 1997).

Fifty-seven percent of (57%) the sample had an Analytic style, 28% an Intermediate style and only 15 % had a Wholistic style.

*The results from the analysis of the Wholistic-Analytic style dimension **supports** the hypothesis that more than 50% of the sample will have an Analytic or Intermediate style on the Wholistic-Analytic style dimension.*

Forty one percent (41%) of the sample had an Imager style, 31% were Bimodal in style and only 28 % were Verbalisers.

*The results from the analysis of the Verbaliser-Imager style dimension do **not** support the hypothesis that more than 50% of the sample will have a Verbaliser or Bimodal style in the Verbaliser-Imager style dimension.*

The sample in this study appeared to lean slightly more to the Imager style. Although I conducted extensive searches across several databases, including MEDLINE, I found no literature that reports a Wholistic-Analytic style profile, using Riding's CSA, for health science students, a finding supported by Cook (2005). Luk (1998), using the Field Independence/Field Dependence category of cognitive style, investigated the relationship between cognitive style and academic achievement in a group of nursing students who were following a distance education programme. Two studies were reported. In the first study 53% of the sample were Field Independent (a style label that can be compared with the Analytic style), while in the second study 65% of the sample were Field Independent. Chapman and Calhoun (2006), in a recent study that validated three learning styles, concluded that the medical students were found to be significantly more Field-Independent, favouring a more Analytic approach to learning. It would appear that medical and dental students appear to be more analytic in style, but there is not enough evidence in the literature to confirm that the results of my study are aligned with a particular cognitive style profile for medical and dental students. Such profiles have not been adequately researched and described in the health science education literature. There are **learning style profiles** reported in the literature (Chapman & Calhoun, 2006; Cook & Smith, 2006; Martin, Stark & Jolly, 2000; Smits et al., 2004) but this study has taken the position that cognitive style and learning style are two different constructs.

I found very little literature that reports on empirical studies that investigated the visual-verbal dimension of style in health science education. Effken and Doyle (2001) explored the role of the visual-verbal cognitive style with a small sample of undergraduate nursing students. The style was measured using an adaptation of Richardson's Verbaliser-Visualizer Questionnaire (VVQ). Seventy percent of the sample were found to be visualisers. Effken & Doyle did not indicate whether this profile was typical of nursing students in general. The sample ($n=18$) was in fact so small it is not possible to generalise to a larger population of health science education students. It is therefore not possible to determine whether the findings of this study are aligned with a general Verbaliser-Imager style profile for health science education students in general, and medical and dental students in particular.

The analysis of cognitive style was extended to consider cognitive style **and** gender and cognitive style **and** culture and cognitive style **and** time spent on the lessons for two reasons:

- The findings in the literature, discussed in Chapter 2, are mixed as far as the relationship between cognitive style and gender is concerned.
- Cognitive style profiles have not been adequately researched and described in the health science education literature. Findings related to gender, culture and time used during multimedia learning extend the knowledge about the cognitive style profile of health education students.

5.4.1.2 Cognitive style and gender

The analysis in this study indicated that there was no significant difference between the gender groups for both the Wholistic-Analytic and Verbaliser-Imager style dimensions. The cognitive style for each gender was determined for each of the three categories of the Wholistic-Analytic cognitive style. The effect size obtained for the Wholistic style ($d = 0.35$) was the only effect size that indicated that there might be a difference in cognitive style for gender. This is still a small effect and is not large enough to conclude without a doubt that the difference between male and female participants, who are health science students, for the Wholistic-Analytic style has any practical significance.

The results reported in the preceding paragraph support the argument that there is, in general, no gender effect for cognitive style (Riding et al., 1995). Graff, Davies and McNorton (2004) explored cognitive style and cross cultural differences in Internet use and attitudes to computers. Cognitive style was measured using the CSI of Allinson & Hayes. They found no gender effect and no interaction between gender and cognitive style. Take note that these conclusions from the literature were made based on empirical research that did not include the health science education field.

Further evidence that there is in fact little difference between the genders for most psychological variables is provided by Hyde (2005), who proposed the Gender Similarities hypothesis. The hypothesis states that most psychological gender differences are in the close-to-zero ($d \leq 0.10$) or small ($0.10 > d < 0.35$) range, a few are in the medium ($0.36 > d < 0.65$) range and very few are in the

large ($d = 0.66 - 1.00$) or even very large ($d < 1.00$) range. Hyde reviewed 46 meta-analyses of gender difference studies, including meta-analyses that assessed cognitive variables, and found support for this Gender Similarities hypothesis.

5.4.1.3 Cognitive style and culture

In a stepwise regression, using Wholistic-Analytic style as the dependent variable, culture was retained as one of the variables entered into the regression. A confirmatory GLM (General Linear Model) returned a statistically significant finding, with the culture of the participants accounting for the significant result. For the Wholistic-Analytic style dimension, Non-White participants were statistically significantly more Analytic than the White participants.

The relationship between cultural groups and cognitive style is not addressed extensively in the current cognitive style literature. Graff, Davies and McNorton (2004) also compared cognitive style and differences in Internet use and attitudes to computers using nationality and age as independent variables. They compared the Chinese and British cultures. Chinese students were found to have significantly more positive behavioural attitudes towards computers than UK students. Chinese students seemed to use the Internet more, but there were no cultural (nationality) differences between knowledge of the Internet and ease of use. Graff et al. (2004) concluded that the practical implications of their study was that designers needs to pay more attention to these cultural differences in designing web-based instruction.

The learning styles literature has addressed the issue of culture and ethnicity (Cho & Forde, 2001). Culture has also received attention in the instructional design literature (Chen, Mashhadi, Ang & Harkrider, 1999; Hedberg & Brown, 2002). Cho and Forde (2001) found a significant relationship between ethnicity (White versus Non-white) and learning styles, although they did point out that the purpose of their study was not to conclude that a certain ethnic group falls into a certain learning style category, but rather to encourage educators to use multiple teaching and assessment methods to accommodate style differences. They did not report on the probable causes for these differences.

The European Learning Styles Information Network (ELSIN) announced their 2008 conference towards the end of 2007. The theme for this conference is 'Style and cultural differences - how can organisations, regions and countries take advantage of style differences' (ELSIN 2008, 2008). The papers presented at this conference might shed more light on the relationship between cognitive style and culture.

What are the implications of the finding of my study, namely that Non-White participants were statistically significantly more analytic than the White participants? The effect size was 0.40, indicating practical significance, which cannot be ignored. The implications of this finding are difficult to determine without knowing what assumptions are made within the research population about the styles and characteristics of students from different cultural groups, especially if these are minority

groups. Cultural differences need to be explored simultaneously from multiple perspectives (psychological, cognitive, social, affective) in order to really understand the impact of culture on behaviour and learning.

5.4.1.4 Cognitive style and time spent on program

The analysis was done at two levels. I first looked at the time for the program as a whole before considering time at screen level.

Time spent on the program as a whole for the Wholistic-Analytic style dimension

For the Wholistic-Analytic style dimension it was predicted that the Analytic learner would spend more time studying the content of the program than the Wholistic learner. The Analytic learners generally spent more time than the Wholistic learners studying the content of both versions of the program. This difference was significant for the static images & text version, but not for the animation version. I also found that the Wholistic learner spent about the same amount of time on each version. This behaviour would seem to confirm the approach a Wholistic learner takes to learning (Riding & Rayner, 1998), in that the Wholistic learner's tendency to scan in order to get a big picture reduced the time spent on the static images & text version, even though there were more screens in this version. Proportionally the Wholistic learner spent more time studying the content of the animation version. Use of animation would constrain this scanning strategy of the Wholistic learner as he/she would have to watch the whole animation at least once in order to understand it.

The hypothesis that the Analytic learner will spend more time studying the content of the program than the Wholistic learner was therefore supported for the static images & text version of the program, but not for the animation version.

Time spent on selected screens for the Wholistic-Analytic style dimension

At screen level there was virtually no difference between the Wholistic and Analytic learner in the amount of time spent studying the animation. The difference in the time spent on the static images & text version for the comparison of the Wholistic and Analytic style was not statistically significant, but the effect size indicated that the difference **was** practically significant.

The analysis of the time **each style group** spent on the different screens in the two versions indicated that the Analytic learner spent significantly more time on the static images & text screens (screens 13-16) compared to the animation screen (screen 12). The Wholistic learner spent more time on the static images & text version but it was not significantly more than the time spent on the animation screen.

A suggestion for the fact that both style groups spent approximately the same amount of time studying the content of the animation could be viewed from two perspectives: either animation is a strategy that suits both styles in this style dimension equally well **OR** animation as an instructional strategy does not accommodate either style as effectively as does the strategy that used static images & text across several screens. The animation allows for less control by the individual learner. The learner could pause and restart the animation and use the progress bar to 'scroll' backwards and forwards to selected points in the animation. The use of the control features of the animation was not observed or tracked electronically. Each learner would have had to look at the entire animation at least once, from start to finish in order to understand it, irrespective of their particular cognitive style. The Wholistic learner would most likely only scan the content in the static images & text version, while the design would allow the Analytic learner better opportunity to study the detail without the interference determined by the pace of an animation. Since the Analytic learner processes more elaborately than the Wholistic learner they would take more time to work through the content and this was confirmed in this study. The instructional design of the static images & text version better accommodated the Analytic style and this type of learner seemed to make use of the opportunity by spending more time studying the content on these screens.

Time spent on the program as a whole for the Verbaliser-Imager style dimension

For the Verbaliser-Imager dimension it was predicted that the Verbaliser will spend less time studying the content of the program than the Imager. The intervention was visually rich. There were very few screens that did not use an image to illustrate the content. It was therefore expected that it would suit the Imager style, even though the literature indicates that the Verbaliser style can also use and benefit from visually rich content (Riding et al., 2003).

Participants with an Verbaliser style did spend less time studying the animation version than the participants with an Imager style, but the difference was not significant. The two style groups spent about the same time on average studying the content of the static images & text version and the difference was predictably not significant.

*The results from the analysis therefore do **not** support the hypothesis that the Verbaliser will spend less time studying the content of the program than the Imager.*

The animation included both images and narration and would therefore suit both style groups. The fact that the mean time for the two styles was not significantly different suggests that a narrated animation, as an instructional strategy, accommodates both the Verbaliser and Imager style. The static images & text version also included both images and text and also appeared to accommodate both the Verbaliser and Imager style.

As discussed in Chapter 3 and repeated here in order to avoid a split-attention effect, there were 29 images across 19 screens in the animation version and 39 images across 23 screens in the static images & text version. There were proportionally more images in the animation version. Verbalisers are more likely to select the text above the image when studying content and so it was predicted that the Verbaliser would spend proportionally more time studying the content of the static images & text version than the animation version.

The analysis of the time **each style group** spent on the two versions indicated that the Verbaliser style spent significantly more time on the static image & text version than the animation version. Proportionally however, because of the difference in the number of screens across the two versions, the Verbaliser learner spent more time studying the content in the animation version than the static images & text version.

*The hypothesis that the Verbaliser will spend proportionally more time studying the content of the static images & text version than the animation version was therefore **not** supported.*

Each style group seemed to spend more time on the version that suited their style. The static images & text version was thought to better suit the Verbaliser learner. While the difference was not statistically significant, the Verbaliser **did** spend more time studying the content of the static images & text version than he/she spent studying the animation version. The animation version was thought to better suit the Imager learner. While the difference was not statistically significant the Imager did spend more time studying the content of the animation version than he/she spent studying the static images & text version.

Time spent on selected screens for the Verbaliser-Imager style dimension

I then compared the time **each style group** spent on selected screens. There was very little difference in the time spent by the Verbalisers and Imagers on screen 12 in the animation version. It was not statistically significant. This was similar to the finding for the Wholistic-Analytic style dimension in this study. The difference in time spent by the Verbalisers and Imagers on screens 13-16 in the static images & text version was larger but still not statistically significant. The hypothesis that the Verbaliser will spend less time studying the content of the program than the Imager was also not supported at screen level for both the animation and the static images & text versions.

In considering each style separately the analysis indicated that both the Verbaliser and Imager style spent significantly more time studying the content of the static images & text version. A possible explanation for this could be that the user had more control over the content in the static images & text version, and used the time to review the text and/or images carefully.

In conclusion

The literature appears to be silent on the issue of the amount of time learners with different cognitive styles need in order to learn from multimedia instructional materials. Most of the studies reviewed gave the participants in the study a fixed amount of time to work through the instructional intervention (Ford & Chen, 2001; Graff, 2003b, 2005). These studies were carried out in formal experimental settings rather than in more authentic settings. I do not think that the amount of time needed for studying the content will impact on the instructional design as such, other than making a decision about what to include in each program or topic, but knowledge of time needed by learners with different cognitive styles has practical implication in authentic learning environments. When use is made of electronic resources in any learning environment, but particularly in a blended learning environment, it becomes necessary to know how much time to allocate for these learning opportunities.

The fact that Verbalisers spent more time studying the content of the programs supports the idea that Verbalisers are similar to Analytic learners in terms of their processing needs (Riding et al., 2003). It is possible that the analysis would have yielded different results and/or the findings would have been better explained if I had used certain style combinations in the research design and analysis. The style combinations for Riding's model would yield four groups namely Analytic-Verbalisers, Analytic-Imagers, Wholistic-Verbaliser and Wholistic-Imagers (Riding & Rayner, 1998). Researchers suggest that the Analytic-Verbaliser and Wholistic-Imager styles are unitary and the Analytic-Imager and Wholistic-Verbaliser styles are complementary (Evans, 2004; Riding & Rayner, 1998), but also call for further investigation into the preferences of these groups. A unitary style means that the strengths or weaknesses of the two style dimensions reinforce each other, while a complimentary style means that the strengths of the one dimension compensates for the weaknesses of the other dimension. The suggestion is that learners with a unitary style combination are less able to compensate for the weaknesses in the particular profile, while learners with a complementary style combination will use the strengths of one style to compensate for the weaknesses of the other style. John & Boucouvalas (2002c) explored the effects of the cognitive styles (using Riding's model) on user performance with audio. They considered the dimensions separately and then in combination and found different results when the sample was grouped differently. They concluded that it is not always possible to draw conclusions by analysing the results of one dimension in isolation as the influence of the other dimension can also effect the results.

The one implication of the findings related to cognitive style is that a special effort is needed to assist the Analytic learners in seeing a big picture view. In the context of a health science profession, where the trend is towards holistic, comprehensive patient care, this would mean making sure they are able to view and manage their patient holistically, by providing them with enough opportunity to improve the skills needed to see the big picture view. Another use for this finding lies in a question the instructional designer should ask, namely 'Will the planned design disadvantage the learner in any way?'

5.4.2 The role of cognitive load in an authentic multimedia learning environment

The question:

How do the participants rate the cognitive load of selected multimedia content?

The cognitive load was determined for the program as a whole and for specific screens in each version. In considering the program as a whole it was hypothesised that the animation version would have a higher cognitive load than the static images and text version.

5.4.2.1 Comparison of cognitive load across the two versions

The results indicated that when considering the cognitive load of the program as a whole, the static images & text version had a statistically significantly higher cognitive load than the animation version. The effect size of 0.33 is in the small to medium range, therefore interpretation of practical significance must be made with caution.

*The hypothesis that the animation version would have a higher cognitive load than the static images & text version was therefore **not** supported.*

This was an unexpected finding. There are several factors that could have contributed to this finding. Firstly the length of the program could have played a role. Secondly there were considerably more pop-ups and screens with text /image combinations in the static images & text version compared to the animation version. Although a definite effort was made to control for the split attention effect on the screens with text and images, it is possible that there was still some split attention that cumulatively influenced the extraneous cognitive load. The navigation in some of the screens in the static images & text version was more complex than the animation version, which presented the content in a linear format. In the static images & text version the user had to use a menu to access the different screens. The section could be navigated in a non-linear fashion. This could have increased the extraneous cognitive load.

The findings of my study were similar to those of Tabbers et al. (2004). Working in an authentic learning environment, they found that the mental effort ratings in their study were also low (in what I have categorised as the medium range) and were not statistically significant across the different versions of the research intervention they used.

However, when looking at recent literature a different perspective could also explain the findings of my study. Hegarty, Kriz & Cate (2003) suggested that when a learner uses animation he/she is only required to perceive the temporal changes taking place in the animation, while when using static representation the learner must infer these changes. This is assumed to be more difficult and requires greater mental effort. If this is the case, in order to understand the content of the static

images and text version the learner would have had to invest more mental effort compared to the animation version. The higher cognitive load found in the static images & text version could therefore be indicative of germane load rather than extraneous load. This load is not detrimental to learning, provided the total load is kept within the limits of working memory. The cognitive load for each version of the program was between 5 and 6, a range I have described as medium load.

Moreno and Marley (2007) investigated verbal and visual abilities and preferences, learning performance and perceptions about learning using three presentation formats. The multimedia program, covering the topic of Astronomy, included a narrated explanation with animated graphics, static graphics or no graphics (narration alone). Participants were given a program rating questionnaire to complete. An exploratory factor analysis of the items included in the program-rating questionnaire revealed an affective factor and a cognitive load (CL) factor. There were no learning or cognitive load differences between the animation and static images group.

The next step in the analysis was to consider a series of screen-wise comparisons across the two versions. These sets of screens isolated the instructional strategies and presentation formats: animation versus static images & text, and the use of pop-ups. It was hypothesised that the screen using animation will have a higher cognitive load than the alternative version using static images & text. It was also hypothesised that at screen level, the cognitive load will be the same in each version where the content and presentation format are the same.

Returning to the discussion of the findings in my study, the cognitive load for screen 12 (animation) was significantly higher than the cognitive load for the alternative static images & text screens (screen 13-16).

*The hypothesis that the screen using animation will have a higher cognitive load than the alternative version using static images & text **was** supported.*

I was only able to find two references in the literature since 2006 where the cognitive load was measured in the process of comparing animation and static images as presentation formats. Van Oostendorp and Beijersbergen (2007) used two sets of learning material (working of the human heart and the flushing and refilling of a toilet system) to compare the understanding, confidence and mental effort of three instructional conditions (animation, static images and guided animation). They found no significant difference in the amount of mental effort invested in the different instructional conditions. Höffler, Sumfleth and Leutner (2006) investigated the role of spatial ability when learning from an animation or a series of static pictures. The cognitive load was measured but not reported. I contacted Höffler and was able to review sections of his doctoral thesis (Höffler, 2007). I established two things: firstly he adapted the subjective rating scale of Paas & van Merriënboer (1993) for the measurement of cognitive load and secondly, and more importantly, he found that the cognitive load

was not statistically significantly different for the two formats of the instructional material. The results of my study therefore contradicts the findings of both studies cited here.

Theoretically, animation appears to be the strategy of choice when explaining processes (in my study the process and path of a stimulus in the autonomic nervous system) but the research comparing animation with static images does not support this assumption (Tversky, Morrison & Bétrancourt, 2002). Higher cognitive load does not imply poorer performance, but there is a risk that a higher load will have a negative impact on learning (cognitive load and performance will be discussed later in this chapter). Paas, van Gerven and Wouters (2007) suggest that in the research reviewed by Tversky et al. (2002) it is possible that the animations were not designed with sensitivity towards the processing limitations of the working memory and could therefore have interfered with the learning process.

Unfortunately the methods used to measure cognitive load do not inform about the relative contributions of each type of cognitive load (Whelan, 2007). If the cognitive load obtained in the animation version of my study was due to germane load then it is unlikely that this higher cognitive load was detrimental to learning, since the total load was in the category I had described as the medium range. Germane load has been shown to facilitate the transfer of learning (Bodemer et al., 2004; Schnotz & Rasch, 2005). However, if the cognitive load was due to extraneous load then methods must be sought to reduce the cognitive load, so that cognitive resources can be freed up for deeper processing.

Paas et al. (2007) discuss some suggestions for designing animation that is sensitive to the limitations of working memory: segmenting animations, cueing learners to specific features during the animation, increasing the level of interactivity with the animation. Paas et al. (2007) compared the instructional efficiency of three instructional strategies that were used **after** learners had studied an animation. They found that interaction that required the learner to either reconstruct or deconstruct the process studied in the animation was superior to the non-interactive strategy that merely required the study of static images extracted from the animation.

5.4.2.2 The relationship between cognitive load and cognitive style

There was no evidence in the literature of attempts to consider the relationship between cognitive load and cognitive style. Using existing empirical evidence about how the design of learning material influences cognitive load **together with** existing evidence about how learners with different cognitive styles are thought to process information and learn, the reasoning and argument behind the hypotheses of the relationship between load and style were presented in detail in Section 3.16.3 in Chapter 3.

Wholistic learners prefer to learn using big picture views, while the Analytic learner prefers step-by-step detail. The animation presented the information in one continuous session, although the user could control the pace of the animation. It was expected that this would suit the learner who had a

Wholistic style. The animations were also very visual and this was expected to suit the learner who had the Imager style. It was therefore hypothesised that the:

- cognitive load of the animation would be lower for the Wholistic learner than for the Analytic learner.
- cognitive load of the text and static image version, used as an alternative for the animation, would be lower for the Analytic learner than for the Wholistic learner.
- cognitive load of the animation would be lower for the learner with an Imager style than for the learner with a Verbaliser style.

Cognitive load and Wholistic-Analytic style

The difference in the cognitive load of the program for the Wholistic and Analytic style was small for both the animation and the static images & text version. The cognitive load of the animation was lower for the Wholistic learner than for the Analytic learner, but the difference was not statistically significant.

*The hypothesis that the cognitive load of the animation would be lower for the Wholistic learner than for the Analytic learner was therefore **not** supported.*

It was hypothesised that the cognitive load of the text and static image screens would be lower for the Analytic learner than for the Wholistic learner. It was found however that the Analytic learner reported a higher cognitive load than the Wholistic version, although the difference was not statistically or practically significant.

*The hypothesis that the cognitive load of the text and static image screens, used as an alternative for the animation, would be lower for the Analytic learner than for the Wholistic learner was therefore **not** supported.*

I analysed the cognitive load from another perspective and compared the cognitive load for each style by version used. Wholistic learners who used the static images & text version reported a higher cognitive load than the Wholistic learners who used the animation version. The difference was not statistically significant, but the effect size, in the small to medium range, indicated possible practical significance. Analytic learners who used the static images & text version also reported a higher cognitive load than the Analytic learners who used the animation version and the difference was statistically significant, with an effect size in the small to medium range, indicating possible practical significance.

Why would there be a statistically significant difference between the two versions for the Analytic learner, but not for the Wholistic learner? I present two explanations for this finding. The first relates to the relationship between cognitive style and cognitive load and the second relates to the issue of sample size.

The ability of the Wholistic learner to scan rather than study detail would mean that this learner would not be influenced too much by the fact that the static images & text version was slightly longer, while the Analytic learner would pay great attention to the detail of every screen. The static images & text version was also more text rich than the animation version. The Analytic learner, with their elaborate processing style, would experience and report that they invested more mental effort (and therefore a higher cognitive load) in the static images & text version compared to the animation version. In reconsidering the literature that describes how Wholistic and Analytic learners process information (Riding et.al, 2003; Riding & Rayner, 1998) the direction of the results (Analytic learners report a higher cognitive load than Wholistic learners) provides support for the findings of Riding et al. (2003) namely that the Analytic learner processes information elaborately. I suggest that this elaborate processing by the Analytic learner would require more investment of mental effort (and therefore higher cognitive load) than the scanning strategy of the Wholistic learner. I also suggest that this increased mental effort is a manifestation of germane cognitive load, which is beneficial for learning. The results of this deeper processing should be reflected in the learning performance, as measured by the posttest. I will consider this later in this chapter.

Regarding the issue of sample size, it is documented in the research methodology literature that statistical significance tests have a tendency to yield small p values as the size of the data set increases in size (Ellis & Steyn, 2003). The two sets of comparisons in this section of the study are a case in point. The value of n for the comparison of the means for the Wholistic learner was 73, while the value of n for the comparison of the means for the Analytic learner was 162, more than double that of the Wholistic group. The absolute difference in the means (mean difference between the animation and static images & text version for the Wholistic learner and mean difference between the animation and static images & text version for the Analytic learner) were approximately the same and yet it was the group with the larger sample size that returned the significant p value. In turning to look at the effect size the data indicates that the effect sizes were very similar. Both were in the small-to-medium range, 0.31 and 0.34 for the Wholistic and Analytic group respectively suggesting that this is a visible effect.

The higher cognitive load for the Wholistic learner in the static images & text version is not necessarily related to their processing style but to the fact that they would have needed more time to study the longer version. Time has been regarded as an indicator of cognitive load (Chandler & Sweller, 1991).

Cognitive load and Verbaliser-Imager style

The total cognitive load for the Verbaliser-Imager style dimension, irrespective of version, was almost exactly the same: mean cognitive load of 5.3070 for the Verbaliser style and mean cognitive load of 5.3472 for the Imager style. When the versions were considered separately the mean cognitive loads for each style were also very close to each other. There does therefore not appear to be any relationship between the Verbaliser-Imager style and cognitive load.

A suggestion for this finding relates to the definition of the Verbaliser-Imager style dimension. The Verbal-Imagery style addresses how the individual is inclined to represent information during thinking. This inclination can be towards verbal representation or thinking by means of mental images (Riding & Rayner, 1998). Representing information and processing information (the focus of the Wholistic-Analytic style dimension) are two different cognitive activities. I suggest that it is the processing of information that imposes the load (germane load) rather than the representation of information.

The learner with an Imager style did report a lower cognitive load than the learner with a Verbaliser style for the animation version, but the difference in the cognitive load was not statistically or practically significant.

*The hypothesis that the cognitive load of the animation version will be lower for the learner with Imager style than for the learner with Verbaliser style was therefore **not supported.***

The relationship between cognitive load and Verbaliser-Imager style was also analysed from another perspective. The cognitive load for both the Verbaliser and the Imager learner was compared by version. Both the Verbalisers and Imagers experienced a higher cognitive load in the static images & text version. This is most likely related to the length of the program. In considering the Verbaliser style, the difference in the cognitive load by version was neither statistically or practically significant.

A comparison of the versions for the Imager style did prove to be statistically significant and the effect size also suggested practical significance. The static images & text version had the higher load and this is most likely related to the length of the program.

5.4.2.3 A comparison across the research interventions where content and presentation format were the same

A comparison of the screens in the two versions that had the same content and the same presentation format was conducted. Two sets of screens were compared: screen 19 (animation version) was compared with screen 23 (static images & text version) and screen 5 (animation

version) was compared with screen 5 (static images & text version). The findings in this group of comparisons were completely unexpected when viewed from a purely cognitive load perspective.

Screen 23 had a significantly higher cognitive load than screen 19. Screen 5 in the static images & text version also had a higher cognitive load than the same screen in the animation version.

*The hypothesis that at screen level, the cognitive load will be the same in each version where the content and presentation format are the same was therefore **not** supported.*

It is possible that participant fatigue contributed to the cognitive load difference between screen 19 and 23. Both screens were the last screen in the program and the static images & text version was longer than the animation version. It is unlikely that the finding with regard to Screen 5 can be explained by participant fatigue, as this was the fifth screen in the program for both versions.

Another explanation for these differences could be related to the method of measuring cognitive load. I reviewed the cognitive load obtained by the direct measurement method for these screens (Smith, 2007). There were very small differences in the mean cognitive load of the screens (screen 19 compared to screen 23, screen 5 in each version). The effect sizes for the comparison of the means were calculated and were found to be too small to have any practical significance. If the direct measurement method is more accurate in determining cognitive load then it can be said that the cognitive load of these screens is similar.

5.4.3 The correlation between self-report and direct measurement as techniques in measuring cognitive load

The question:

What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?

I compared two methods of determining cognitive load and investigated the correlation between the cognitive load scores obtained from each technique. The first hypothesis was that the two methods used to measure cognitive load would return results that are the same for each version.

Using the direct method of measurement it was found that the animation version had a significantly higher load than the static images and text version. This contradicts the finding of the subjective rating method where the static images and text version had the significantly higher cognitive load.

The cognitive load measured obtained from the two techniques were compared by version. For the animation version the direct method of measurement returned a significantly higher load than the

subjective rating technique, while for the static images & text version the difference between the cognitive load measurements was small and not significant.

The hypothesis that the two methods used to measure cognitive load will return results that are the same for each version was supported for the static images & text version, but not for the animation version.

The next question is 'Which method is the most accurate?' The one advantage of the direct measurement method is that it measures the cognitive load at the time when the load is induced, while the subjective rating is made after the learning event (Brünken, Plass & Leutner, 2003). Paas et al. (2003) state that there is empirical evidence that reliable measures can be obtained with unidimensional scales (like the one used in my study) and that these scales are sensitive to small variations in cognitive load. They are regarded as valid, reliable and unintrusive measures. In the same article they state that the direct method using secondary tasks is also a highly sensitive and reliable measure. Brünken, Plass & Leutner (2003), in their discussion of the different methods used to measure cognitive load, indicate that although the subjective method appears to be reliable, it is still unclear how mental effort relates to actual cognitive load.

Another avenue to explore is to look at what type of cognitive load each method is supposedly measuring. The literature on studies that have compared different methods of measuring cognitive load is sparse. Paas et al. (1994b) compared physiological measures (heart rate variability) and the subjective rating technique and found that the method using heart rate variability was neither reliable or sensitive enough to detect differences in task complexity. The subjective rating scale proved to be a reliable and sensitive measure. Task complexity relates to intrinsic load. Whelan (2006) also compared two approaches to the measurement of cognitive load – self-report questionnaires and dual-task methodology. The different instruments did not produce uniform results. He concluded that each method reflects a different type of cognitive load. The dual task methodology showed its strength in assessing extraneous load, while Paas' instrument was only sensitive at the mean levels of significance to variations in extraneous load and did not show adequate sensitivity to high and low extraneous cognitive load conditions. Whelan (2007) reviewed the literature, looking for alternative approaches to measuring cognitive load. His overview of the efforts in the field of functional magnetic resonance imaging (fMRI) and his subsequent argument for using neuroimaging techniques adds to this new direction for cognitive load measurement research (Tomasi, Chang, Caparelli & Ernst, 2007).

Using the conclusion of Whelan (2006) to explain my findings, I suggest that if the intrinsic load of the material used in my study is regarded as equal for the two versions, there should have been no difference in the cognitive load of the two formats, as measured using the direct method of measurement. This was not the case, suggesting therefore that the extraneous load was different for

the two versions and this influenced the outcome of the measurement. If the direct method using dual-task methodology is the more sensitive measure for extraneous load then I suggest that there were extraneous load issues in the animation that outweighed the extraneous load imposed by the navigation through the static images & text version, as discussed in Section 5.4.2.1 of this chapter.

If this is the case, then the design of the animation in my study did not apply enough of the techniques thought to create animations that are sensitive to the limitations of working memory. Research into the efficacy of these techniques was published in the 2007 special issue of the *Applied Cognitive Psychology* journal and include attention cueing (de Koning, Tabbers, Rikers & Paas, 2007) adequate learner control (Hasler, Kersten & Sweller, 2007), segmenting methods (Moreno, 2007) and interactive manipulation of static images following animation (Paas et al., 2007). The animation in my study did allow for learner control in the form of stop, pause and play buttons being available to the participants while watching the animation. The use of these controls was not tracked however.

If the screens presenting content in the static images & text format, used as an alternative for the animation is seen as manipulation of intrinsic load, then I should have found that the static images & text version had the lower cognitive load. This was not so when considering the program as a whole, but it **was** true when I only considered the relevant screens that were of particular interest in the study (screen 12 compared to screens 13-16).

It was also hypothesised that there will be a positive correlation between the self-report method and direct measurement method for determining cognitive load. The correlation between the two methods was found to be positive, but very low ($r = 0.07$) and not statistically significant.

*The hypothesis that there will be a positive correlation between the self-report method and direct measurement method for determining cognitive load was **not** supported.*

If the finding of Whelan (2006) is used to explain this result then this low correlation is not surprising as the two methods are not equally sensitive to differences in cognitive load. There is also the possibility that the two methods are not measuring the same type of cognitive load: intrinsic, extraneous and germane load.

5.4.4 Presentation formats and their influence on cognitive load

The question:

To what extent do the presentation formats influence cognitive load?

This reflection looks at the cognitive loads of the individual screens, both within a particular version and between versions, and considers the design differences between the screens. How did the

design of the screens influence cognitive load, and what design strategy might have been responsible for the differences in cognitive load?

5.4.4.1 Comparing screen 12 and screen 19 in the animation version

Both screen 12 and 19 consisted of a narrated animation only. Screen 12 had a statistically significantly higher cognitive load than screen 19, although the effect size was in the small to medium range ($d = 0.31$). This was unexpected, since screen 12 included content that had already been covered in the curriculum, while screen 19 covered new content.

Factors that could have influenced this difference in cognitive load between screen 12 and 19 include the duration of the animation and the transitory nature of the animation. There were a number of different visual views in each animation: parts of the image changed (some of these changes were quite subtle), there were moving elements (excluding text animation), text labels appeared and disappeared, some of the text labels were highlighted and the highlighting technique often moved to another label. Even though narration was added to reduce the processing in visual memory and therefore decrease the extraneous cognitive load, there were still enough visual elements changing to induce extraneous load, over and above the intrinsic load of the content.

The number of changes in the visual views (or scenes) were counted for each of the animations. In screen 12 there were approximately 44 scene changes across the 1 minute 45 seconds. Some of these changes occurred in rapid succession, so close that at times it was difficult to count the changes. In screen 19, on the other hand, there were approximately 30 scene changes across the 1 minute 15 secs. On closer examination of the scene changes in screen 19 it appeared that the time intervals between these scene changes was in most cases longer than those in screen 12. It was at least possible to keep up with the count. While the count is quantitative in nature, the description of the scene changes and reflection on the techniques used is a qualitative observation, but it has provided some valuable insight into the design and subsequent use of animation by learners. Proportionally, the number of scene changes were equal for the two animations, but looking at the design from a cognitive load perspective it seems that these fast scene changes increased the extraneous cognitive load of the animation. The cognitive load experienced by a learner using an animation therefore comes from several sources: the need to search for relevant information within each visual view (extraneous load), the transitory nature of the animation requiring the learner to hold information in memory and process new information (extraneous load) (Ayres & Paas, 2007) and the difficulty of the content (intrinsic load).

However, since the mean cognitive load for screen 12 was within the medium load range ($M = 5.76$) it can be concluded that not all the design techniques were ineffective. Designers do need to be careful not to reduce the cognitive load too much. The study conducted by Schnotz and Rasch (2005) found that some of the animation used in their study unnecessarily reduced the germane cognitive load

associated with deeper more meaningful cognitive processing. In the end, the participants did not invest enough mental effort in order to improve their learning performance sufficiently.

The animation in screen 12 was only 30 seconds longer than the animation in screen 19. Further investigation is necessary to determine the point at which time duration becomes a factor in increased cognitive load of an animation.

5.4.4.2 Comparing screen 5 and screen 13 in the animation version

Screen 5 had the higher mean cognitive load. There were several design issues on screens 5 and 13 that could have influenced the cognitive load. This section briefly discusses each of these issues.

The first issue relates to the images on the screen. Screen 5 and Screen 13 each required the participant to click on a textual hotspot in order to change the image. Participants were required to compare the images presented in each view. Screen 5, illustrated in Figures 5.1 – Figure 5.3, required the learner to toggle between four different images in contrast to the two views of screen 13 (illustrated in Appendix H).

The differences between the images in both screen 5 and 13 were subtle, except for the image illustrated in Figure 5.3. These subtle differences will increase the intrinsic load of the content. When considering the images only, screen 5 could be expected to have a higher cognitive load than screen 13, simply because there were more images to hold in working memory while making the cognitive comparison.

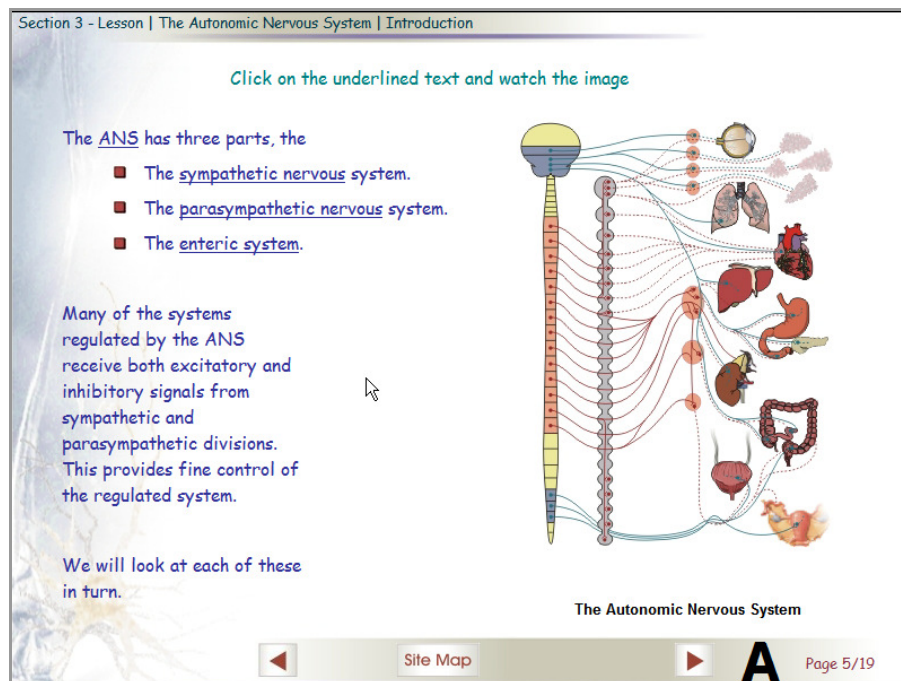


Figure 5.1: Screen 5 – First image

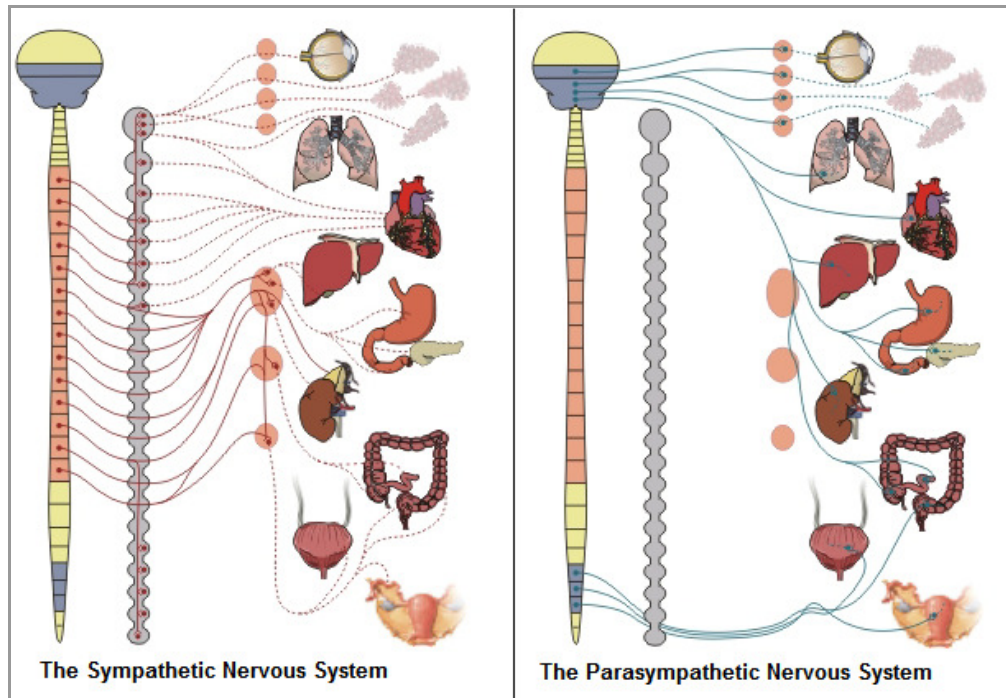


Figure 5.2: Screen 5 – Second and third image

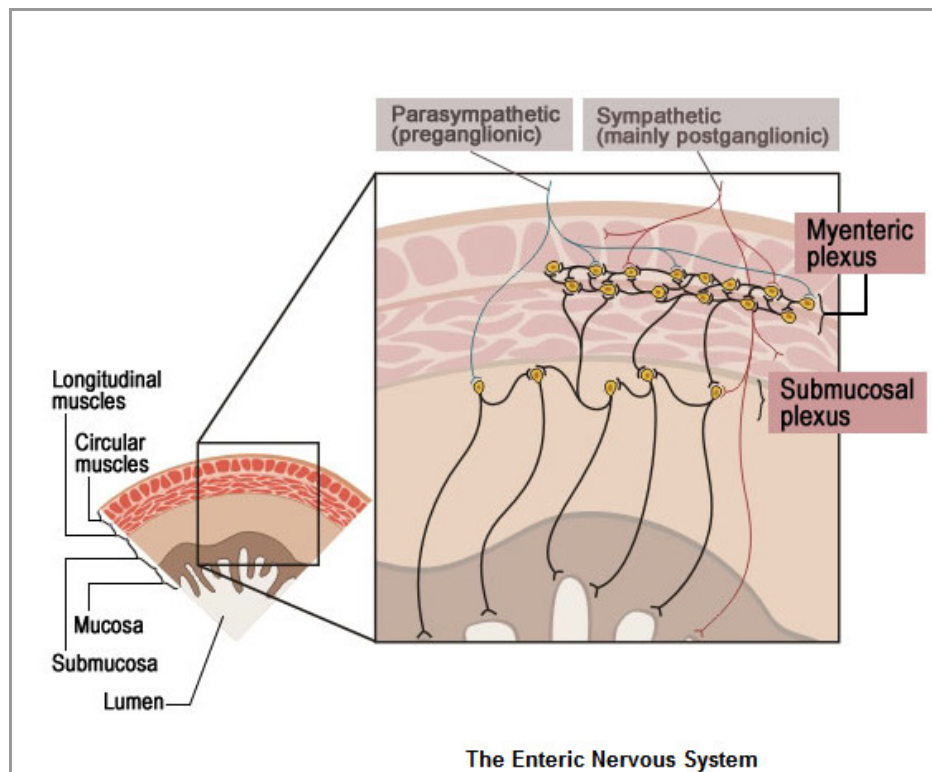


Figure 5.3: Screen 5 – Fourth image

The second issue relates to the design of the image itself. Carney and Levin (2002) reviewed recent research on the use of text and pictures in learning materials. Their review and the ten guidelines

they propose for designing text-picture combinations that will facilitate learning led me to the conclusion that at least one of these guidelines was ignored in the images in screen 5, while this was not the case for the image on screen 13. There was poor congruency between the text and images in screen 5, leading to increased cognitive load for screen 5. In fact, the text on screen 5 did not even relate to the images. The images in screen 5, with the exception of the image illustrated in Figure 5.3, were not labeled, while the images in screen 13, although not labeled, had explanatory text that complemented the image.

The third design issue concerns the non-visual content on these screens. Screen 13 included textual information **and** an image, whereas Screen 5 only changed the image. It could be expected that the additional text in the views for screen 13 had the potential to create a split-attention effect, which would increase the extraneous cognitive load of screen 13. However, the contiguity principle was deliberately applied to offset a split-attention effect for screen 13. Since screen 5 had a significantly higher cognitive load than screen 13 ($p = 0.0008$) it suggests that the intrinsic cognitive load contributed more to the total cognitive load than did the extraneous cognitive load embedded in the click actions to view and hide the content. The role that the contiguity effect plays in reducing the potential for a split attention effect on a screen similar to screen 13, where images must be compared, remains an area for further investigation.

5.4.4.3 Comparing screen 14-16 and screens 23 in the static images & text version

The cognitive load of these four screens (13 - 16) was not significantly higher than the cognitive load of the animation of screen 23 ($M = 6.14$). The animation in screen 23, with its 30 'scene' changes (described in Section 5.4.4.1) still resulted in a higher cognitive load than the load experienced across four separate screens of text and images. Chunking of the content, which could be viewed as a manipulation of the intrinsic load, seemed to lower the cognitive load.

5.4.4.4 Comparing screen 20 and screen 13 in the static images & text version

When comparing screen 20 and screen 13 (illustrated in Figure 5.4), where there was a significant difference in the cognitive load (screen 20 had the higher cognitive load) we see another difference in design technique. On entering screen 13 the user can read the text and study the image, which facilitates the understanding of the text. The contiguity principle (Mayer & Moreno, 2003) was applied in this design, in order to keep the extraneous load as low as possible.

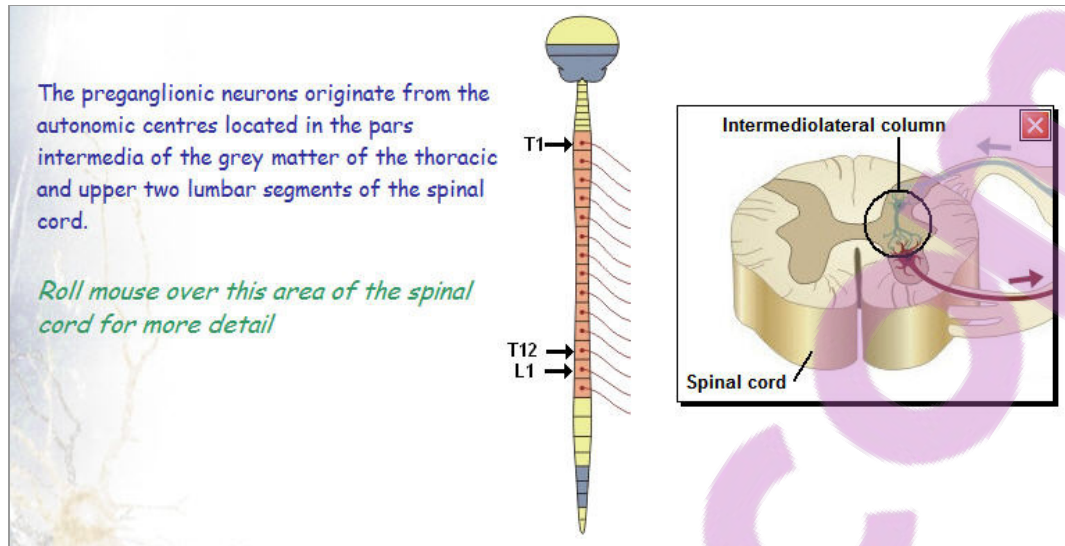


Figure 5.4: Section of screen 13 in Version 2

The initial image is relevant to the last part of the sentence on the screen. Using a mouse-over technique, the user is able to study a second image, which is a cross-section of the spinal cord. The aim of using this image is to provide a better visual in order to understand the text. The image displayed in the mouse-over facilitates an understanding of the first half of the text on the screen. Unlike many pop-ups, which are displayed using the mouse-over technique and which disappear when the user moves the mouse off the hot spot of the image, in this program the user had to click on the icon in the pop-up in order to hide this second image. Although the user must divide his/her attention between the text and the images the load is predicted to be low as there is very little text and the diagrams are not complex. Carney & Levin (2002) suggest that the use of pop-ups is a more effective technique to use when displaying images that need to be used with text. Labeling of the images was kept to a minimum. The user does not have to hold any information in working memory while searching for either the image or the text, due to the application of the contiguity principle. The amount of user control provided also allowed the user to spend as much time as needed to review both the text and the images.

On screen 20 (illustrated in Figure 5.5) the user must compare two concepts. Each concept is explained using both text and a visual image. The user was required to click on the text on the screen (Divergence or Convergence) in order to display the relevant text and image. The aim of this design technique was to facilitate comparison of the images, without the user needing to branch back and forth between screens. The content for the one concept must be held in working memory while the second concept is being studied before any cognitive comparison can take place.

Section 3 - Lesson | The Autonomic Nervous System | Divergence and Convergence

Divergence and convergence

There are two phenomena in the sympathetic nervous system, known respectively as:

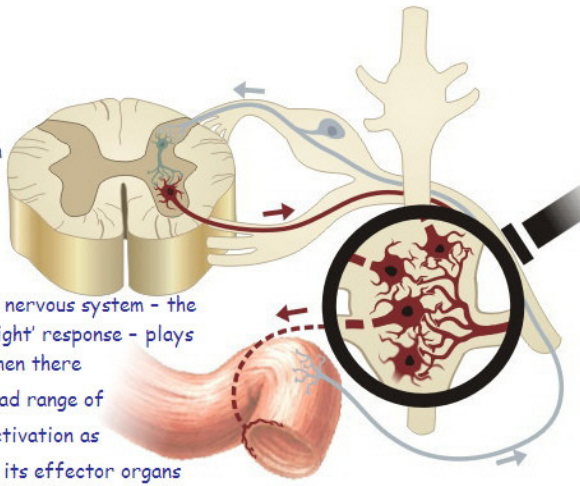
- Divergence
- Convergence

<<<< Click on each concept for an illustrated explanation

Divergence

This is the phenomenon where a single preganglionic neuron synapses with an average of ten postganglionic neurons in the sympathetic ganglion. This **SIMULTANEOUSLY** activates more organ systems. Divergence occurs primarily in the SNS.

This mass activation of the sympathetic nervous system - the so-called mass discharge or 'fight-or-flight' response - plays an important role in stress responses when there is a need to coordinate changes in a broad range of systems. Because of mass activation (activation as a single unit) the SNS can affect all of its effector organs at the same time.



A Page 20/23

Figure 5.5: Section of screen 20 in Version 2

Considerably less text was displayed for the concept 'Convergence'. A comparison of Figure 5.4 and Figure 5.5 clearly shows that the volume of the content was greater for screen 20 than for screen 13. An alternative design for screen 20 might be to place both images next to each other on the screen, but in order to adhere to the contiguity principle the images would have to be very small. Would reducing the size of the image interfere with learning? Further investigation would need to be done in order answer to this question.

5.4.4.5 Comparing screen 5 and screens 17-19 in the static images & text version

The design of these two screens is discussed since the findings were unexpected. It was expected that the cognitive load of these screens would be significantly different. Screens 17-19 had a higher cognitive load than screen 5 but the difference was neither statistically nor practically significant. Interestingly, the p value before the application of the Bonferroni correction was statistically significant. The effect size does not change however, providing support for the value of the effect size in interpreting findings.

There were two differences between these screens: one difference focuses on the content, where screens 17–19 had considerably more content than screen 5. The second difference relates to the design, specifically regarding the user action required to access the content. In screen 5 the user had a one-click sequence to view the relevant content: click on text → view image. There were three new

images to view. Each image illustrated the structure of one of the three sub-systems of the ANS. There are subtle differences between the SNS and PNS, but the Enteric system is obviously different. This design was used to allow the user to compare the structure of the sub-systems visually. In screen 18 a similar design approach was used, but the user had at least a two-click sequence in order to view the relevant content: click on text/image → view baseline image → click on one of three buttons to see changes in image. The changes allowed the user to compare the effects of SNS and PNS innervation on the specific organ. The innervation of 5 different organs was presented on this one screen. The comparison was not between the organs, but between the effects of the SNS or PNS on the individual organ. One of the views of screen 18 is illustrated in Figure 5.6.

Although only one image was present on screen 18 at a time (the smaller images functioned as icons), there were a total of 15 different images, with three views for each organ. In both screen 5 and 18 the participant had to hold one or two images in working memory while looking at a second or third image in order to make the comparison. In screen 18 the image of each organ had three different states, which changed dynamically depending on the button clicked to initiate the animation, while in screen 5 there were three **different** images. It may be that the extraneous load imposed by the need to click several times while studying the content imposed a higher load than just clicking on text and being able to view the image without worrying about clicking on another button to change the view.

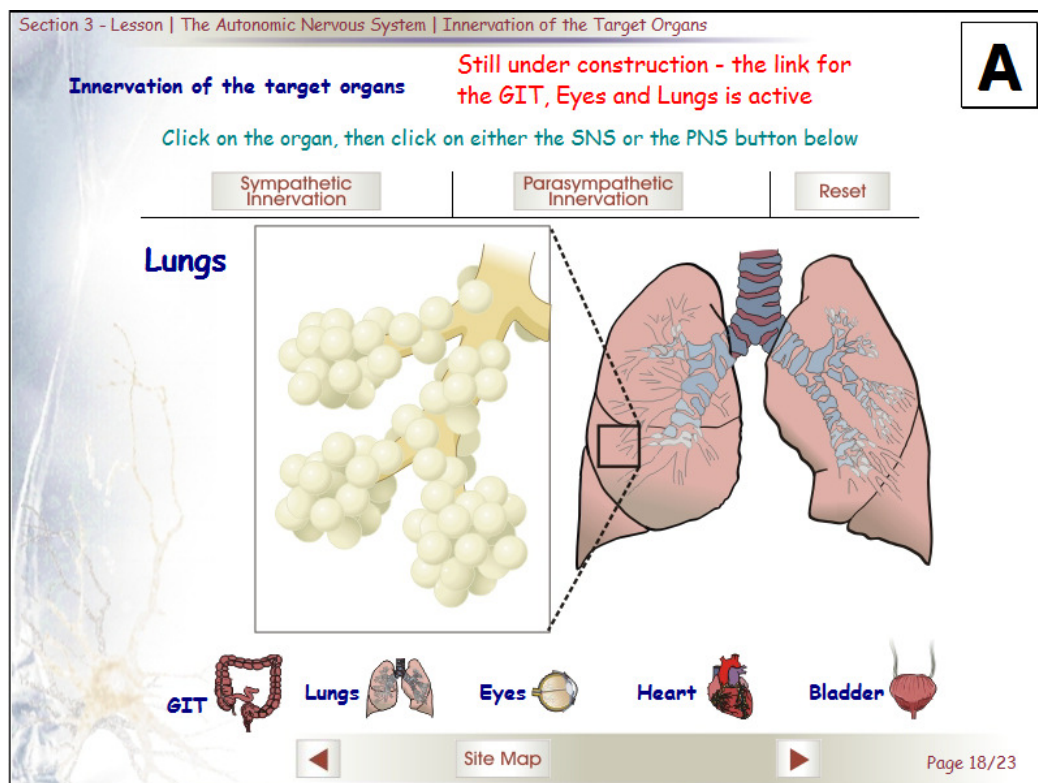


Figure 5.6: View from screen 18 in the static images & text version

In spite of this there was not a significant or practical difference in the cognitive load of these screen groups. Design issues that could account for this include the effective use of learner control, the careful use of text and images in combination with each other and the attention that was paid to the contiguity effect.

5.4.4.6 In summary

A consistent finding was that screens with animation had a higher cognitive load than the screens with static images & text. This finding continues to support the arguments in the literature that animation requires far deeper cognitive processing and requires more cognitive resources from working memory than do static images and text (Chan & Black, 2005; Lowe, 2004).

The presentation formats of the screens that had statistically significant **and** practically significant differences in the cognitive load were discussed. It would appear that the following factors could have played a role in the cognitive load:

- Length of the program and time needed to work through the entire program
- Length of the individual animations
- Number of 'scene' changes in an animation
- Amount or density of the content
- User actions required to navigate through the screens: non-linear navigation appears to impose higher extraneous load.

5.4.5 Cognitive load, cognitive style and learning performance

The question:

How is learning performance influenced when content with different cognitive load is studied by learners with different cognitive styles?

There was no evidence in the literature of attempts to consider the relationship between cognitive style, cognitive load and learning. The reasoning and argument behind the hypotheses of the relationship between style, load and learning were presented in detail in Section 3.16.2 in Chapter 3.

5.4.5.1 Learning performance in general

Due to the fact that this study was conducted in an authentic learning environment it was important to determine whether there was a learning gain in general, irrespective of cognitive load, cognitive style, time spent on the program and other variables that could influence the outcome. Each participant did a pretest at the start of the experiment and a posttest after studying the content of one of the two versions.

There was a difference in the mean values of the computer-based pre- and posttest and a univariate analysis, which compared the difference between the pretest and posttest results, indicated that this difference was statistically significant. Learning did take place and from an ethical point of view it can be concluded that no participant was disadvantaged by the version they used. Whether this learning was sustained over a period of time was not determined as part of this study.

5.4.5.2 A look at performance in the pretest

Performance in the pretest was categorised into low, average and high. There was no relationship between the pretest results and grouping by version. The groups could therefore be regarded as equal in terms of prior knowledge as measured by the pretest. Two further analyses were conducted using the pretest results.

The first analysis focused on the question: How did prior knowledge, determined by asking whether the participants had studied the topic previously, compare with the pretest results? The expectation was that proportionally more participants who had **not** studied the topic previously would fall in the 'Low' performance group. Similarly, proportionally more participants who **had** studied the topic previously would fall in the 'High' performance group. The results indicated that this was not the case and it was concluded that there is no difference, for all three levels of performance, between the group who answered 'Yes' to this question and the group who answered 'No'.

The second analysis asked: How did prior knowledge, determined by the subjective self-report rating of the knowledge about the topic, compare with the pretest results? The expectation was that proportionally more participants who indicated that they knew nothing about the topic would fall in the 'Low' performance group. Similarly, proportionally more participants who indicated that they had more than a basic understanding of the content would fall in the 'High' performance group. The results supported this expectation. The majority of the participants in each of the 'Rating of knowledge' groups scored in the average range of performance.

The results in this study suggest that self-report ratings of knowledge are not necessarily reliable predictors of prior knowledge. I concluded that actual pretest results are a better criterion to use than self-report measures for determining prior knowledge.

5.4.5.3 Open-ended posttest scores

The posttest included a pencil and paper test at the end of the experiment. Two open-ended questions, testing application of knowledge, were given to the participants (see Appendix E). Many of the participants appeared to put far more effort into Question 1 than Question 2, scoring a relatively good mark for Question 1, but failing (less than 40%) Question 2. The possibility exists that they did not understand the question. In retrospect it was a difficult question, although it had been reviewed by the faculty member responsible for teaching this 2nd year course. It is possible that the sample was truly not able to apply the knowledge they had just learned. None of the participants had had any

exposure to the specific clinical situation used in the question. Experience of the condition used in Question 2 would be limited to theoretical knowledge.

Some of the reasons for the low mean for Question 2, not directly related to performance, knowledge and learning, but which can impact this, could include the following:

- Participants were under considerable time pressure at this stage of the study. This test was taken at the end of the session and many of them had a class to attend.
- Fatigue could also have contributed to the low score and they merely scribbled down an answer in order to get it over and done with.
- Lack of interest or commitment to answer to question.

None of these reasons were explored with the sample, as it was not possible to contact them again after the experimental session.

5.4.5.4 Learning performance and cognitive style

General Linear Modeling analysis was conducted to determine if cognitive style had any influence on learning performance. The two style dimensions were considered separately. The pretest was entered as a covariate in this analysis, and in both analyses it accounted for the significant finding. Neither the Wholistic-Analytic or the Verbaliser-Imager style appeared to have any influence on learning performance in the posttest.

The findings of my study support those of two studies and appear to contradict the results of two recent studies.

Calcaterra, Antonietti & Underwood (2005) investigated the effect of the Analytic-Sequential versus Holistic-Intuitive style and hypermedia navigation behaviours on learning outcomes. This study, using undergraduate students, concluded that performance outcomes were related to particular search and navigation patterns and not to the time spent studying the content or the particular cognitive style of the learner. Massa and Mayer (2006) investigated the types of help given to participants categorised as Verbalisers and Visualisers. Using college and non-college students they found that the behaviour of the participants was consistent for their self-reported style (verbalisers tended to reply on textual help and visualisers tend to rely on pictorial help). Verbalisers and visualisers did not differ on the learning test and the researchers concluded that there was no evidence that verbalisers and visualisers should be given different multimedia instructions.

On the other hand Lee (2007) explored the effects of visual metaphor and cognitive style in a hypermedia-based environment and found that the participant's cognitive styles substantially affected learning performance. The Analytic-Verbalisers performed significantly better than the Wholistic-Imagers and Wholistic-Verbalisers and the largest difference was between the Analytic-Verbalisers and the Wholistic-Imagers. The cognitive style literature (Evans, 2004; Riding et al., 2003; Riding &

Grimley, 1999) describes the Wholistic-Analytic and Verbaliser-Imager style combinations as either unitary or complimentary and in Lee's study we see that the unitary style group outperformed the complementary style group. Grimley (2007), in a study that was very similar to my study, found a main effect for the Wholistic-Analytic style on the overall recall of knowledge. Wholists tended to answer more questions correctly than Analytics. Grimley also found gender effects in this study (all style groupings except Wholistic-Verbalisers showed gender differences) and went into some depth in exploring these gender differences. Grimley (2007) however did not measure the cognitive load of the two interventions (text and pictures and narration and pictures) used in his study.

5.4.5.5 Learning performance and cognitive load

General Linear Modeling analysis was conducted, with the posttest as the dependent variable and the cognitive load and version as the predictor variables. The pretest results were entered as a covariate. The analysis indicated that cognitive load, as measured using the self-reporting technique did not appear to influence the learning performance in this study in any way.

Tabbers et al. (2004) measured mental effort during instruction and testing in an authentic learning environment. The mental effort during instruction was found to be relatively low. Learners who used the visual mode reported higher mental effort than those who used the audio mode, but the difference was not significant. Tabbers et al. (2004) did not report on any analysis of the relationship between mental load **during** instruction and the posttest results. They did report on the analysis of the relationship between the mental effort scores during testing and the subsequent test results. They found that there were significant differences in the mental effort reported for the retention test (learners in the visual condition reported more effort than students in the audio condition) but not for the transfer test. One of the conclusions these researchers discussed was that the authentic environment could have introduced confounding variables that influenced the outcome of the study.

Höffler (2007) investigated learning from instructional animation or a series of static pictures and found that spatial ability accounted for the difference in learning outcome rather than cognitive load, which was not significantly different for the two versions. Höffler, Sumfleth and Leutner (2006) found a strong correlation between spatial ability and learning outcomes in the group that used static pictures, but the correlation between these variables was weak for the group that used animation. Using general linear modeling and controlling for grade-point average they found that the interaction between spatial ability and type of learning material was statistically significant, and this did not change when cognitive load was entered into the linear model.

Another possible explanation for the finding in my study is that since the cognitive load for each version was within an acceptable range it did not influence learning performance negatively. A different finding might have been obtained if the cognitive load for one version was in the low range and the other in the very high range.

5.4.5.6 Cognitive style, cognitive load and learning performance

The relationship and dynamic between style and load appeared to be more complex than initially thought, influenced by different factors that were difficult or even impossible to control in the authentic learning environment. With this in mind, the final set of hypotheses must be viewed as a first attempt to investigate this complex relationship.

We have seen in the preceding sections that neither cognitive style or cognitive load seemed to influence learning performance in this study. Another variable was then introduced into the analysis: the time used by the participants. Early cognitive load studies considered time to be an indicator of cognitive load (Chandler & Sweller, 1991). The results will be discussed briefly before turning to the discussion of the relationship between cognitive style, cognitive load and learning performance.

Two hypotheses were investigated:

- Analytic learners who spend less time on the program will rate the cognitive load more highly.
- Verbalisers and Imagers who spend less time on the program will rate the cognitive load more highly.

The results of a regression analysis used to investigate the hypotheses indicated that cognitive style did not influence the cognitive load. The interaction between version and time spent on the program was retained in the regression, but it only approached significance. A confirmatory GLM (General Linear Model) was conducted and it was established that the only predictor of cognitive load in the model tested was rating of knowledge about the topic. The participants who indicated that they knew nothing about the topic rated the cognitive load significantly higher than the rest of the sample.

*The hypothesis that Analytic learners who spend less time on the program will rate the cognitive load more highly was therefore **not** supported*

*The hypothesis that Verbalisers and Imagers who spend less time on the program will rate the cognitive load more highly was therefore **not** supported.*

The final two hypotheses investigated in this study were:

- The Analytic learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest.
- The Verbaliser and Imager learner who spends inadequate time on the program, and who rated the cognitive load as high, will perform more poorly on the posttest.

The analyses looked at the posttest results for the Analytic, Verbaliser and Imager learners, grouped by time spent on the program and rating of cognitive load. Significance testing was difficult as the sample size was small. The data was organised and analysed (using descriptive statistics) in order to determine if a trend could be observed in the data. There was a trend that suggested that the higher the load the poorer the learning performance. There was no clear trend for the amount of time spent on the program.

This trend was confirmed in a subsequent regression analysis where version of the multimedia, cognitive style (both dimensions), subjective rating of cognitive load and total time spent on the lesson were included in the model. The variables retained in the stepwise regression were version and subjective rating of cognitive load. There was a statistically significant main effect for cognitive load and a marginally significant main effect for version. The confirmatory GLM (General linear model) confirmed that the subjective rating of cognitive load accounted for the significant result. The posttest scores were higher in the animation version compared to the static images & text version. This study also determined that the static images & text version had a higher cognitive load than the animation version. I was therefore able to confirm the trend that within an authentic learning environment, where the cognitive load of the entire program was considered, the higher the cognitive load the poorer the learning performance.

The only study found in the literature since February 2006 that is similar to my study (investigates cognitive style and cognitive load in multimedia learning) was that of Grimley (2007). Grimley compared the performance of the cognitive style groups from each dimension, using two different multimedia designs, each with different cognitive loads. The cognitive loads were not specifically measured, but were assumed to be different due to the use of a split-attention effect in the one version. Grimley (2007) argued that, in terms of overall performance, the multimedia material used in his study seemed to suit the Wholistic learner better. Grimley ascribed this to the ability of the Wholistic learner to see the big picture, which required an understanding of both the images and text. Grimley also used gender to explain the differences in the results and this made precise comparison with my study difficult.

The relationship between cognitive load, cognitive style **and** learning performance in an authentic learning environment is less clear than the relationship between cognitive load and style, without the consideration of performance. It is likely that different results will be obtained in a more controlled experimental setting (a dedicated research laboratory) using materials that have larger differences in cognitive load.

5.5 Scientific reflection

This final section of Chapter 5 focuses on four topics:

- Contribution of this study to the scientific body of knowledge in the cognitive style, cognitive load and multimedia learning fields
- Implications of this study for instructional design
- Implications of this study for using multimedia in the learning environment
- Recommendations for future research

5.5.1 Contribution of this study to the body of knowledge

I believe that the most significant contribution of this study lies in the fact that it is one of the first attempts to empirically explore the relationship between cognitive style and cognitive load. The cognitive load theory was discussed in depth and illustrated in Chapter 2. According to cognitive load theory, cognitive style is one of the subject characteristics thought to influence mental effort and performance. Reports of mental effort invested by an individual, together with their subsequent performance, provides an indication of the cognitive load of learning material.

A second unique contribution relates to the context selected for this study. Empirical studies that address either cognitive load or cognitive style have been under-researched in the health science education context. There are a few more recent publications that address the theory of cognitive load (Khalil, Paas, Johnson & Payer, 2005a, 2005b). My study has thoroughly investigated the role of both cognitive load and cognitive style **and** the relationship between these two factors in a health science education context.

A third unique contribution lies in the fact that this study has addressed the call for research in more authentic learning environments. Such environments are complex and the control of variables are difficult. Authentic environments are not usually the first choice for the context of cognitive load and/or cognitive style research. I was able to undertake a rigorous quantitative study in an authentic environment.

A fourth unique contribution is that my study is one of the few studies that compares one of the most widely used techniques to measure cognitive load (subjective rating of mental effort) with a technique that is not widely used, but which is thought to be a more accurate method of determining cognitive load, since it measures cognitive load at the time at which it occurs (direct method using dual-task methodology).

I therefore present the following theses:

1 While health education learners appear to be more Analytic in the way they process information, and while they display more of an Imager style in the way they represent information during thinking, this profile is not yet conclusive for the health education field, due to the lack of empirical studies with which to compare the findings of this study.

2 It cannot be assumed that there are no cultural differences in the way learners process information during multimedia learning, as this study found a significant difference in the Analytic style of learners from different cultural groups.

3 Cognitive style does not appear to make any difference to the amount of time needed to study from animations. There was no significant difference in the amount of time the Wholistic and Analytic learners spent studying the animation, and there was no significant difference in the amount of time Verbaliser and Imager learners spent studying the animation.

Use of animation will most likely constrain the scanning strategy of the Wholistic learner as he/she would have to watch the whole animation at least once in order to understand it.

4 Cognitive style does make a difference to the amount of time needed to study multimedia material that uses static images and text. The Analytic learners spent significantly more time than the Wholistic learner studying the content of program. Since the Analytic learner is not constrained by the temporal nature of the animation he/she can (and therefore does) take all the time he/she needs to process the content. The Verbalisers appeared to need more time than Imagers, but the difference was not significant.

5 The static images & text version of the program as a whole, used in an authentic learning environment, had a significantly higher cognitive load than the animation version. Since the learner must infer the processes explained in the static images & text rather than merely perceiving them in an animation, more mental effort is required to understand and learn the content in the static images & text version. The non-linear nature of the static images & text version and the total length of the program are also thought to have increased the extraneous load of the program as a whole.

6 Animation as an isolated learning event requires more cognitive resources from working memory than do screens presenting the same information using static images and text. This resulted in a higher self-report rating of mental effort invested for the animation screen. This higher cognitive load for the animation was a consistent finding whenever an animation screen was compared with a non-animation screen using static images & text.

- 7** If the extraneous cognitive load can be controlled and minimised the Analytic learner will still report a higher cognitive load than the Wholistic learner, due to the more elaborate processing strategies used by the Analytic learner, which implies deeper processing and a higher germane cognitive load.
-
- 8** Measurements of cognitive load, using the same content, but different measurement techniques were not correlated. The possibility exists that each measurement technique focuses more strongly on a different type of cognitive load.
-
- 9** The more mental images the learner is required to compare, which are not presented contiguously, the higher the rating of mental effort.
-
- 10** If cognitive load is kept low as possible learning will take place irrespective of the learning strategy used. This study showed that significant learning took place in both versions of the program.
-
- 11** There is no conclusive evidence that animation is a better instructional strategy for learning in health science education. The use of animation must be carefully considered in order to justify the cost and time of such development.
-
- 12** There is no clear and simple relationship between cognitive load, cognitive style and learning performance in an authentic learning environment in health science education. The relationships are complex and require further investigation.

5.5.2 Implications of this study for instructional design

This study confirms the validity of existing instructional design guidelines for multimedia learning material. These include the practice of controlling for the split attention effect and paying attention to the contiguity principle. I would like to suggest that instructional designers should **first** control for the split-attention effect and **then** consider the contiguity principle. Design strategies that include the use of pop-ups can pay attention to both of these effects. Using a small pop-up that the user can drag and position in such a way that any underlying text or image can still be viewed, rather than placing the information on two separate screens, will address both the split-attention and contiguity effects. The extra click action required to open and close pop-ups does not seem to influence extraneous cognitive load adversely.

My study demonstrated clearly that the learners had significant gains in learning irrespective of the version used. Animation is not necessarily a superior strategy. The development of animation is time-consuming and costly. The use of animation must be carefully considered when there are time and

funding constraints. This study did not provide enough convincing evidence that animation was more beneficial in teaching the topic 'The Autonomic Nervous System'.

There is still much to learn about the effective design of animation. When designing animation, attention needs to be paid to the speed at which the views or scenes change. It appeared as if the views in some of the animations used in this program changed too fast. Learners did have control over the pace of the animation, but few learners were observed as actually using the controls. The designer could consider introducing more learner control at specific points in the animation, for example where there is a change in focus of the content. Learners need to be specifically instructed in the use of control features. An alternative is to develop several shorter animations that are viewed in a specific sequence.

5.5.3 Implications of this study for using multimedia in the learning environment

A finding in this study was that an unacceptably high percentage of participants spent inadequate time on the multimedia programs. Some of these participants seemed to merely click through the program, spending less than 1 minute on a screen. The amount of learning that takes place in such a situation must be queried. While this study did not follow these participants up in order to obtain some qualitative data about the reasons for this pattern of use, it is clear that more guidance needs to be provided to learners on how to use multimedia programs when learning. Note-taking while they are using the multimedia program could be encouraged. Learners need to be advised on how much time they should spend on each screen. It was noted that some of the learners clicked very quickly through the entire program before going back to review each screen in more detail. This might be related to cognitive style and a qualitative follow-up of multimedia use might provide a better perspective on this pattern.

5.5.4 Recommendations for future research

I have identified various avenues for future research during the course of this study. The sources of these recommendations include:

- Limitations in the research methodology of my study
- New literature that appears to be aligned with my study
- Findings in my study that suggest other avenues for inquiry

I have divided the recommendations into three areas, although there may be overlap and some of the recommendations can be combined in a single study.

5.5.4.1 Exploring the relationships between cognitive load, cognitive style and multimedia learning

Wholistic learners are unlikely to invoke germane load to the same extent that Analytic learners with their elaborate processing style do. Therefore, if the extraneous load is low the Analytic learner should perform better than the Wholistic learner, since the deeper processing (germane load) should result in better performance. If the extraneous load is high we can expect to see the Analytic learner perform more poorly compared to the Wholistic learner, as their style will still result in some germane load and the total cognitive load might reach a level where it is detrimental to learning. In order to test these predictions I recommend moving the research back into a more controlled experimental setting, using shorter interventions that have large differences in cognitive load.

I would also recommend that future research controls the amount of time the learner is given to study the content, as time seemed to play some role in this study.

Since it appears as if the Wholistic learner will not process as deeply as the Analytic learner future research could investigate and compare strategies that can assist the Wholistic learner with deeper processing.

The relationship between cognitive load, cognitive style and learning performance in multimedia learning needs further investigation in a variety of disciplines, with a variety of different learner groups. Further studies could replicate the current design and/or improve on the design, using the suggestions made in Section 5.3 of this chapter.

5.5.4.2 Instructional design practice

The participants in this study were allowed to make notes while using the multimedia program. These notes were only scanned to get a high level overview of the nature of the note-taking. A detailed analysis of these notes was beyond the scope of this study. Detailed analysis of the note-taking practices of learners using multimedia is an avenue for further research. Of particular interest would be to determine the extent to which different style groups approach note-taking, and how note-taking influences the subjective ratings of mental effort. Electronic and paper-based note-taking could also be compared.

The role that the contiguity effect plays in reducing the potential for a split attention effect, and subsequent cognitive load, on screens where comparison of images is required also needs further investigation. An alternative design could be considered where both images are placed on the same screen, but in order to adhere to the contiguity principle the images would have to be very small. Would this interfere with learning? Would this influence the cognitive load of the instruction? Further investigation would need to be done in order answer to these questions.

One screen in this program provided an overview of the content and it functioned as a menu to facilitate non-linear navigation. More detailed analysis of the path of navigation by each participant after they had accessed this overview/menu screen was beyond the scope of this study, but could provide an avenue for future research that considers how best to design such screens. The navigation path could also provide an indication of which sections of the content were accessed more often, and this analysis could be followed up using both a qualitative and quantitative approach.

5.5.4.3 Cognitive load research

This study only measured the cognitive load the first time the learner accessed the screen. The learners were however allowed to move back and forth through the program. If the cognitive load had been measured with each entry into the program it would have been possible to determine if the rating of mental effort invested decreased as they became more familiar with the content.

Further investigation is necessary to determine the point at which time duration becomes a factor in the cognitive load of an animation.

The one change I will recommend in future research is that the self-report rating of cognitive load only require the learner to consider one screen per rating. The fact that the participant had to keep two screens, and even three, in memory might have inadvertently influenced the cognitive load rating.

5.5.4.4 Cognitive style research

Further investigation is necessary into the cultural differences in cognitive style. This research should focus specifically on environments where many different cultures co-exist and co-learn, rather than looking at differences between cultures separated by distance (Graff, Davies & McNorton, 2004).

The two style dimensions (WA and VI styles) could be combined to reflect four different possible groupings (Wholistic-Verbaliser, Wholistic-Imager, Analytic-Verbaliser and Analytic-Imager) as each learner has a particular style on each of the dimensions. This would allow the investigation of the impact of unitary and complementary cognitive styles on the cognitive load of instructional material.

5.5.4.5 Use of multimedia material for learning

The patterns of navigation in the program were only explored briefly. Reasons for viewing each screen only once, or several times, which could be explored in future research, include attitudes toward the use of a computer to study, personal issues that influence the time participants wanted to spend with the program, fatigue, motivation and interest in topic and general study habits and learning style.

5.6 Conclusion

This study explored the role that cognitive load and cognitive style play in multimedia learning from a number of different perspectives: time spent on the program, comparison of the cognitive load for a particular style across the two versions (illustrated in Figure 5.7 using the yellow arrow), comparison of the cognitive load of a particular version across the two styles (illustrated in Figure 5.7 using the blue arrow), results of learning performance.

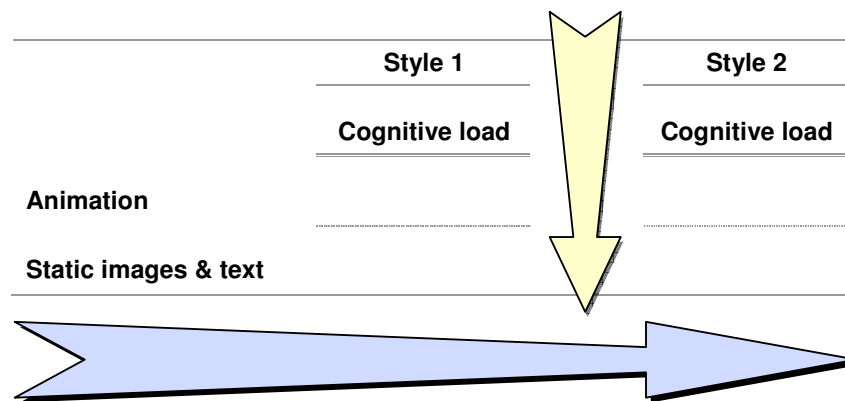


Figure 5.7: Perspectives of the study

The strategies of animation and static images & text were compared within an authentic learning environment. Since the environment was authentic it was important to establish whether learning actually took place.

The cognitive style profile of the research sample leaned towards the Analytic style in the Wholistic-Analytic dimension and toward the Imager style in the Verbaliser-Imager dimension. It has yet to be determined whether this is a typical cognitive style profile for health education learners. The cognitive loads of the respective research interventions were significantly different, yet neither version appeared to have an excessive cognitive load that negatively influenced learning. Significant learning took place for all the participants in this study. Surprisingly, it was found that when the program was considered as a whole, the version that used predominantly animation had the lower cognitive load as measured using the subjective rating technique. When the analysis drilled down to specific screens and compared animation and static images & text, the results consistently showed that animation had a higher cognitive load than static images & text.

This study established that there is once again empirical evidence that cognitive load influences learning performance. It was possible to establish that it seems as if the Analytic cognitive style influences the subjective rating of cognitive load, but further empirical investigation of this relationship is necessary. The proposal is that the Analytic style influences the germane load experienced during learning. Since researchers are currently unable to measure the three different types of load separately, this proposal remains an area for further investigation.

The subjective cognitive load rating of the program was compared with the cognitive load rating measured using the direct measurement method. The direct measurement method (also applied in an authentic learning environment) found that the animation version had the higher cognitive load. The correlation between these two methods of measurement was very low and not significant, thereby confirming a suggestion in recent literature that each method might be measuring different aspects of cognitive load.

A final comment on my experience as a researcher is adequately summed up in the quote by Lloyd Alexander.

We learn more by looking for the answer to a question and not finding it than we do from learning the answer itself.'

(Lloyd Alexander, n.d).

References

- Abraham, R.R., Kamath, A., Upadhy, S. & Ramnarayan, K. (2006). Learning approaches to physiology of undergraduates in an Indian Medical School. *Medical Education*, 40(9), 916–923.
- Abouserie, R. & Moss, D. (1992). Cognitive Style, Gender, Attitude Toward Computer-Assisted Learning and Academic Achievement. *Educational Studies*, 18(2), 151-160.
- Ainsworth, S. & Van Labeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, 14(3), 241–255.
- Alessi, S.M. & Trollip, S. (2001). *Multimedia for Learning: Methods and Development*. 3rd Edition. Needham MA: Allyn & Bacon.
- Alexander, L. (n.d.). Retrieved January 05, 2008, from http://thinkexist.com/quotes/lloyd_alexander/.
- Ally, M. (2004). Foundations of Educational Theory for Online Learning. In T. Anderson & F. Elloumi (Eds.), *Theory and Practice of Online Learning* (pp 3-31). Athabasca University: Canada.
- Allinson, C.W. & Hayes, J. (1996). The Cognitive Style Index: A Measure Of Intuition-Analysis For Organizational Research. *Journal of Management Studies*, 3(1), 119-135.
- Aly, M., Elen, J. & Willems, G. (2004). Instructional multimedia program versus standard lecture: a comparison of two methods for teaching the undergraduate orthodontic curriculum. *European Journal of Dental Education*, 8(1), 43–46.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review* 111, (4). 1036-1060. Retrieved 10 July 2008 from <http://act-r.psy.cmu.edu/publications/pubinfo.php?id=526>
- Armstrong, S.J. (2004). The impact of supervisor's cognitive styles on the quality of research supervision in management education. *British Journal of Educational Psychology*, 74(4), 599–616.
- Armstrong, S.J. & Priola, V. (2001). Individual Differences in Cognitive Style and Their Effects on Task and Social Orientations of Self-Managed Work Teams. *Small Group Research*, 32(3), 283–312.
- Ayres, P. (2006a). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology*, 20(3), 287–298.
- Ayres, P. (2006b). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, 16(5), 389-400.

-
- Ayres, P. & Paas, F. (2007). Making Instructional Animations More Effective: A Cognitive Load Approach. *Applied Cognitive Psychology, 21*(6), 695-700.
- Azer, S.A. (2005). A multimedia CD-ROM tool to improve student understanding of bile salts and bilirubin metabolism: evaluation of its use in a medical hybrid PBL course. *Advances in Physiology Education, 29*(1), 40–50.
- Bannert, M. (2002). Managing cognitive load-recent trends in cognitive load theory. *Learning and Instruction, 12*(1), 139–146.
- Bodemer, D., Ploetzner, R., Feuerlein, I. & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction, 14*(3), 325–341.
- Boling, N. C. & Robinson, D. H. (1999). Individual Study, Interactive Multimedia, Or Cooperative Learning: Which Activity Best Supplements Lecture-Based Distance Education? *Journal of Educational Psychology, 91*(1), 169–174.
- Borg, G. (2004). Level-anchored ratio scaling and applications. In: Armando M. Oliviera et al. (Eds.), *Fechner Day 2004*, Coimbra: International Society for Psychophysics.
- Borg, G., Bratfisch, O. & Dornič, S. (1971). On The Problems of Perceived Difficulty. *Scandinavian Journal of Psychology, 12*, 249-260.
- Brann, D.W. & Sloop, S. (2006). Curriculum development and technology incorporation in teaching neuroscience to graduate students in a medical school environment. *Advances in Physiology Education, 30*(1), 38–45.
- Brueckner, J.K. & Traurig, H. (2003). Students' responses to the introduction of a digital laboratory guide in medical neuroscience. *Medical Teacher, 25*(6), 643–648.
- Brünken, R., Plass, J.L. & Leutner, D. (2004). Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects. *Instructional Science, 32*(1), 115-132.
- Brünken, R., Plass, J.L. & Leutner, D. (2003). Direct Measurement of Cognitive Load in Multimedia Learning. *Educational Psychologist, 38*(1), 53-61.
- Brünken, R., Steinbacher, S., Plass, J.L. & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49*(2), 109-119.
- Buzzell, P.R., Chamberlain, V.M. & Pintauro, S.J. (2002). The Effectiveness of Web-Based, Multimedia Tutorials for Teaching Methods of Human Body Composition Analysis. *Advances in Physiology Education, 26*(1), 21-29.
-

-
- Calcaterra, A., Antonietti, A. & Underwood, J. (2005). Cognitive style, hypermedia navigation and learning. *Computers in Education*, 44(4), 441–457.
- Carney, R.N. & Levin, J.R. (2002). Pictorial Illustrations Still Improve Students' Learning From Text. *Educational Psychology Review*, 14(1), 5–26.
- Catrambone, R. & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction*, 16(2), 139–153.
- Chan, M. & Black, J. (2005). When can animation improve learning? Some implications on human computer interaction and learning. In P. Kommers & G. Richards (Eds.), *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005* (pp. 933-938). Chesapeake, VA: AACE.
- Chang, Chun-Yen. (2002). Does Computer-Assisted Instruction + Problem Solving = Improved Science Outcomes? A Pioneer Study. *Journal of Educational Research*, 95(3), 143–150.
- Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction*, 14(3), 353-357.
- Chandler, P. & Sweller, J. (1996). Cognitive Load While Learning to Use a Computer Program. *Applied Cognitive Psychology*, 10(1), 151–170.
- Chandler, P. & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, 8(4), 293-332.
- ChanLin, L. (1999). Visual Control for Dynamic Presentation in Multimedia Learning. *Education Media International*, 36(4), 258–262.
- Chapman, D. M. & Calhoun, J.G. (2006). Validation of learning style measures: implications for medical education practice. *Medical Education*, 40(6), 576–583.
- Chen, Ai-Yen, Mashhadi, A., Ang, D. & Harkrider, N. (1999). Cultural issues in the design of technology-enhanced learning systems. *British Journal of Educational Technology*, 30(3), 217-230.
- Chen, S.Y., Ghinea, G. & Macredie, R. (2006). A cognitive approach to user perception of multimedia quality: An Empirical Investigation. *International Journal of Human-Computer Studies*, 64(12), 1200–1213.
- Cho, M. & Forde, E. (2001). Designing Teaching and Assessment Methods for Diverse Student Populations. *International Journal of Art & Design Education*, 20(1), 86–95.
-

- Chong, E., Balen, R. M. & Jewesson, P. J. (2003). Delivery of multimedia educational presentations via streaming media. *American Journal of Health-System Pharmacy*, 60(23), 2475–2478.
- Clark, D. R. (2007). Learning Domains or Bloom's Taxonomy. Retrieved December 26, 2007 from <http://www.nwlink.com/~donclark/hrd/bloom.html>
- Clarke, T., Ayres, P. & Sweller, J. (2005). The Impact of Sequencing and Prior Knowledge on Learning Mathematics Through Spreadsheet Applications. *Educational Technology Research and Development*, 53(3), 15-24.
- Cohen, L., Manion, L. & Morrison, K. (2000). *Research Methods in Education*. 5th edition. London: RoutledgeFalmer.
- Cohen, J. (1988). *Statistical power analysis for behavioural sciences*. 2nd edition. Hillsdale, NJ: Erlbaum.
- Collaud, G., Gurtner, J-L. & Coen, P-F. (2000). Design and use of a hypermedia system at the University level. *Journal of Computer Assisted Learning*, 16(2), 137-147.
- Colom, R., Flores-Mendoza, C. & Rebello, I. (2003). Working memory and intelligence. *Personality and Individual Differences*, 34(1), 33–39.
- Cook, D.A. (2005). Learning and Cognitive Styles in Web-based Learning: Theory, Evidence and Application. *Academic Medicine*, 80(3), 266–278.
- Cook, D.A. & Smith, A.J. (2006). Validity of index of learning style scores: multitrait - multimethod comparison with three cognitive style/learning style instruments. *Medical Education*, 40(9), 900–907.
- Creswell, John W. (2005). *Educational research : planning, conducting, and evaluating quantitative and qualitative research*. Welwyn Garden City, Herts, UK.
- Creswell, John W. (2002). *Educational research : planning, conducting, and evaluating quantitative and qualitative research*. Welwyn Garden City, Herts, UK.
- Cunningham-Atkins, H., Powell, N., Moore, D., Hobbs, D. & Sharpe, S. (2004). The role of cognitive style in educational computer conferencing. *British Journal of Educational Technology*, 35(1), 69-80.
- Curry, L. (1990). A critique of the research on learning styles. *Educational Leadership*, 48(2), 50–57.

-
- Dalgarno, B. (2001). Interpretations of constructivism and consequences for Computer Assisted Learning. *British Journal of Educational Technology*, 32(2), 183–194.
- Davis, M.J., Wythe, J., Rozum, J.S. & Gore, R.W. (1997). Use of World Wide Web Server and Browser Software To Support a First-Year Medical Physiology Course. *Advances in Physiology Education*, 27(2), 1–14.
- de Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P. & Paas, F. (2007). Attention Cueing as a Means to Enhance Learning from an Animation, *Applied Cognitive Psychology*, 21(6), 731-746.
- Dempsey, John V. & Van Eck, R. (2003). Modality and placement of a pedagogical adviser in individual interactive learning. *British Journal of Educational Technology*, 34(5), 585–600.
- Department of Education (1997). Education White Paper 3: A Programme for Higher Education Transformation. South African Government Gazette, Notice No. 18207, 15 August 1997. Retrieved March 15, 2007, from <http://www.education.gov.za>.
- Diao, Y. & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction*, 17(1), 78–88.
- Dixon, P. (1991). From Research to Theory to Practice: Commentary on Chandler and Sweller, *Cognition and Instruction*, 8(4), 343–350.
- Donnon, T., DesCôteaux, J. & Violato, C. (2005). Impact of cognitive imaging and sex differences on the development of laparoscopic suturing skills. *Canadian Journal of Surgery*, 49(5), 387–393.
- Douglas, G. & Riding, R. J. (1993). The Effect of Pupil Cognitive Style and Position of Prose Passage Title on Recall. *Educational Psychology*, 13(3/4), 385-394.
- Dutke, S. & Rinck, M. (2006). Multimedia learning: Working memory and the learning of word and picture diagrams. *Learning and Instruction*, 16(6), 526–537.
- Dwyer, F.M., Moore, David M. (1991). Effect of Color Coding on Visually Oriented Tests With Students of Different Cognitive Styles. *Journal of Psychology*, 125(6), 677-680.
- Dwyer, T.M., Fleming, J., Randall, J.E. & Coleman, T.G. (1997). Teaching Physiology and the World Wide Web: Electrochemistry and Electrophysiology on the Internet. *Advances in Physiology Education*, 18(1), 2–13.
- DxR Development Group, Inc. (2005). Retrieved August 12, 2007, from <http://www.dxrgroup.com>.
- Effken, J.A. & Doyle, N. (2001). Interface Design and Cognitive Style in Learning an Instructional Computer Simulation. *Computers in Nursing*, 19(4), 164–171.

-
- Ehrman, M. & Leaver, B.L. (2003). Cognitive styles in the service of language learning. *System*, 31(3), 393–415.
- Ellis, S.M. & Steyn, H.S. (2003). Practical significance (effect sizes) versus or in combination with statistical significance (p-values). *Management Dynamics*, 12(4): 51-53.
- Ellis, T. & Cohen, M. (2001). Integrating multimedia into a distance learning environment: Is the game worth the candle? *British Journal of Educational Technology*, 34(4), 495–497.
- Erwin, T. D. & Ricardo, R. (1999). Comparing Multimedia and Traditional Approaches in Undergraduate Psychology Classes. *Teaching of Psychology*, 26(1), 58-61.
- European Learning Styles Information Network (ELSIN). (n.d.). Retrieved 3 March 2008 from <http://www.elsin2008.com>
- Evans, C. (2004). Exploring the Relationship between Cognitive Style and Teaching Style. *Educational Psychology*, 24(4), 509–530.
- Fink, A. & Neubauer, A.C. (2005). Individual differences in time estimation related to cognitive ability, speed of information processing and working memory. *Intelligence*, 33(1), 5–26.
- Fink, A. & Neubauer, A.C. (2001). Speed of information processing, psychometric intelligence: and time estimation as an index of cognitive load. *Personality and Individual Differences*, 30(6), 1009–1021.
- Ford, N. & Chen, S.Y. (2001). Matching/mismatching revisited: an empirical study of learning and teaching styles. *British Journal of Educational Technology*, 32(1), 5-23.
- Frith, V., Jaftha, J. & Prince, R. (2004). Evaluating the effectiveness of interactive computer tutorials for an undergraduate mathematical literacy course. *British Journal of Educational Technology*, 35(2), 159–171.
- Garg, A., Norman, G. & Spero, L. (2001). How medical students learn spatial anatomy. *The Lancet*, 357(9253), 363-364.
- Garg, A., Norman, G., Spero, L. & Taylor, I. (1999). Learning anatomy: do new computer models improve spatial understanding. *Medical Teacher*, 21(5), 519–522.
- Garson, G.D. (2006). *Significance Tests for Two Dependent Samples: McNemar, Marginal Homogeneity, Sign, and Wilcoxon Tests*, from *StatNotes: Topics in Multivariate Analysis*. Retrieved 12 March 2008, from <http://www2.chass.ncsu.edu/garson/pa765/mcnemar.htm>
- Garson, G.D. (2006). *Correlation*, from *StatNotes: Topics in Multivariate Analysis*. Retrieved 28 August 2007, from <http://www2.chass.ncsu.edu/garson/pa765/statnote.htm>.
-

-
- Garson, G.D. (2006). *Scales and Standard Measures*, from *StatNotes: Topics in Multivariate Analysis*. Retrieved 05 September 2007, from <http://www2.chass.ncsu.edu/garson/pa765/statnote.htm>.
- Garson, G.D. (2006). *Research Designs*, from *StatNotes: Topics in Multivariate Analysis*. Retrieved 13 April, 2006, from <http://www2.chass.ncsu.edu/garson/pa765/statnote.htm>.
- Genovese, J.E.C. (2005). Hemispheric Cognitive Style: A Comparison of Three Instruments. *The Journal of Genetic Psychology*, 166(4), 467–481.
- Gerjets, P., Scheiter, K. & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, 16(2), 104–121.
- Gerjets, P., Scheiter, K. & Catrambone, R. (2004). Designing Instructional Examples to Reduce Intrinsic Cognitive Load: Molar versus Modular Presentation of Solution Procedures. *Instructional Science*, 32(1), 33-58.
- Gerjets, P. Scheiter, K. (2003). Goal Configurations and Processing Strategies as Moderators Between Instructional Design and Cognitive Load: Evidence From Hypertext-Based Instruction. *Educational Psychologist*, 38(1), 33–41.
- Gevins, A., Smith, M.E., Leong, H., McEvoy, L., Whitfield, S., Du, R. & Rush, G. (1998). Monitoring Working Memory Load during Computer-Based Tasks with EEG Pattern Recognition Methods. *Human Factors*, 40(1), 79–92.
- Ghinea, G. & Chen, S.Y. (2003). The impact of cognitive styles on perceptual distributed multimedia quality. *British Journal of Educational Technology*, 34(4), 393-406.
- Gunns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15(4), 313–331.
- Gunns, P., Chandler, P. & Sweller, J. (2003). When imagining information is effective. *Contemporary Educational Psychology*, 28(2), 229–251.
- Goldman, S. R. (2003). Learning in complex domains: when and why do multiple representations help? *Learning and Instruction*, 13(2), 239–244.
- Goldman, S. R. (1991). On the Derivation of Instructional Applications From Cognitive Theories: Commentary on Chandler and Sweller. *Cognition and Instruction*, 8(4), 333-342.
- Gómez-Arbonés, X., Ferreira, A., Piqué, M., Roca, J., Tomás, J. Frutos, J.L., Vinyas, J., Prat, J. & Ballester, M. (2004). A cardiological web as an adjunct to medical teaching: prospective analysis. *Medical Teacher*, 26(2), 187-191.
-

-
- Graff, M. (2005). Individual differences in hypertext browsing strategies. *Behaviour & Information Technology*, 24(2), 93–99.
- Graff, M., Davies, J. & McNorton, M. (2004). Cognitive Style and Cross Cultural Differences in Internet Use and Computer Attitudes. *European Journal of Open, Distance and E-Learning*, Retrieved 21 March 2008 from <http://www.euodl.org/>
- Graff, M. (2003a). Cognitive Style and Attitudes Towards Using Online Learning and Assessment Methods. *Electronic Journal of e-Learning*, 1(1), 21–28.
- Graff, M. (2003b). Learning from web-based instructional systems and cognitive style. *British Journal of Educational Technology*, 34(4), 407-418.
- Griffin, J.D. (2003). Technology in the Teaching of Neuroscience. *Advances in Physiology Education*, 27(3), 146–155.
- Griffin, M. M. & Griffin, B. W. (1996). Situated cognition and cognitive style: Effects on students' learning as measured by conventional tests and performance assessments. *The Journal of Experimental Education*, 64(4), 293–309.
- Grimley, M. (2007). Learning from Multimedia Materials: The relative impact of individual differences. *Educational Psychology*, 27(4), 465-485.
- Große, C.S. & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction*, 16(8), 122–138.
- Hall, R. (2002). Aligning learning, teaching and assessment using the web: an evaluation of pedagogic approaches. *British Journal of Educational Technology*, 33(2), 149-158.
- Harden, R.M. (2002). An international virtual medical school (IVIMEDS): the future for medical education? *Medical Teacher*, 24(3), 261–267.
- Hasler, B.S., Kersten, B. & Sweller, J. (2007). Learner Control, Cognitive Load and Instructional Animation. *Applied Cognitive Psychology*. 21(6), 713-729.
- Hayes, J., Allinson, C. W., Hudson, R.S. & Keasey, K. (2003). Further reflections on the nature of intuition-analysis and the construct validity of the Cognitive Style Index. *Journal of Occupational & Organizational Psychology*, 76(2), 269–278.
- Health Education Assets Library (HEAL). (n.d.) Retrieved August 20, 2007 from <http://www.healcentral.org>.
- Hedberg, J.G. & Brown, I. (2002). Understanding Cross-Cultural Meaning through Visual Media. *Education Media International*, 39(1), 23–30.

-
- Hegarty, M. (2004). Dynamic visualizations and learning: getting to the difficult questions. *Learning and Instruction, 14*(3), 243–351.
- Hegarty, M., Kriz, S. & Cate, C. (2003). The Roles of Mental Animations and External Animations in Understanding Mechanical Systems. *Cognition and Instruction, 21*(4), 325-360.
- Hodgkinson, G.P. & Sadler-Smith, E. (2003a). Complex or unitary? A critique and empirical re-assessment of the Allinson--Hayes Cognitive Style Index. *Journal of Occupational & Organizational Psychology, 76*(2), 243–268.
- Hodgkinson, G.P. & Sadler-Smith, E. (2003b). Reflections on reflections...on the nature of intuition, analysis and the construct validity of the Cognitive Style Index. *Journal of Occupational & Organizational Psychology, 76*(2), 279–281.
- Höffler, T. N. (2007). *Lernen mit dynamischen Visualisierungen: Metaanalyse und experimentelle Untersuchungen zu einem naturwissenschaftlichen Lerninhalt. [Learning with dynamic visualizations: A meta-analysis and experimental studies about a scientific topic.]* Unpublished doctoral dissertation, University of Duisburg-Essen, Essen, Germany.
- Höffler, T.N., Sumfleth, E. & Leutner, D. (2006). The Role of Spatial Ability when Learning from an Instructional Animation or a Series of Static Pictures. *Proceedings from the 2006 NYU Steinhardt Technology and Learning Symposium*, April 2006. Retrieved March 15, 2008 from <http://create.alt.ed.nyu.edu/symposium2006/proceedings.html>
- Howard, W. G., Ellis, H. H. & Rasmussen, K. (2004). From the arcade to the classroom: capitalizing on students' sensory rich media preferences in disciplined-based learning. *College Student Journal, 38*(3), 431-441.
- Hulsman, R.L., Ros, W.J.G., Winnubst, J.A.M. & Bensing, J.M. (2002). The effectiveness of a computer-assisted instruction programme on communication skills of medical specialists in oncology. *Medical Education, 36*(2), 125-134.
- Hyde, J.S. (2005). The Gender Similarities Hypothesis. *American Psychologist, 60*(6), 581–592.
- Inglis, A., Ling, P. & Joosten, V. (1999). *Delivering Digitally: Managing the Transition to the Knowledge Media*. London: Kogan Page.
- Issenberg, S.B., McGaghie, W.C., Gordon, D.L., Symes, S., Petrusa, E.R., Hart, I.R. & Harden, R.M. (2002). Effectiveness of a Cardiology Review Course for Internal Medicine Residents Using Simulation Technology and Deliberate Practice. *Teaching and Learning in Medicine, 14*(1), 223–228.

-
- Issenberg, S. B. & Scalese, R. J. (2004). Multimedia computer-based measures of basic science and clinical skills. *Medical Education*, 38(5), 545–575.
- Jeung, Hyun-Ju & Chandler, P. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17(3), 329-344.
- Jha, V., Widdowson, S. & Duffy, S. (2002). Development and evaluation of an interactive computer-assisted learning program—a novel approach to teaching gynaecological surgery. *British Journal of Educational Technology*, 33(3), 323-331.
- John, D. & Boucouvalas, A. (2002a). Multimedia Tasks and User Cognitive Styles. Proceedings of the Second International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP 2002), Staffordshire University, UK: 15–17 July, 2002.
- John, D & Boucouvalas, A.C. (2002b). The effect of the Visual-Auditory dimension of cognitive style. Proceedings of the 24th International Conference on Information Technology Interfaces, ITI 2002, Cavtat, Croatia, June 24-27, 2002, 179-184.
- John, D & Boucouvalas, A.C. (2002c). User Performance with Audio: The effect of subjects' cognitive styles. *Educational Psychology*, 22(2), 133–147.
- Khalil, M.K., Paas, F., Johnson, T.E. & Payer, A.F. (2005a). Interactive and dynamic visualizations in teaching and learning of anatomy: A cognitive load perspective. *The Anatomical Record*, 286(1), 8-14.
- Khalil, M.K., Paas, F., Johnson, T.E. & Payer, A.F. (2005b). Design of interactive and dynamic anatomical visualizations: The implication of cognitive load theory. *The Anatomical Record*, 286(1), 15-20.
- Kalyuga, S. (2006). Rapid cognitive assessment of learner's knowledge structures. *Learning and Instruction*, 16(1), 1-11.
- Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, 38(1), 23-31.
- Kalyuga, S., Chandler, P., Tuovinen, J. & Sweller, J. (2001). When Problem Solving Is Superior to Studying Worked Examples. *Journal of Educational Psychology*, 93(3), 579 – 588.
- Kalyuga, S., Chandler, P. & Sweller, J. (2001). Learner Experience and Efficiency of Instructional Guidance. *Educational Psychology*, 21(1), 5-23.
- Kalyuga, S., Chandler, P. & Sweller, J. (2000). Incorporating Learner Experience Into the Design of Multimedia Instruction. *Journal of Educational Psychology*, 92(1), 126 – 136.
-

-
- Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13(4), 351-371.
- Kalyuga, S., Chandler, P. & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, 40(1), 1 – 17.
- Kalyuga, S. & Sweller, J. (2005). Rapid Dynamic Assessment of Expertise to Improve the Efficiency of Adaptive E-learning. *Educational Technology Research and Development*, 53(3), 83 – 93.
- Kalyuga, S. & Sweller, J. (2004). Measuring Knowledge to Optimize Cognitive Load Factors During Instruction. *Journal of Educational Psychology*, 96(3), 558 – 568.
- Kearsley, G. (1994 - 2004). *Explorations in Learning & Instruction: The Theory Into Practice Database*. Retrieved: 18 October 2004, from <http://tip.psychology.org/index.html>.
- Kekkonen–Moneta, S, & Moneta, G.B. (2002). E–Learning in Hong Kong: comparing learning outcomes in online multimedia and lecture versions of an introductory computing course. *British Journal of Educational Technology*, 33(4), 423 – 433.
- Kerfoot, B. P., Masser, B. A. & Hafler, J. P. (2003). How does the introduction of computers and wall-mounted plasma screens Impact small group tutorials? Preliminary results. *Medical Education*, 37(5), 478 - 479.
- Kester, L., Kirschner, P. A. & van Merriënboer, J. J.G. (2004). Information presentation and troubleshooting in electrical circuits. *International Journal of Science Education*, 26(2), 239 - 256.
- Kiewra, K.A. & Mayer, R.E. (1997). Effects of advance organizers and repeated presentations on students' learning. *Journal of Experimental Education*, 65(2), 147-160.
- Kirk, R.E. (2001). Promoting Good Statistical Practices: Some Suggestions. *Educational and Psychological Measurement*, 61(2), 213–218.
- Kirschner, P.A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12(1), 1-10.
- Kneebone, R. & ApSimon, D. (2001). Surgical skills training: simulation and multimedia combined. *Medical Education*, 35(9), 909-915.
- Kozhevnikov, M., Hegarty, M. & Mayer, R.E. (2002). Revising the Visualizer-Verbalizer Dimension: Evidence for Two Types of Visualizers. *Cognition and Instruction*, 20(1), 47–77.
- Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 16(2), 205–226.
-

-
- Kukolja Taradi, S., Taradi, M., Radić, K. & Pocrjac, N. (2005). Blending problem-based learning with Web technology positively impacts student learning outcomes in acid-base physiology. *Advances in Physiology Education*, 21(2), 35–39.
- Kukolja Taradi, S. & Taradi, M. (2004). Expanding the traditional physiology class with asynchronous online discussions and collaborative projects. *Advances in Physiology Education*, 28(2), 73–78.
- Laight, D. W. (2004). Attitudes to concept maps as a teaching/learning activity in undergraduate health professional education: influence of preferred learning style. *Medical Teacher*, 26(3), 229–233.
- Leahy, W., Chandler, P. & Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Applied Cognitive Psychology*, 17(4), 401-418.
- Lee, J. (2007). The effects of visual metaphor and cognitive style for mental modeling in a hypermedia-based environment. *Interacting with Computers*, 19(5/6), 614-629.
- Levinson, A.J., Weaver, B., Garside, S., McGinn, H. & Norman, G. (2007). Virtual reality and brain anatomy: a randomized trial of e-learning instructional designs. *Medical Education*, 41(5), 495–501.
- Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction*, 13(2), 177–189.
- Lewis, M.J. (2003). Computer-assisted learning for teaching anatomy and physiology in subjects allied to medicine. *Medical Teacher*, 25(2), 204-206.
- Liu, M. (2004). Examining the performance and attitudes of sixth graders during their use of a problem-based hypermedia learning environment. *Computers in Human Behaviour*, 20(3), 357–379.
- Lowe, R.K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14(3), 257-274.
- Lowe, R.K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13(2), 157–176.
- Lowry, R. (2007). Concepts and Application of Inferential Statistics. Retrieved July 8, 2007, from <http://faculty.vassar.edu/lowry/webtext.html>.
- Luk, S. C. (1998). The relationship between cognitive style and academic achievement. *British Journal of Educational Technology*, 29(2), 137–147.
-

-
- Lynch, T.G., Steele, D.J., Palensky, J.E.J., Lacy, N.L. & Duffy, S.W. (2001). Learning preferences, computer attitudes, and test performance with computer-aided instruction. *The American Journal of Surgery*, 81(4), 368-371.
- Macnab, W., Hansell, M. H. & Johnstone, A. H. (1991). Cognitive style and Analytical ability and their relationship to competence in the biological sciences. *Journal of Biological Education*, 25(2), 135–139.
- Mantovani, F., Castelnuovo, G., Gaggioli, A. & Riva, G. (2003). Virtual Reality Training for Health-Care Professionals. *CyberPsychology & Behavior*, 6(4), 389–395.
- Marcus, N., Cooper, M. & Sweller, J. (1996). Understanding Instructions. *Journal of Educational Psychology*, 88(1), 49–63.
- Martin, I.G., Stark, P. & Jolly, B. (2000). Benefiting from clinical experience: the influence of learning style and clinical experience on performance in an undergraduate objective structured clinical examination. *Medical Education*, 34(7), 530-534.
- Massa, L.J. & Mayer, R.E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences*, 16, 321-335.
- Massa, L.J. & Mayer, R.E. (2005). Three obstacles to validating the Verbal-Imager Subtest of the Cognitive Styles Analysis. *Personality and Individual Differences*, 39(4), 845–848.
- Mayer, R.E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13(2), 125-139.
- Mayer, R.E. (1997a). Multimedia Learning: Are We Asking The Right Questions? *Educational Psychologist*, 19(1), 1-19.
- Mayer, R.E. (1997b). Searching for the perfect ISM: An unproductive activity for educational researchers. *Issues in Education*, 3(2), 225 - 228.
- Mayer, R.E. & Anderson, R.B. (1992). The Instructive Animation: Helping Students Build Connections Between Words And Pictures In Multimedia Learning. *Journal of Educational Psychology*, 84(4), 444–452.
- Mayer, R.E., Bove, W., Bryman, A., Mars, R. & Tapangco, L. (1996). When Less Is More: Meaningful Learning From Visual And Verbal Summaries Of Science Textbook Lessons. *Journal of Educational Psychology*, 88(1), 64–73.
- Mayer, R.E. & Chandler, P. (2001). When Learning Is Just a Click Away: Does Simple User Interaction Foster Deeper Understanding. *Journal of Educational Psychology*, 93(2), 390–397.
-

- Mayer, R.E., Dow, G.T. & Mayer, S. (2003). Multimedia Learning in an Interactive Self-Explaining Environment: What Works in the Design of Agent-Based Microworlds? *Journal of Educational Psychology*, 95(4), 806–812.
- Mayer, R.E., Fennell, S., Farmer, L. & Campbell, J. (2004). Personalization Effect in Multimedia Learning: Students Learn Better When Words Are in Conversational Style Rather Than Formal Style. *Journal of Educational Psychology*, 96(2), 389-395.
- Mayer, Richard E. & Gallini, J.K. (1990), When Is an Illustration Worth Ten Thousand Words? *Journal of Educational Psychology*, 82(4), 715–716.
- Mayer, R.E. & Massa, L. (2003). Three Facets of Visual and Verbal Learners: Cognitive Ability, Cognitive Style, and Learning Preference. *Journal of Educational Psychology*, 95(4), 833-846.
- Mayer, R.E., Mautone, P. & Prothero, W. (2002). Pictorial Aids for Learning by Doing in a Multimedia Geology Simulation Game. *Journal of Educational Psychology*, 94(1), 171–185.
- Mayer, R.E. & Moreno, R. (2003). Animation as an Aid to Multimedia Learning. *Educational Psychology Review*, 14(1), 43–52.
- Mayer, R.E. & Moreno, R. (2002). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 87-99.
- Mayer, R.E. & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working. *Journal of Educational Psychology*, 90(2), 312-320.
- Mayer, R.E., Moreno, R., Boire, M. & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, 91(4), 638–643.
- Mayer, R.E., Sobko, K. & Mautone, P. D. (2003). Social Cues in Multimedia Learning: Role of Speaker's Voice. *Journal of Educational Psychology*, 9(2), 419 – 425.
- McDonough, M. & Marks, I.M. (2002). Teaching medical students exposure therapy for phobia/panic - randomized, controlled comparison of face-to-face tutorial in small groups vs. solo computer instruction. *Medical Education*, 36(5), 412-417.
- McGrath, P., Kucera, T. & Smith, W. (2003). Computer Simulation of Introductory Neurophysiology, *Advances in Physiology Education*, 27(3), 120–129.
- McKay, E. (1999). An Investigation of Text-based Instructional Materials Enhanced with Graphics. *Educational Technology*, 19(3), 323–335.

-
- McMillan, J.H. & Schumacher, S. (2001). *Research in Education: A Conceptual Introduction*. 5th edition. Addison Wesley Longman, Inc.
- McNeal, G.H. & Dwyer, F. (1999). Effect of learning style on consistent and inconsistently designed instruction. *International Journal of Instructional Media*, 26(3), 337-345.
- Mergel, B. (1998). *Instructional Design & Learning Theory*. Retrieved: 18 October 2004, from <http://www.usask.ca/education/coursework/802papers/mergel/brenda.htm>.
- Merrill, M.D. (1983). Component Display Theory. In C. Reigeluth (ed.), *Instructional Design Theories and Models*. Hillsdale, NJ: Erlbaum Associates.
- Meyer, B.J., van Papendorp, D.H., Meij, H.S. & Viljoen, M. (2002). *Human Physiology*. 3rd edition. Cape Town: Juta.
- Michael, J. (2007). What makes physiology hard for students to learn? Results of a faculty survey. *Advances in Physiology Education*, 31(1), 34-49.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4), 159-167.
- Mikk, J. & Luik, P. (2003). Characteristics of multimedia textbooks that affect posttest scores. *Journal of Computer Assisted Learning*, 19(4), 528-537.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychology Review*, 63(2), 81-97.
- Moreno, R. (2007). Optimising Learning from Animations by Minimising Cognitive Load: Cognitive and Affective Consequences of Signalling and Segmentation Methods. *Applied Cognitive Psychology*, 21(6), 765-781.
- Moreno, R. (2006). When worked examples don't work: Is cognitive load theory at an Impasse? *Learning and Instruction*, 16(2), 170–181.
- Moreno, R. (2004). Decreasing Cognitive Load For Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science*, 32(1), 99–113.
- Moreno, R., & Marley, S. (2007). *Do students' verbal and visual abilities and preferences affect their learning and perceptions about learning astronomy with static and animated graphics?* Paper presented at the European Association for Research on Learning and Instruction Annual Conference, August 2007, Budapest, Hungary.
- Moreno, R. & Mayer, R.E. (2002). Verbal Redundancy in Multimedia Learning: When Reading Helps Listening. *Journal of Educational Psychology*, 94(1), 156–163.
-

-
- Moreno, R. & Mayer, R.E. (2000a). Engaging Students in Active Learning: The Case for Personalized Multimedia Messages. *Journal of Educational Psychology*, 92(4), 724-733.
- Moreno, R. & Mayer, R.E. (2000b). A Coherence Effect In Multimedia Learning: The Case For Minimizing Irrelevant Sounds In The Design Of Multimedia Instructional Messages. *Journal of Educational Psychology*, 92(1), 117–125.
- Moreno, R. & Mayer, R.E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2), 358–368.
- Moreno, R., Mayer, R.E., Spires, H.A. & Lester, J.C. (2001). The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents? *Cognition and Instruction*, 19(2), 177–213.
- Moreno, R. & Valdez, A. (2005), Cognitive Load and Learning Effects of Having Students Organise Pictures and Words in Multimedia Environments: The Role of Student Interactivity and Feedback. *Educational Technology Research and Development*, 53(3), 35–45.
- Morton, D.A., Foreman, K.B., Goede, P.A., Bezzant J.L. & Albertine, K.H. (2007). Tk3 eBook software to author, distribute, and use electronic course content for medical Education. *Advances in Physiology Education*, 31(3), 55–61.
- Murphy, E. (2003). Moving from theory to practice in the design of web-based learning from the perspective of Constructivism. *The Journal of Interactive Online Learning*, 1, (4). Retrieved January 14, 2004 from <http://www.ncolr.org/jiol/archives/2003/spring/4/MS02028.pdf>
- Murphy, H. J. & Casey, B. (1997). Field dependence/independence and undergraduates' academic performance in an information management program. *College Student Journal*, 31(1), 45-51.
- Nietfeld, J. & Bosma, A. (2003). Examining the self-regulation of impulsive and reflective response styles on academic tasks. *Journal of Research in Personality*, 37(3), 118–140.
- O'Donnell, A.M., Dansereau, D.E. & Hall, R.H. (2002). Knowledge Maps as Scaffolds for Cognitive Processing. *Educational Psychology Review*, 14(1), 71–86.
- Ong, Y.W. & Milech, D. (2004). Comparison of the Cognitive Styles Analysis and the Style of Processing Scale. *Perceptual and Motor Skills*, 99(1), 155–162.
- Ong, Y.W. & Milech, D. (2001). The Style Of Processing Scale: Normative And Reliability Data.. *Perceptual and Motor Skills*, 93(3), 595-598.
- Onwuegbuzie, A.J. & Leech, N. L. (2004). Enhancing the Interpretation of “Significant” Findings: The Role of Mixed Methods Research. *The Qualitative Report*, 9(4), 770–792.

-
- Oxford English Dictionary Online. (2006). (2nd Edition). Oxford University Press. Retrieved: 18 August 2006, from <http://www.oed.org>.
- Paas, F.G.W.C. (1992). Training Strategies For Attaining Transfer Of Problem-Solving Skill In Statistics: A Cognitive Load Approach. *Journal of Educational Psychology*, 82(4), 429–434.
- Paas, F. & Kester, L. (2006). Learner and information characteristics in the design of powerful learning environments. *Applied Cognitive Psychology*, 20(3), 281–285.
- Paas, F., Renkl, A. & Sweller, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, 32(1), 1-8.
- Paas, F., Renkl, A. & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1-4.
- Paas, F., Tuovinen, J.E., van Merriënboer, J.J.G. & Darabi, A. A. (2005). A Motivational Perspective on the Relation Between Mental Effort and Performance: Optimizing Learner Involvement in Instruction. *Educational Technology Research and Development*, 53(5), 25–34.
- Paas, F., Tuovinen, J.E., Tabbers, H. & Van Gerven, P.W.M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, 38(1), 63-71.
- Paas, F., Van Gerven, P.W.M, & Wouters. P. (2007). Instructional Efficiency of Animation: Effects of Interactivity through Mental Reconstruction of Key Static Frames. *Applied Cognitive Psychology*, 21(6), 783-793.
- Paas, F. & van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. *Learning and Instruction*, 16(2), 87-91.
- Paas, F.G.W.C. & van Merriënboer, J.J.G. (1994a). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351-371.
- Paas, F.G.W.C., van Merriënboer, J.J.G. & Adam, J.J. (1994b). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, 79(1), 419-430.
- Paas, F.G.W.C., van Merriënboer, J.J.G. & Adam, J.J. (1994c). Variability of Worked Examples and Transfer of Geometrical Problem-solving Skills: A Cognitive Load Approach. *Journal of Educational Psychology*, 86(1), 122-133.
- Paas, F.G.W.C. & van Merriënboer, J.J.G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors*, 35(4), 373–743.
-

-
- Palmquist, R. A. (2001). Cognitive Style and Users' Metaphors for the Web: An Exploratory Study. *Journal of Academic Librarianship*, 27(1), 24–32.
- Paivio, A. (1986). *Mental Representations: A dual coding approach*. Oxford, England: Oxford University Press.
- Parkinson, A., Mullally, A.A.P. & Redmond, J.A. (2004). Test–retest reliability of Riding's cognitive styles analysis test. *Personality and Individual Differences*, 37(6), 1273-1278.
- Peterson, E. R. & Deary, I. J. (2006). Examining Wholistic–Analytic style using preferences in early information processing. *Personality and Individual Differences*, 41(1), 3-14.
- Peterson, E. R., Deary, I. J. & Austin, E.J. (2005a). A new measure of Verbal-Imagery Cognitive Style: VICS. *Personality and Individual Differences*, 38(6), 1269 – 1281.
- Peterson, E. R., Deary, I. J. & Austin, E.J. (2005b). Are intelligence and personality related to verbal-Imagery and Wholistic-Analytic cognitive styles?. *Personality and Individual Differences*, 39(1), 201-213.
- Peterson, E.R., Deary, I J. & Austin, E.J. (2003a). The reliability of Riding's Cognitive Style Analysis test. *Personality and Individual Differences*, 34(5), 881-891.
- Peterson, E.R., Deary, I J. & Austin, E.J. (2003b). On the assessment of cognitive style: four red herrings. *Personality and Individual Differences*, 34(5), 899-904.
- Ploetzner, R. & Lowe, R. (2004). Dynamic Visualisations and Learning. *Learning and Instruction*, 14(3), 235-357.
- Pollock, E., Chandler, P. & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12(1), 61-86.
- Quealy, J. & Langan-Fox, J. (1998). Attributes of delivery media in computer-assisted instruction. *Ergonomics*, 41(3), 257-279.
- Rayner, S. & Riding, R. (1997). Towards a categorisation of cognitive styles and learning styles. *Educational Psychology*, 17(1/2), 5–28.
- Rawson, R.E. & Quinlan, K.M. (2002). Evaluation Of A Computer-Based Approach To Teaching Acid/Base Physiology. *Advances in Physiology Education*, 26(2), 85–97.
- Rees, L. H. (2000). Medical education in the new millennium. *Journal of Internal Medicine*, 248(2), 95–101.
-

-
- Regnard, C. (2000). Using videoconferencing in palliative care. *Palliative Medicine*, 14(6), 519-528.
- Reimann, P. (2003). Multimedia learning : beyond modality. *Learning and Instruction*, 13(2), 245–252.
- Reisslein, J., Atkinson, R.K., Seeling, P. & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction*, 16(2), 92–103.
- Rendas, A.B., Pinto, P.R. & Gamboa, T. (1998). Problem-Based Learning in Pathophysiology: Report of a Project and Its Outcome. *Teaching & Learning in Medicine*, 10(10), 34–39.
- Renkl, A., Atkinson, R.K. & Große, C.S. (2004). How Fading Worked Solution Steps Works – A Cognitive Load Perspective, *Instructional Science*, 32(1), 59–82.
- Renkl, A. & Atkinson, R.K. (2003). Structuring the Transition from Example Study to Problem Solving in Cognitive Skill Acquisition: A Cognitive Load Perspective. *Educational Psychologist*, 38(1), 15–22.
- Rezaei, A.R. & Katz, L. (2004). Evaluation of the reliability and validity of the cognitive styles analysis. *Personality and Individual Differences*, 36(6), 1317-1327.
- Riding, R. (2005a). *Cognitive Styles Analysis: Administration*. Learning & Training Technology: Stourbridge, UK.
- Riding, R. (2005b). *Cognitive Styles Analysis: Style Nature and Assessment*. Learning & Training Technology: Stourbridge, UK.
- Riding, R.J. (2003). On the assessment of cognitive style: a commentary on Peterson, Deary, and Austin. *Personality and Individual Differences*, 34(5), 893–897.
- Riding, R. J. (1997). On the nature of cognitive style. *Educational Psychology*, 17(1/2), 29-49.
- Riding, R. & Agrell, T. (1997). The effect of cognitive style and cognitive skills on school subject performance. *Educational Studies*, 23(2), 311–324.
- Riding, R.J. & Ashmore, J. (1980) Verbaliser-Imager learning style and children's recall of information presented in pictorial versus written form, *Educational Studies*, 6, 141-145.
- Riding, R., Burton, D., Rees, G. & Sharratt, M. (1995). Cognitive style and personality in 12-year-old children. *British Journal of Educational Psychology*, 65, 113–124.
- Riding, R.J. & Caine, T. (1993). Cognitive style and GCSE performance in Mathematics, English Language and French. *Educational Psychology*, 13(1), 59-68.
-

-
- Riding, R.J. & 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*, 79, 59-64.
- Riding, R. & Cheema, I. (1991). Cognitive styles--an overview and integration. *Educational Studies*, 18(3), 323–351.
- Riding, R.J. & Glass, A. (1997). Cognitive style and individual differences in EEG alpha during information processing, *Educational Psychology*, 17(1/2), 219–234.
- Riding, R.J., Glass, A. & Douglas, G. (1993). Individual Differences in Thinking: cognitive and neurophysiological perspectives. *Educational Psychology*, 13(3/4), 267-280.
- Riding, R.J., Grimley, M., Dahraei, H. & Banner, G. (2003). Cognitive style, working memory and learning behaviour and attainment in school subjects. *British Journal of Educational Psychology*, 73(2), 149-169.
- Riding, R.J & Grimley, M. (1999). Cognitive style, gender and learning from multi-media materials in 11-year-old children. *British Journal of Educational Technology*, 30(1), 43-56.
- Riding, R.J. & Mathias, D. (1991). Cognitive styles and preferred learning mode, reading attainment and cognitive ability in 11-year-old children. *Educational Psychology*, 11(3/4), 383-394.
- Riding, R. & Pearson, F. (1995). The Relationship Between Cognitive Style and Intelligence. *Educational Psychology*, 14(4), 413–425.
- Riding, R.J., & Rayner, S.G. (1998). *Cognitive Styles and Learning Strategies*. David Fulton Publisher, London.
- Riding, R.J. & Read, G. (1996). Cognitive style and pupil learning preferences. *Educational Psychology*, 16(1), 81-107.
- Riding, R. & Sadler-Smith, E. (1992). Type of instructional material, cognitive style and learning performance. *Educational Studies*, 18(3), 323-341.
- Riding, R. & Staley, A. (1998). Self-Perception As Learner, Cognitive Style And Business Studies Student's Course Performance. *Assessment and Evaluation in Higher Education*, 23(1), 43-59.
- Rieber, L.P., Tzeng, S. & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: finding the right mix. *Learning and Instruction*, 14(3), 307–323.
- Rikers, R.M.J.P., van Gerven, P.W.M. & Schmidt, H.G. (2004). Cognitive Load Theory as a Tool for Expertise Development. *Instructional Science*, 32(1), 173–182.
-

-
- Robinson, D.H.(Ed). (2002). Spatial Text Adjuncts and Learning: An Introduction to the Special Issue. *Educational Psychology Review*, 14(1), 1-120.
- Ross, J. & Schulz, R. (1999). Can computer-aided instruction accommodate all learners equally? *British Journal of Educational Technology*, 30(1), 5-24.
- Rush, G. M. & Moore, D.M. (1991). Effects of restructuring training and cognitive style. *Educational Psychology*, 11(3/4), 309–322.
- Rutherford, D. & Lloyd, W. J. (2001). Assessing a Computer-aided Instructional Strategy in a World Geography Course. *Journal of Geography in Higher Education*, 25(3), 341-355.
- Sabry, K. & Baldwin, L. (2003). Web-based learning interaction and learning styles. *British Journal of Educational Technology*, 34(4), 443–454.
- Sadler-Smith, E. Spicer, D.P. & Tsang, F. (2000). Validity of the Cognitive Style Index: Replication and Extension. *British Journal of Management*, 11, 75–81.
- Salden, R.J.C.M., Paas, F., Broers, N.J. & van Merriënboer, J.J.G. (2004). Mental Effort and Performance as Determinants for the Dynamic Selection of Learning Tasks in Air Traffic Control Training, *Instructional Science*, 32(1), 153–172.
- Scherly, D., Roux, L. & Dillenbourg, P. (2000). Evaluation of hypertext in an activity learning environment. *Journal of Computer Assisted Learning*, 16(2), 125–136.
- Schmidt, H.G., Vermeulen, L. & van der Molen, H.T. (2006). Long-term effects of problem-based learning: a comparison of competencies acquired by graduates of a problem-based and a conventional medical school. *Medical Education*, 40(6), 562–567.
- Schnotz, W. & Lowe, R. (2003). External an internal representations in multimedia learning. *Learning and Instruction*, 13(2), 117-123.
- Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13(2), 141–156.
- Schnotz, W. & Rasch, T. (2005). Enabling, Facilitating, and Inhibiting Effects of Animations in Multimedia Learning: Why Reduction of Cognitive Load Can Have Negative Results on Learning. *Educational Technology Research and Development*, 53(3), 47-58.
- Schultze-Mosgau, S. Zielinski, T. & Lochner, J. (2004). Interactive, web-based e-lectures with a multimedia online examination. *Medical Education*, 38(11), 1184-1184.
- Schwan, S. & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, 14(3), 293–305.
-

-
- Sefton, A.J. (1998). The Future of Teaching Physiology: An International Viewpoint. *Advances in Physiology Education*, 20(1), 53–58.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13(2), 227–237.
- Sherwood, L. (2004). *Human physiology: from cells to systems*. 5th edition. Belmont, CA: Thomson.
- Silverthorn, D.U. (2004). *Human physiology : an integrated approach*. 3rd edition. San Francisco : Pearson/Benjamin Cummings.
- Sinclair, K. J., Renshaw, C.E. & Taylor, H. A. (2004). Improving computer-assisted instruction in teaching higher-order skills. *Computers and Education*, 42(2), 169–180.
- Smith, M.E. (2007). *Factors in the measurement of cognitive load of multimedia learning*. Unpublished M Ed thesis: University of Pretoria.
- Smith, Stephen M. & Woody, Paul C. (2000). Interactive Effect of Multimedia Instruction and Learning Styles. *Teaching of Psychology*, 27(3), 220–224.
- Smits, P.B.A., Verbeek, J.H.A.M., Nauta, M.C.E., Ten Cate, Th.J., Metz, J.C.M. & van Dijk, F.J.H. (2004). Factors predictive of successful learning in postgraduate medical education. *Medical Education*, 38(x), 758-766.
- Sonnenwald, D. & Li, B. (2003). Scientific laboratories in higher education: exploring learning style preferences and perceptions of technology. *British Journal of Educational Technology*, 34(3). 419–431.
- Spickard 111, A., Smithers, J., Cordray, D., Gigante, J. & Wofford, J.L. (2004). A randomized trial of an online lecture with and without audio. *Medical Education*, 38(7), 787-790.
- Stark, R., Mandl, H., Gruber, H. & Renkl, A. (2002). Conditions and effects of example elaboration. *Learning and Instruction*, 12(1), 39–60.
- Statistics South Africa. (2005). Demography. *Bulletin of Statistics*, 39(4), 1.1–1.2.
- Steele, D.J., Palensky, J.E.J., Lynch, T.G., Lacy, N.L. & Duffy, S.W. (2002). Learning preferences, computer attitudes, and student evaluation of computerised instruction. *Medical Education*, 36(3), 225 – 232.
- Stern, D.T., Mangrulkar, R.S., Gruppen, L.D., Lang, A.L., Grum, C.M. & Judge, R.D. (2001). Using a Multimedia Tool to Improve Cardiac Auscultation Knowledge and Skills. *Journal of General Internal Medicine*, 16(11), 763–769.

-
- Sternberg, J.R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, England: Cambridge University Press.
- Sturmberg, J.P., Crowe, P. & Hughes, C. (2003). Computer-assisted instruction: guiding learning through a key features approach in a community-based medical course. *Medical Teacher*, 25(3), 332-335.
- Sweller, J. (2006a). Cognitive Load Theory: Theoretical Underpinnings. *Conference proceedings of The 2006 Technology & Learning Symposium @NYU*. University of New York.
- Sweller, J. (2006b). Discussion of 'emerging topics in cognitive load research: using learner and information characteristics in the design of powerful learning environments'. *Applied Cognitive Psychology*, 20(3), 353–357.
- Sweller, J. (2006c). The worked example effect and human cognition. *Learning and Instruction*, 16(2), 165–169.
- Sweller, J. (2004). Instructional Design Consequences of an Analogy between Evolution by Natural Selection and Human Cognitive Architecture. *Instructional Science*, 32(1/2), 9–31.
- Sweller, J. (1988). Cognitive load during problem solving: effects on Learning. *Cognitive Science*, 12, 257-285.
- Sweller, J. & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and Instruction*, 12(3), 185-233.
- Sweller, J. & Chandler, P. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, 8(4), 293 – 332.
- Sweller, J. & Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition and Instruction*, 8(4), 351-362.
- Sweller, J., van Merriënboer, J.J.G & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251-296.
- Tabbers, H.K., Martens, R.L. & van Merriënboer, J.J.G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, 74(1), 71-81.
- The Virtual Labs Project. (2003). Retrieved August 12, 2007 from <http://virtuallabs.stanford.edu/>.
<http://virtuallabs.stanford.edu/demo/index.html>.
- Thornell, J. G. (1976). Research on Cognitive Styles: Implications for Teaching and Learning. *Educational Leadership*, 33(7), 502-504.
-

Tinajero, C. & Paramo, M. F. (1993). Field dependence-independence and brain organization: The confluence of two different ways of describing general forms of cognitive functioning? A theoretical Review. *Perceptual and Motor Skills*, 77(3), 787–802.

Tomasi, D., Chang, L., Caparelli, E.C. & Ernst, T. (2007). Different activation patterns for working memory load and visual attention load. *Brain Research*, 1132, 158-165.

Treadwell, I., de Witt, T.W. & Grobler, S. (2002). The impact of a new educational strategy on acquiring neonatology skills. *Medical Education*, 36(5), 441-448.

Triantafyllou, E., Pomportsis, A., Demetriadis, S. & Georgiadou, E. (2004). The value of adaptivity based on cognitive style: an empirical study. *British Journal of Educational Technology*, 35(1), 95–106.

Triantafyllou, E., Pomportsis, A. & Demetriadis, S. (2003). The design and the formative evaluation of an adaptive educational system based on cognitive styles. *Computers & Education*, 41(1), 87–104.

Trinder, R. (2002). FORUM: Multimedia in the Business English Classroom: The Learners' Point of View. *Computer Assisted Language Learning*, 15(1), 69-84.

Tuovinen, J.E. & Paas, F. (2004). Exploring Multidimensional Approaches to the Efficiency of Instructional Conditions. *Instructional Science*, 32(1), 133–153.

Tuovinen, J. E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334–341.

Tversky, B., Morrison, J.B. & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 247 – 262.

Vacha-Haase, T. (2001). Statistical Significance should not be Considered one of Life's Guarantees: Effect Sizes are needed. *Educational and Psychological Measurement*, 61(2), 219-225.

Valcke, M. (2002). Cognitive load: updating the theory, *Learning and Instruction*, 12(1), 147–154.

van der Meij, J. & de Jong, T. (2006), Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction*, 16(3), 199–212.

Van Gog, T., Paas, F. & van Merriënboer, J.J.G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction*, 16(2), 154-164.

van Gog, T., Ericsson, K. A., Rikers, R.M.J.P. & Paas, F. (2005). Instructional Design for Advanced Learners: Establishing Connections Between the Theoretical Frameworks of Cognitive Load and Deliberate Practice. *Educational Technology Research & Development*, 53(3), 73–81.

Van Gog, T., Paas, F & van Merriënboer, J.J.G. (2004). Process-Oriented Worked Examples: Improving transfer Performance Through Enhanced Understanding. *Instructional Science*, 32(1), 83-98.

van Merriënboer, J.J.G. (1990). Instructional strategies for teaching computer programming: interactions with the cognitive style reflection-impulsivity. *Journal of Research on Computing in Education*, 23(1), 45–53.

Van Merriënboer, J.J.G. & Ayres, P. (2005). Research on Cognitive Load Theory and Its Design Implications for E-Learning. *Educational Technology Research and Development*, 53(3), 1–104.

van Merriënboer, J.J.G. , Kirschner, P. A. & Kester, L. (2003). Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning. *Educational Psychologist*, 38(1), 5-13.

Van Merriënboer, J.J.G. & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, 17(2), 147–177.

van Oostendorp, H. & Beijersbergen, M.J. (2007). Animated diagrams: Their effect on understanding, confidence and mental effort. *Proceedings from the 12th Biennial Conference for Research on Learning and Instruction*, August 28 - September 1, 2007. Retrieved 25 March 2008 from <http://earli2007.hu/ng/home/>

Vaske, J. (2002). Communicating Judgments About Practical Significance: Effect Size, Confidence Intervals and Odds Ratios. *Human Dimensions of Wildlife*, 7(4), 287–300.

Verdi, M.P., & Kulhavy, R.W. (2002). Learning with Maps and texts: An Overview. *Educational Psychology Review*, 14(1), 27–46.

Vernon, T. & Peckham, D. (2002). The benefits of 3D modelling and animation in medical teaching. *Journal of Visual Communication in Medicine*, 25(4), 142–148.

Vichitvejpaisal, P., Sitthikongsak, S., Parakkamodom, S., Manon, C. & Petcharatana, S. (2001). Does computer-assisted instruction really help to improve the learning process? *Medical Education*, 35(10), 983-989.

Vivekananda-Schmidt, P., Hassell, A. & McLean, M. (2004). The evaluation of multimedia learning packages in the education of health professionals: experience of a musculo-skeletal examination package. *Nurse Researcher*, 11(3), 43-55.

Wallen, E., Plass, J.L. & Brünken, R. (2005). The Function of Annotations in the Comprehension of Scientific Texts: Cognitive Load Effects and the Impact of Verbal Ability. *Educational Technology Research and Development*, 53(3), 59–72.

Westendorp, M.W. & McGraw, R.C. (2002). Computer-assisted instruction of carpal bone radiograph interpretation. *Medical Teacher*, 24(6), 605-608.

Whelan, R.R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review*, 2(1), 1–12.

Whelan, R.R. (2006). Measuring Cognitive Load: A comparative study of self-report and direct measurement techniques. Unpublished manuscript.

Wieseman, R. A. & Portis, S. C. (1990). Strategies for putting dependent and independent learners in the winner's circle. *Journal of Instructional Psychology*, 17(3), 161-168.

Wikipedia - Online Encyclopedia. (2006). Retrieved 18 August, 2006, from <http://en.wikipedia.org>.

Williams, C., Aubin S., Harkin, P. & Cottrell, D. (2001). A randomised, controlled, single blind trial of teaching provided by a computer-based multimedia package versus lecture. *Medical Education*, 35(9), 847-854.

Winkelman C. (2001). Effect Size: Utility and Application in Neuroscience Nursing. *Journal of Neuroscience Nursing*, 33(4), 216.

Witkin, H.A., Moore, C.A., Goodenough, D.R. & Cox, P.W. (1977). Field-Dependent and Field-Independent Cognitive Styles and their Educational Implications. *Review of Educational Research*, 47(1), 1–64.

Wittrock, M.C. (1992). Generative Learning Processes of the Brain. *Educational Psychologist*, 27(4), 531–541.

Wooten, K.C., Barner, B.O. & Silver, N.C. (1994). The influence of cognitive style upon work environment preferences. *Perceptual and Motor Skills*, 79(1), 307–314.

Workman, M. (2004). Performance and perceived effectiveness in computer-based and computer-aided education: do cognitive styles make a difference? *Computers in Human Behaviour*, 20(4), 527 – 534.

Yeung, A.S., Jin, P. & Sweller, J. (1998). Cognitive Load and Learner Expertise: Split-Attention and Redundancy Effects in Reading with Explanatory Notes. *Contemporary Educational Psychology*, 23(1), 1–21.

Zahn, C., Barquero, B. & Schwan, S. (2004). Learning with hyperlinked videos - design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction*, 14(3), 275–291.

Zhang, Li-fang. (2004). Field-dependence/independence: cognitive style or perceptual ability?-- validating against thinking styles and academic achievement. *Personality and Individual Differences*, 37(6), 1295–1311.

**APPENDIX A: SUMMARY OF COGNITIVE STYLES RESEARCH REVIEWED
BY RIDING AND CHEEMA (1991)**

Holistic – Analytic Style Dimension		
Style labels	Theorist	Basic description
Field dependence (FD) – independence (FI)	Witkin (1962)	Field Dependence (FD) - rely on external frames of reference, prefer group interaction, like structure. Field Independence (FI) – rely on internal frames of reference, prefer to work on individual tasks, like to impose their own structure.
Impulsivity – Reflectivity	Kagan (1964)	Impulsive style characterised by quick responses to request to undertake task, whereas the reflective style will deliberate over the issue before providing a response.
Convergent - Divergent thinkers	Guilford (early 1950's)	Convergent thinkers deal best with problems which require the ability to generate one correct answer, while divergent thinkers are perform well when required to generate several equally acceptable answers.
Levellers - Sharpeners	Holzman & Klein (1954)	Focuses primarily on how a visual task is perceived. Levellers tend to perceive a task very simply and assimilate new events with previously stored ones, while sharpeners perceive a task in a complex and differentiated fashion, with little assimilation.
Holists - serialists	Pask (1972)	Holists will scan large amounts of data and look for patterns, while serialists will examine less data and use a step-by-step approach when completing a task.
Verbaliser-Imager Style Dimension		
Style Label	Theorist	Basic description
Abstract – concrete	Harvey (1961)	Preferred level and capacity of abstraction.
Verbaliser - visualiser	Paivio (1971), Riding and Taylor (1976), Riding and Calvey (1981)	Visualisers better than verbalisers in the recall of high imagery material. Will use either verbal or visual strategies to represent knowledge and thinking.

APPENDIX B: LETTER REQUESTING PERMISSION TO CONDUCT THE STUDY AND LETTER PROVIDING THIS CONSENT

12 September 2005

Head of Department: Department of Physiology
 Faculty of Health Sciences
 Basic Medical Sciences Building – 9-8
 University of Pretoria
 0002

Dear Prof van Papendorp

PERMISSION TO CONDUCT PhD RESEARCH USING PHYSIOLOGY AS CONTEXT AND CONTENT

I am currently registered as a PhD student in the Faculty of Education at the University of Pretoria. My research proposal outlines a study which aims to investigate the relationship between cognitive load and cognitive styles when using animations as learning resources within a specific content domain. I attach an executive summary of my proposal. I successfully defended my proposal on the 22 April 2005.

My work in the field of multimedia development over the last five years has primarily been in the health sciences education field. This is a field which makes extensive use of multimedia learning resources. As such I would like to use the context of health sciences education in general, and physiology education in particular for my research. It is also my aim to base this research in a setting which is as authentic as possible – in other words, to take content which must be studied by all students who take physiology as part of their curriculum. Many research studies in this field use content which is not part of the student's normal curriculum. There is a need for research which uses more authentic learning experiences.

I therefore request permission to develop multimedia content which covers various sections of the physiology of the renal system, and to use the students at UP who take physiology as a subject as participants in the study. On completion of the study all content developed will be given to the University of Pretoria.

There will be two phases in this research. The detail about the times and particular student groups are outlined on the table below.

Phase	Student group	No of participants	Time frame	Duration of experiment
Pilot study	Students doing Physiology which is not part of the MBChB program	80 – 120	February 2006	2 sessions First – 20 – 30min Second – 1.5 - 2 hours
Larger study	MBChB students – during	400 - 500	April / May 2006	2 sessions

	Other students in the Faculty of Health Sciences		Whenever it fits into their curriculum	First – 20 – 30min Second – 1.5 - 2 hours
--	--	--	--	--

I will therefore require approximately 2 hours of time from each student who participates in the study. These sessions will be conducted in the computer laboratories on with the Hatfield campus or the Prinshof campus. Participation will be voluntary. I furthermore request permission to approach the lecturers in your department who teach the physiology of the renal system to serve as the content experts who will guide the development. I will continue with the necessary application and ethical review with the appropriate committees should you give me permission to undertake this study.

Yours sincerely

Anne Strehler

Student No: 77006799

APPENDIX C: PERMISSION FROM SMITH (2007) TO USE DATA FROM HER STUDY

P O Box 74000
Lynnwood Ridge
0040
20 February 2008

Department of Statistics
University of Pretoria
Pretoria

Dr M v. d. Linde

I, M E Smith, hereby grant Anne Strehler permission to use any of the datasets from the research project DPG9077 - OD425993 - T06028 as required.

Regards

M E Smith

Student nr 72224089

APPENDIX D: PRE TEST / POSTTEST: COMPUTER-BASED TEST**Knowledge questions**

Section 2 | Pre-test

Pre-test

There are 9 questions in this test. Please try and answer all the questions. The questions will be scored, and the mark displayed at the end.

Instructions on how to answer each question are provided on the screen.

Click on the **START** button to begin. There is a time limit of 10 min for completing the test.

You **CANNOT** go back to a previous question.

Start


Image cropped for illustrative purposes

Section 2 | Pre-test

Pre-test

Question 1: Which area of the brain is most directly involved in the reflex control of the autonomic system?

Select the answer in the drop-down list.


Click on the  button to continue.

Cerebellum
Hypothalamus
Medulla Oblongata
Cerebral cortex

Correct answer / Score = 1

Image cropped for illustrative purposes

Time remaining:

Click here when you are finished 

Section 2 | Pre-test

Pre-test


Question 2: Which organs are innervated **mainly** by the sympathetic system?

Select the answer(s) by clicking on the check box next to the option

- Salivary glands
- Stomach glands **Correct answer / Score = 1**
- Sweat glands
- Blood vessels of the skin **Correct answer / Score = 1**
- Pancreas glands

Image cropped for illustrative purposes

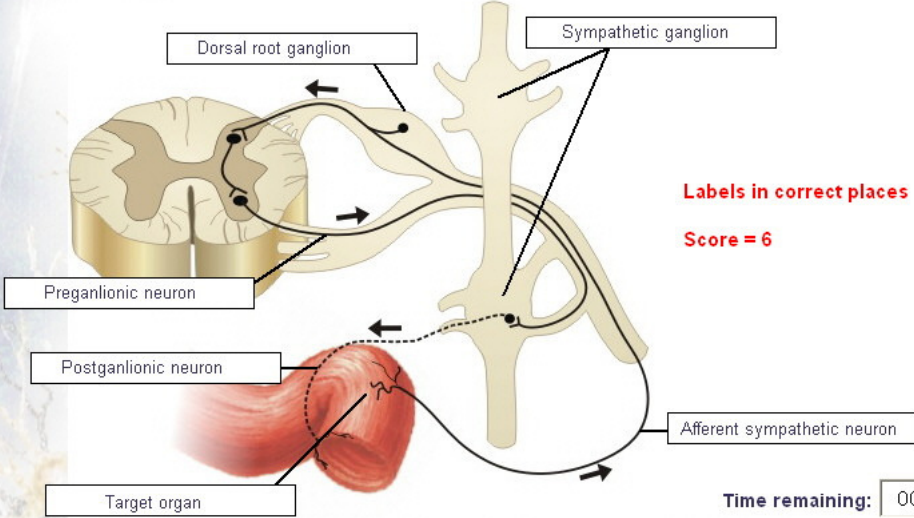
Time remaining: 00:09:41

[Click here when you are finished](#) 

Section 2 | Pre-test


Pre-test

Question 3: Click on the label and drag it to its place on the diagram.



Labels in correct places
Score = 6

Time remaining: 00:09:18

[Click here when you are finished](#) 

Section 2 | Pre-test

Pre-test

Question 4: Which group of receptors are stimulated when the bladder is full?


Select the answer in the drop-down list.

Correct answer / Score = 1

Baroreceptors
Stretch receptors
Volume receptors

Image cropped for illustrative purposes

Time remaining: 00:09:01

Click here when you are finished 

Section 2 | Pre-test

Pre-test

Question 5: Parasympathetic ganglia are located...


Select the answer from the drop down list

Correct answer / Score = 1

in a chain parallel to the spinal cord.
in the dorsal roots of spinal nerves.
next to or within the organs innervated.
in the brain.

Image cropped for illustrative purposes

Time remaining: 00:08:44

Click here when you are finished 

Section 2 | Pre-test

Pre-test

Question 6:
This diagram illustrates

Correct answer / Score = 1

Select the answer in the drop-down list.

Time remaining: 00:08:29

Click here when you are finished

Section 2 | Pre-test

Pre-test

Question 7: Which neurotransmitters are released by the neurons at the synapses in this diagram?

Type on your answer in each field. Use the abbreviations provided in the legend. Use the TAB key to move from one field to the other

Legend

- Ach - Acetylcholine
- Nor - Noradrenaline
- Dop - Dopamine
- NO - Nitrogen Oxide

Time remaining: 00:08:11

Click here when you are finished

Section 2 | Pre-test

Pre-test

Question 8: What effect does increased sympathetic stimulation have on the following organs?

Select the answer from the drop-down list next to each organ

Heart **Correct answer / Score = 1**
 Decreases heart rate

Pupils **Correct answer / Score = 1**
 Constrict

Bladder **Correct answer / Score = 1**
 Wall contracts - bladder empties

Salivary glands **Correct answer / Score = 1**
 No effect whatsoever
 Secretes large amount of watery enzyme-rich saliva

Time remaining: 00:07:57

[Click here when you are finished](#) ▶

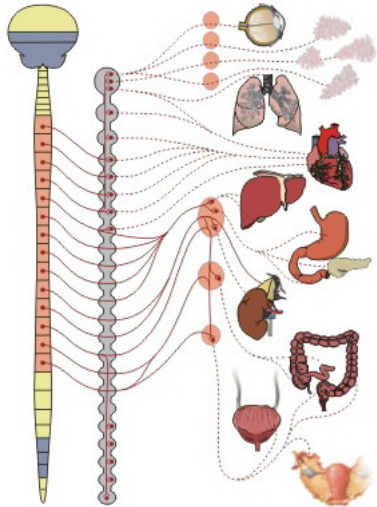
Section 2 | Pre-test

Pre-test

Question 9:

This diagram illustrates the **Correct answer / Score = 1**

Select the answer in the drop-down list.
 Click on the ▶ button to continue



Time remaining: 00:07:02

[Click here when you are finished](#) ▶

Section 2 | Pre-test

Pre-test results

Your score for this pre-test: 11 / 22

You are now going to work through the content of the lesson. You will be assessed again at the end of the lesson. You may go back to previous screens as often as you like in order to master the content. Take your time. You may use the paper provided to make notes - these will be collected before the next assessment.



Click on the  button at the bottom of the screen to continue

Image cropped for illustrative purposes



The Autonomic Nervous System | Menu


Post-test results


Your score for the post-test was: 12 / 22

This is not your final score. The two open-ended questions must still be assessed. I will send the results to your lecturer, if you are interested in knowing what your final score was.

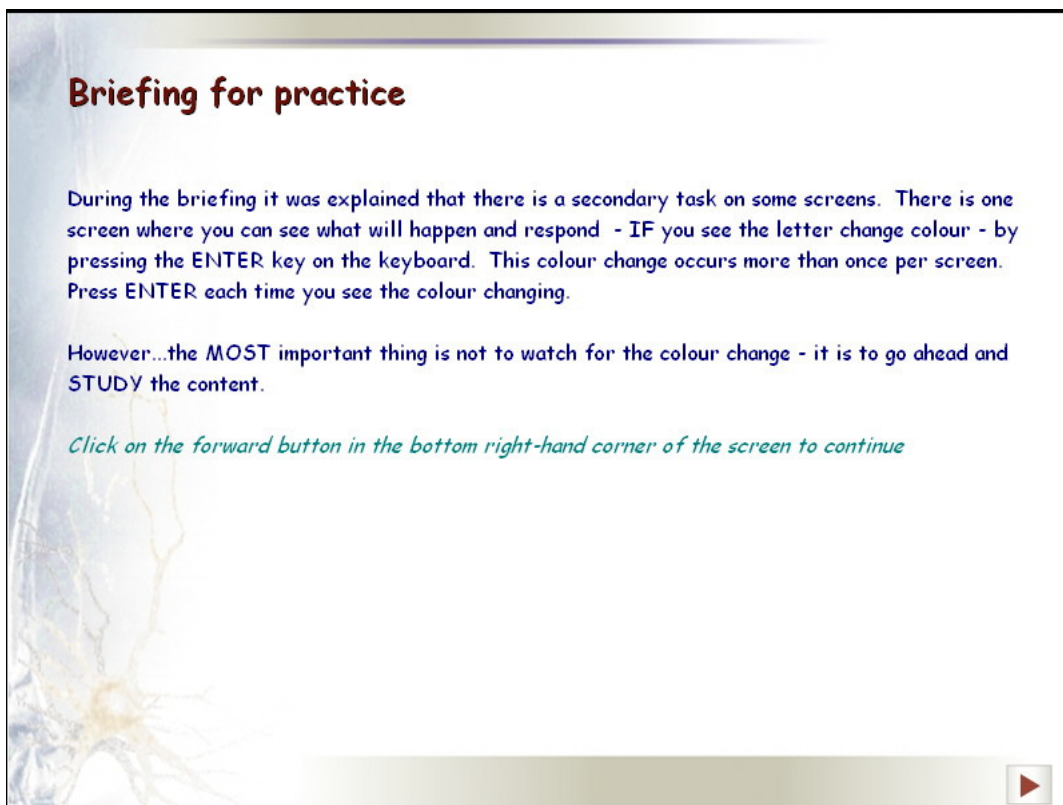
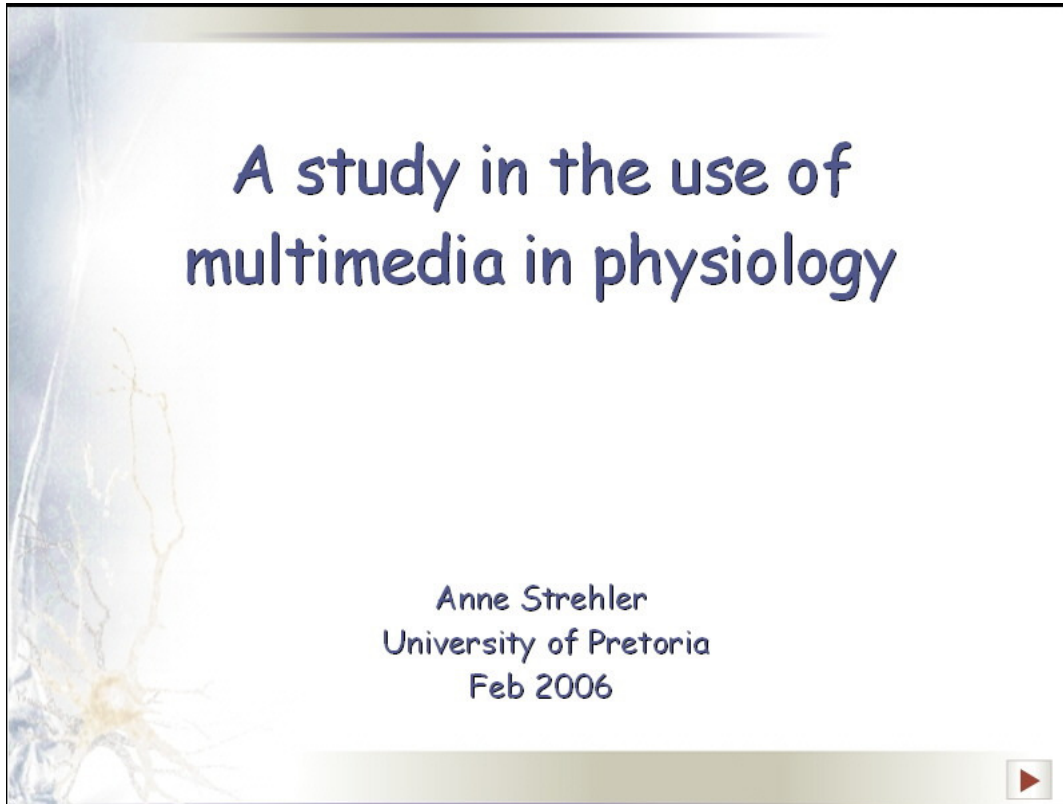
Did this score come close to your estimation at the start of the post-test?

There are another four screens to work through. You need to answer a series of questions about your experience today. Please read the question and the options carefully before your attempt to answer.

Click on the  button at the bottom of the screen to continue



APPENDIX F: TITLE SCREEN AND PRACTICE SESSION SCREEN

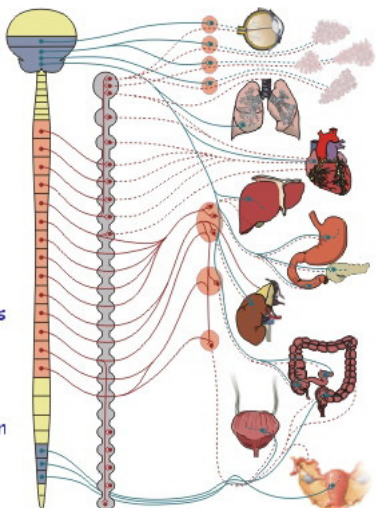


Practice screen

This system was described by Gaskell and Langley around the end of the 19th - beginning of the 20th century. At first it was thought that the autonomic nervous system functioned independently of the central nervous system, but this is not the case.

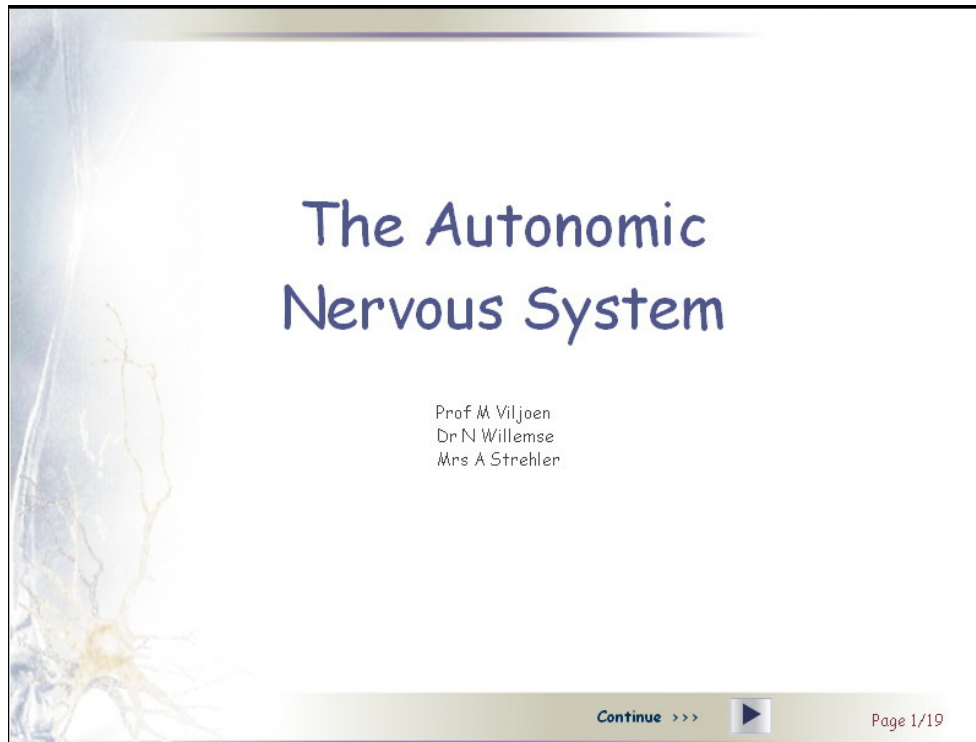
The ANS is regulated by the central nervous system (CNS), especially the hypothalamus. We also know that cognitive and emotive processes from the higher brain centra also influence the ANS. The ANS plays an important role in homeostasis.

The autonomic nervous system is often seen only as an efferent or motor system, but there are afferent or sensory pathways and inflows to the central nervous structures, which are involved in the regulation of the efferent or motor function. We will now look at both these pathways in more detail in the next few screens.

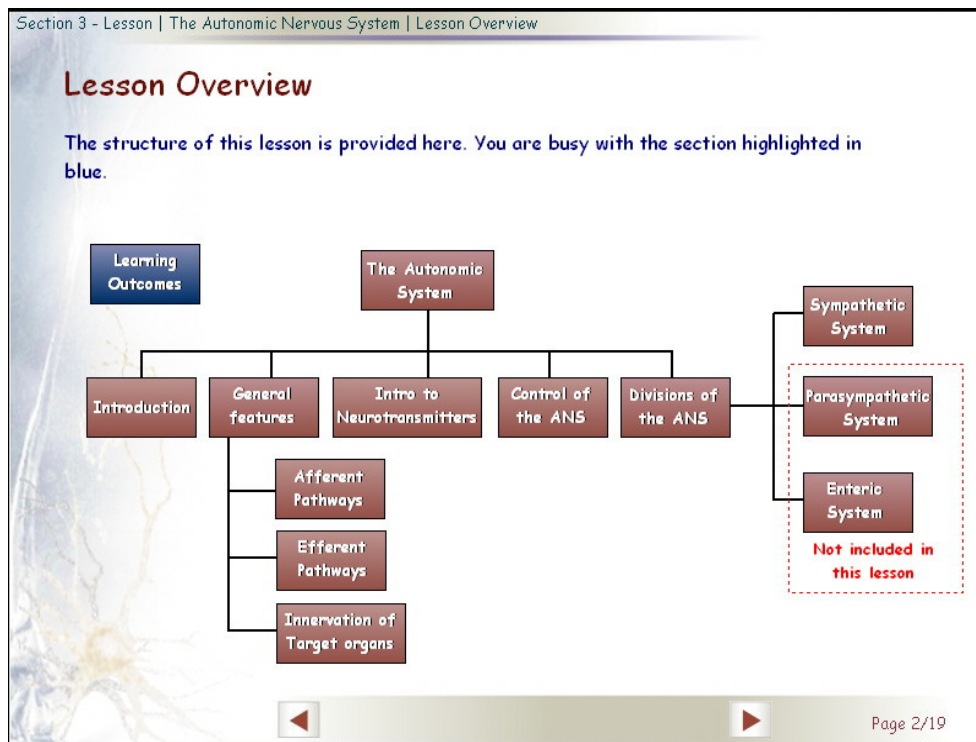


The diagram illustrates the autonomic nervous system (ANS) and its connections to the central nervous system (CNS). On the left, the brain and spinal cord are shown. The brain is divided into the cerebrum (yellow), cerebellum (blue), and brainstem (red). The spinal cord is shown in cross-section, with the brainstem at the top. On the right, various organs are depicted, including the lungs, heart, liver, stomach, intestines, and bladder. Lines represent the pathways of the ANS, showing how the brain and spinal cord regulate these organs. A green letter 'A' is visible in the bottom right corner of the screen.

APPENDIX G: MULTIMEDIA INTERVENTION: SCREENS COMMON TO BOTH FORMATS



Animation version: Screen 1/19 and Static images & text version: Screen 1/23




Animation version: Screen 2/19 and Static images & text version: Screen 2/23



Section 3 - Lesson | The Autonomic Nervous System | Learning Outcomes

Learning outcomes

The learning outcomes for this lesson are to:

- Describe the structure of the Autonomic Nervous System, using basic illustrations.
- Understand the control of the Autonomic Nervous System.
- Compare the structure and function of the Sympathetic and Parasympathetic divisions.
- Describe the function of the Sympathetic Nervous System.
- Describe how the function of the Autonomic reflexes can apply to patient care management for selected problems.

Click on the  button to continue

 Site Map  Page 3/19

Animation version: Screen 3/19 and Static images & text version: Screen 3/23

Section 3 - Lesson | The Autonomic Nervous System | Introduction

Introduction

Click on the underlined text



The Autonomic Nervous System (ANS) is involved in maintaining homeostasis of the internal environment of the body through the regulation and modification of the following activities.

- All involuntary muscle tissue.
- The secretory activity of all exocrine glands.
- Some endocrine glands.
- Some adipose tissue.

The **TWO** main functions of the system are to:

- Regulate the activity of the visceral organs and glandular structure responsible for the basic (vegetative) bodily processes - during rest and digest.
- Respond to a stressor - also known as the typical 'fight or flight' response.

Examples of this type of tissue include the heart muscle, smooth muscle fibers.

 Site Map  **A** Page 4/19

Animation version: Screen 4/19 and Static images & text version: Screen 4/23

Section 3 - Lesson | The Autonomic Nervous System | Introduction

Click on the underlined text and watch the image

The ANS has three parts, the

- The sympathetic nervous system.
- The parasympathetic nervous system.
- The enteric system.

Many of the systems regulated by the ANS receive both excitatory and inhibitory signals from sympathetic and parasympathetic divisions. This provides fine control of the regulated system.

We will look at each of these in turn.

The Enteric Nervous System

Site Map Page 5/19

Animation version: Screen 5/19 and Static images & text version: Screen 5/23

Section 3 - Lesson | The Autonomic Nervous System | Introduction

This system was described by Gaskell and Langley around the end of the 19th - beginning of the 20th century. At first it was thought that the autonomic nervous system functioned independently of the central nervous system, but this is not the case.

The ANS is regulated by the central nervous system (CNS), especially the hypothalamus. We also know that cognitive and emotive processes from the higher brain centra also influence the ANS. The ANS plays an important role in homeostasis.

The autonomic nervous system is often seen only as an efferent or motor system, but there are afferent or sensory pathways and inflows to the central nervous structures, which are involved in the regulation of the efferent or motor function. We will now look at both these pathways in more detail in the next few screens.

Site Map Page 6/19

Animation version: Screen 6/19 and Static images & text version: Screen 6/23

Section 3 - Lesson | The Autonomic Nervous System | Afferent pathways

Afferent pathways

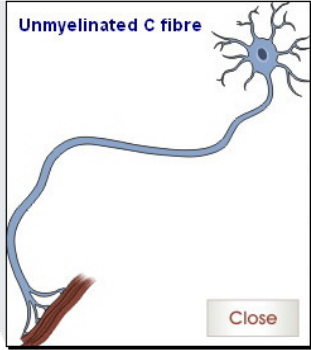
Click on the underlined text

These autonomic fibres are unmyelinated C fibres. Stimulation is both internal and external.

Internal stimulation

The fibres originate in receptors that provide functions in the visceral organs. These receptor groups have further categorisation. This is summarised in the table below.

Pressure changes	Baroreceptors
	Stretch receptors
	Volume receptors
Chemical changes	Chemoreceptors
	Osmoreceptors



Unmyelinated C fibre

Each

Close

◀ Site Map ▶ Page 7/19

Animation version: Screen 7/19 and Static images & text version: Screen 7/23

Section 3 - Lesson | The Autonomic Nervous System | Afferent pathways

Activity: Afferent pathways

Match the example with the type of receptor. Drag the example to the correct place in the table. Click on the Check Answer button to mark your effort

Pressure changes	Baroreceptors	Stimulated when the bowel is full
	Stretch receptors	Decrease in O ₂ saturation of blood
	Volume receptors	Changes in electrolyte balance
Chemical changes	Chemo-receptors	
	Osmo-receptors	

An increase in blood pressure

A full bladder

Check Answer

◀ Site Map ▶ **A** Page 8/19

Animation version: Screen 8/19 and Static images & text version: Screen 8/23

Section 3 - Lesson | The Autonomic Nervous System | Afferent pathways

External stimulation

- Changes in the external environment can also stimulate the afferent pathways - temperature changes, extreme climates and environments e.g. underwater pressure during diving.
- Psychological stressors.

The afferent pathway is illustrated here (using blue).

Labels in diagram: Dorsal root ganglion, Sympathetic ganglion, Afferent sympathetic fibre, Preganglionic neuron, Postganglionic neuron, Target organ.

Navigation: Site Map, Page 9/19

Animation version: Screen 9/19 and Static images & text version: Screen 9/23

Section 3 - Lesson | The Autonomic Nervous System | Efferent pathways

Efferent pathways

Click on the underlined text

The functional unit of the **efferent division** of the autonomic nervous system is a two-neuron motor pathway, which consists of a preganglionic neuron and a postganglionic neuron. The preganglionic fibers are typically myelinated B fibres, whereas the postganglionic fibers are unmyelinated C fibres.

Labels in diagram: Preganglionic neuron, Postganglionic neuron, Vascular smooth muscle.

Myelinated C fibre

Close

Navigation: Map, Page 10/19

Animation version: Screen 10/19 and Static images & text version: Screen 10/23

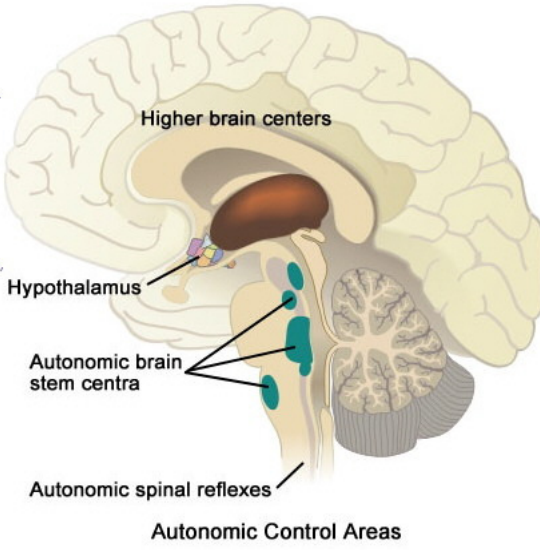
Section 3 - Lesson | The Autonomic Nervous System | Control

Control of the ANS

Click on the underlined text

The activity of the sympathetic nervous system is controlled by:

- autonomic spinal reflexes - first level of control,
- autonomic centra in the brain stem - second level of control,
- the autonomic nuclei of the hypothalamus - third level of control - and
- by higher brain centra.



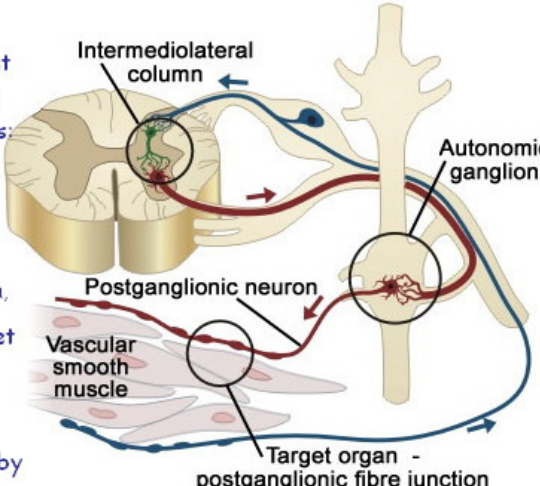
Page 15/19

Animation version: Screen 15/19 and Static images & text version: Screen 22/23

First level of control

Autonomic spinal reflexes represent the first level of control of sympathetic activity. The control centres involved at this level includes:

- the preganglionic sympathetic nuclei in the intermediolateral grey matter,
- the autonomic ganglia in the paravertebral and prevertebral ganglia,
- and even the junction between the target organ and postganglionic fibre.



The first level of control can be influenced by the second level of control.

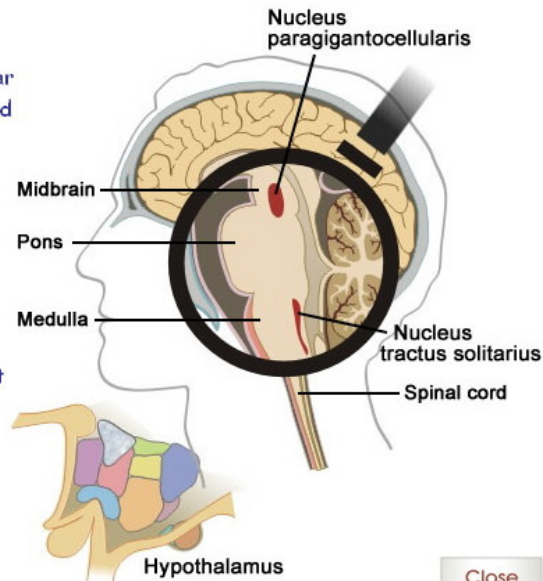
Close

Second level of control

The second level of control, is the reticular formation in the brain stem and spinal cord that include the

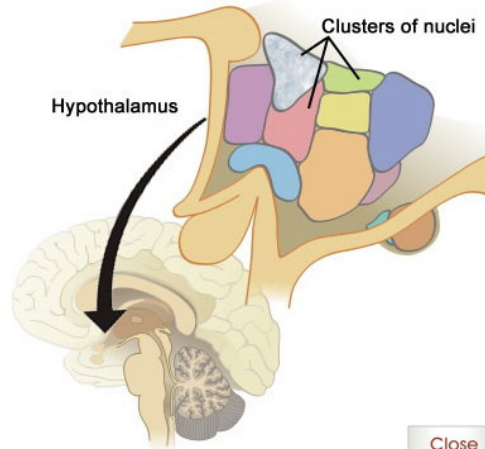
- nucleus tractus solitarius, and the
- nucleus paragigantocellularis.

These autonomic centra of the midbrain, pons, medulla and spinal cord are in turn under the influence of the third level, that is, the hypothalamus.



Third level of control

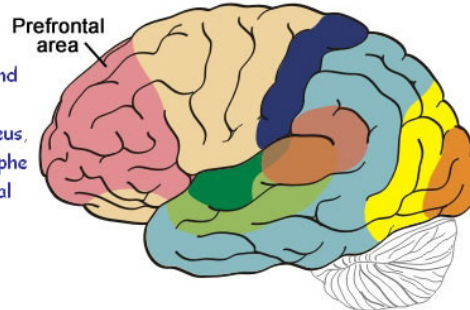
The autonomic nuclei of the hypothalamus represents the third level of control. The hypothalamus can in a sense be seen as the head autonomic ganglion and consists of clusters of nuclei that control endocrine and autonomic nervous system function. The hypothalamus can, however, be influenced by the higher brain centra.



Control by higher brain

Higher brain centres can influence autonomic control by the hypothalamus and brain stem through descending tracts associated with the paraventricular nucleus, the central noradrenergic system, the raphe nuclei and others. Cognitive and emotional events can in this way alter sympathetic activity and chronic stress can through conditions of sustained sympathetic stimulation even shorten the life time of humans and animals.

The degree to which external and cortico-limbic inputs activate the sympathetic system is largely determined by the individual's perception of the situation, which is often a function of previous experience and conditioning.



Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

Innervation of the target organs Still under construction - the link for the GIT, Eyes and Lungs is active

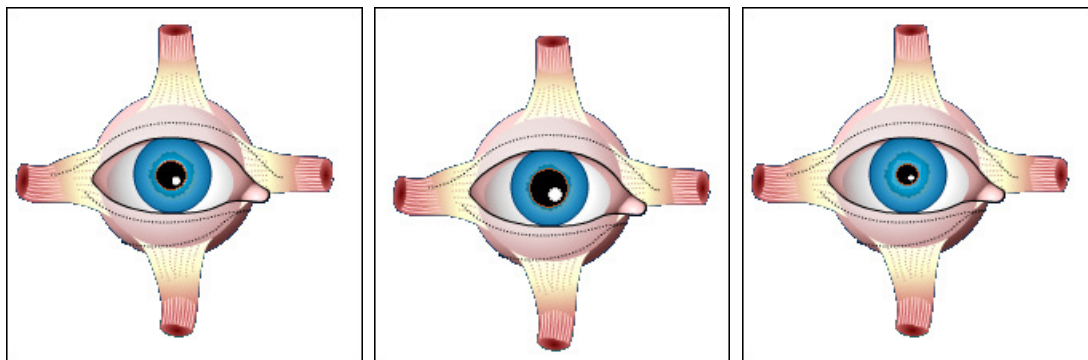
Click on the organ, then click on either the SNS or the PNS button below

Sympathetic Innervation | Parasympathetic Innervation | Reset

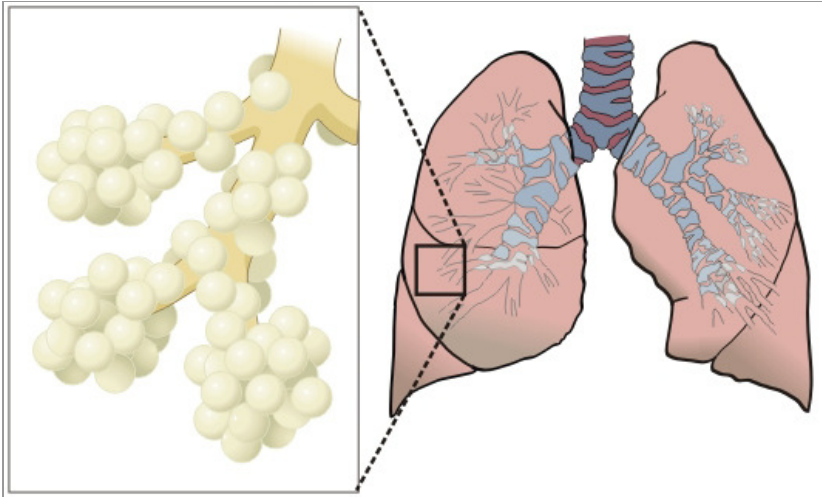
GIT Lungs Eyes Heart Bladder

◀ Site Map ▶ Page 17/19

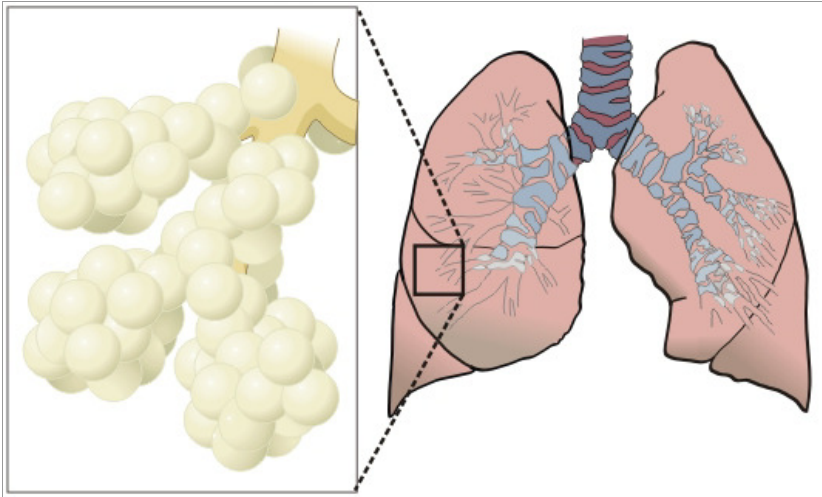
Animation version: Screen 17/19 and Static images & text version: Screen xx/23 – opening view



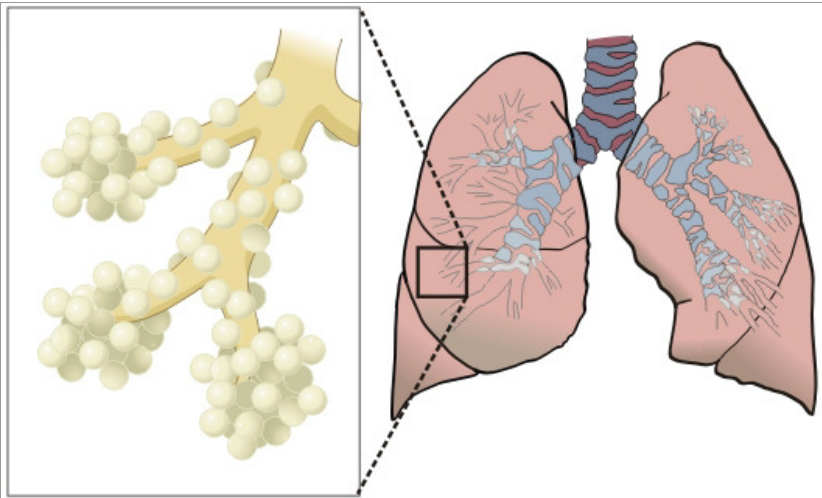
Three views of the innervation of the eye



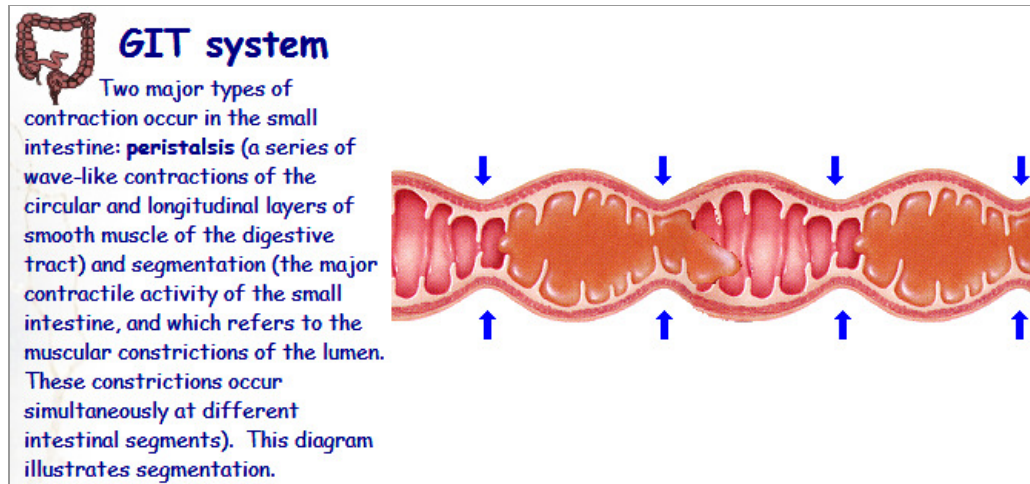
First view of the lung



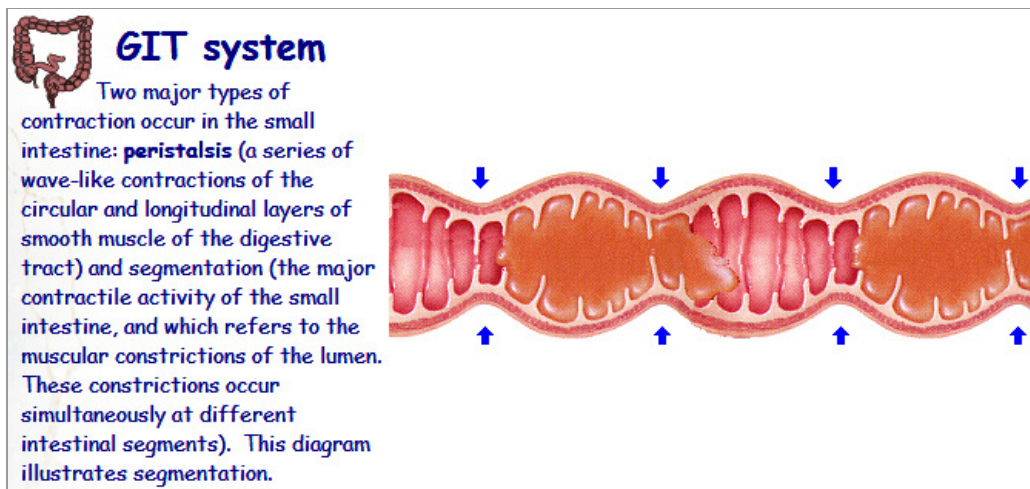
Innervation by the Sympathetic Nervous System



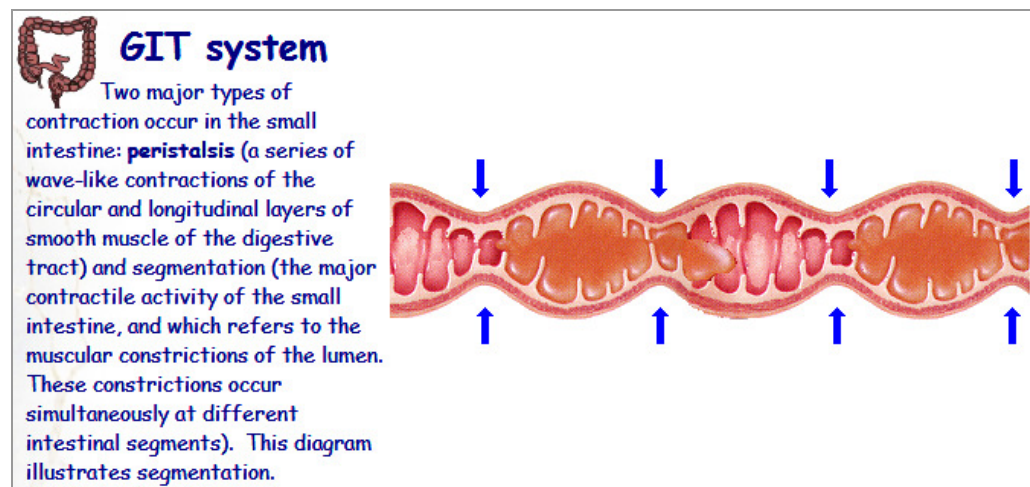
Innervation by the Parasympathetic Nervous System



First view of the GIT system



Innervation by the Sympathetic Nervous System



Innervation by the Parasympathetic Nervous System

APPENDIX H: MULTIMEDIA INTERVENTION: DIFFERENT STRATEGIES TO DISPLAY SAME CONTENT

Animation version – Screen 13

Section 3 - Lesson | The Autonomic Nervous System | Divergence and Convergence

Divergence and convergence

There are two phenomena in the sympathetic nervous system, known respectively as:

- Divergence
- Convergence

Click on each concept for an illustrated explanation

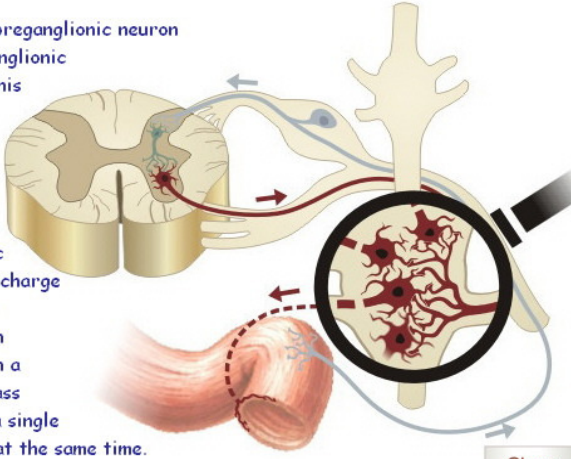
Image cropped for illustrative purposes

◀ Site Map ▶ **A** Page 13/19

Divergence

This is the phenomenon where a single preganglionic neuron synapses with an average of ten postganglionic neurons in the sympathetic ganglion. This **SIMULTANEOUSLY** activates more organ systems. Divergence occurs primarily in the sympathetic nervous system (SNS).

This mass activation of the sympathetic nervous system - the so-called mass discharge or 'fight-or-flight' response - plays an important role in stress responses when there is a need to coordinate changes in a broad range of systems. Because of mass activation the SNS can (activation as a single unit) affect all of its effector organs at the same time.

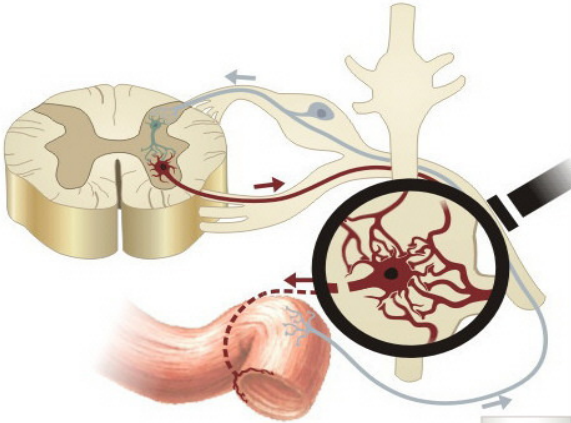


Close

Convergence

This is the phenomenon where many preganglionic neurons converge on one postganglionic cell body in the sympathetic ganglion.

Convergence allows for modulation and integration of control.



Close

Static images & text version – Screen 20

Section 3 - Lesson | The Autonomic Nervous System | Divergence and Convergence

Divergence and convergence

There are two phenomena in the sympathetic nervous system, known respectively as:

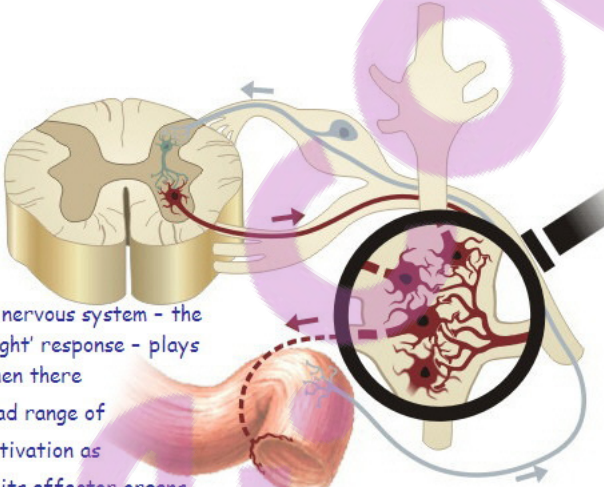
- Divergence
- Convergence

<<<< Click on each concept for an illustrated explanation

Divergence

This is the phenomenon where a single preganglionic neuron synapses with an average of ten postganglionic neurons in the sympathetic ganglion. This **SIMULTANEOUSLY** activates more organ systems. Divergence occurs primarily in the SNS.

This mass activation of the sympathetic nervous system - the so-called mass discharge or 'fight-or-flight' response - plays an important role in stress responses when there is a need to coordinate changes in a broad range of systems. Because of mass activation (activation as a single unit) the SNS can affect all of its effector organs at the same time.



◀ Site Map ▶ **A** Page 20/23

Section 3 - Lesson | The Autonomic Nervous System | Divergence and Convergence

Divergence and convergence

There are two phenomena in the sympathetic nervous system, known respectively as:

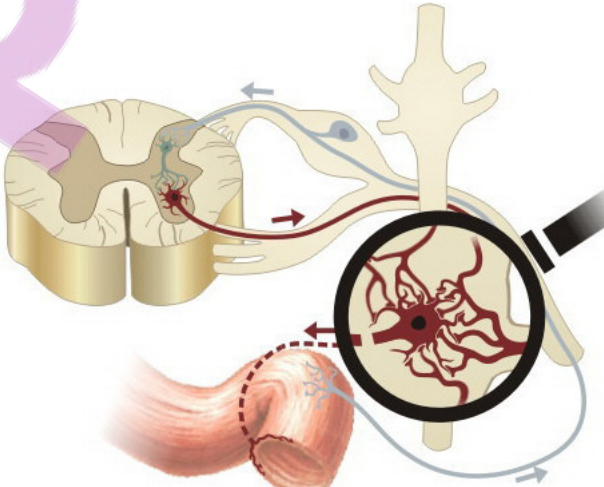
- Divergence
- Convergence

<<<< Click on each concept for an illustrated explanation

Convergence

This is the phenomenon where many preganglionic neurons converge on one postganglionic cell body in the sympathetic ganglion.

Convergence allows for modulation and integration of control.



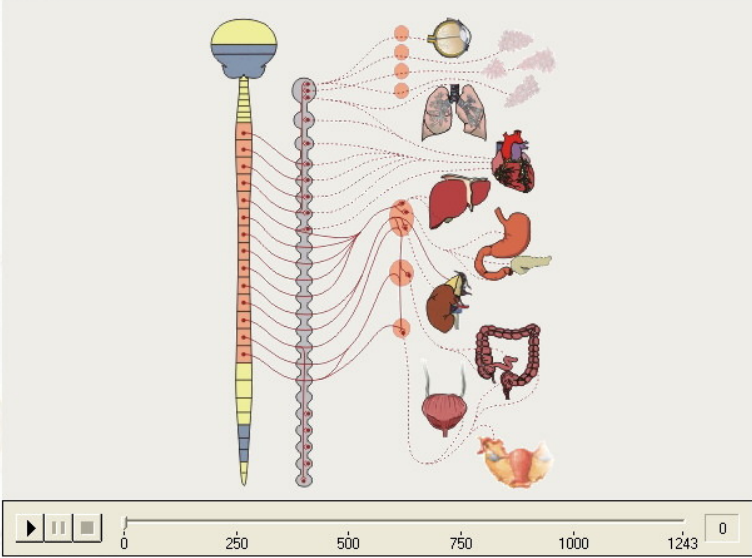
◀ Site Map ▶ **A** Page 20/23




APPENDIX I: MULTIMEDIA: ANIMATION VERSUS STATIC IMAGES

Animation version: Screen 12 (Animation)

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System

Structure of the system This is a narrated animation - listen using the earphones



Click on the  button to play the animation  [Site Map](#)  **A** Page 12/19

Static images & text version: Screens 12 – 16

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system


This section will take you through a step-by-step, illustrated explanation of the structure of the sympathetic nervous system. There are a total of 7 screens. You need to work through all 7 screens, and return to this screen to continue.


The content, displayed schematically here, is explained in three sections:

- The origins of the efferent fibres
- The ganglia and synaptic connections
- The effector organs

Follow the instructions on the screen. Explore the content on your own as well. Rolling the mouse over the images will often display more information.

Click on the white text below to get to the different sections. When you have completed a section you will see the word 'Completed' under the relevant section. Once all are complete you can use the forward button to move on in the lesson



 [Site Map](#) Page 12/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

You are studying the section highlighted in blue

The preganglionic neurons originate from the autonomic centres located in the pars intermedia of the grey matter of the thoracic and upper two lumbar segments of the spinal cord.

Roll mouse over this area of the spinal cord T1 - T12 for more detail

Page 13/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

You are studying the section highlighted in blue

The preganglionic neurons originate from the autonomic centres located in the pars intermedia of the grey matter of the thoracic and upper two lumbar segments of the spinal cord.

Roll mouse over this area of the spinal cord T1 - T12 for more detail

Intermediolateral column

Spinal cord

Page 13/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

Origins →
 Synapses & ganglia →
 Effector organs

You are studying the section highlighted in blue

The preganglionic neurons which originate from the CNS terminate in autonomic ganglia where they form synaptic connections with neurons in the ganglia. These autonomic ganglia are located in three areas - illustrated here in the diagram and on the next screen.

Roll mouse over the diagram for more detail

Paravertebral chain ganglia

▶ Page 14/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

Origins →
 Synapses & ganglia →
 Effector organs

You are studying the section highlighted in blue

The preganglionic neurons which originate from the CNS terminate in autonomic ganglia where they form synaptic connections with neurons in the ganglia. These autonomic ganglia are located in three areas - illustrated here in the diagram and on the next screen.

Roll mouse over the diagram for more detail

2

The preganglionic fibres supplying the digestive tract do **not** relay in the sympathetic chain but pass right through and synapse in collateral ganglia in nerve plexuses on the arteries supplying the intestines. There are three large ganglia:

- Celiac ganglion
- Superior mesenteric ganglion
- Inferior mesenteric ganglion

Paravertebral chain ganglia

▶ Page 14/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

Origins →
 Synapses & ganglia →
 Effector organs

You are studying the section highlighted in blue

The preganglionic neurons which originate from the CNS terminate in autonomic ganglia where they form synaptic connections with neurons in the ganglia. These autonomic ganglia are located in three areas - illustrated here in the diagram and on the next screen.

Roll mouse over the diagram for more detail

1

The **sympathetic ganglia**. These are **two** paravertebral or sympathetic chains, one on each side of the spinal cord. Each chain consists of about **22 ganglia**. Most of the preganglionic fibres terminate in neurons within the ganglia, but some may pass up or down the chain before establishing synaptic connections.

Paravertebral chain ganglia

▶ Page 14/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

Origins →
 Synapses & ganglia →
 Effector organs

You are studying the section highlighted in blue

In this example.....

T12

When it enters the region of the chain of autonomic ganglia it passes downwards in this chain.....

Reset
▶

▶ Page 16/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

A

You are studying the section highlighted in blue

The preganglionic fibres (illustrated in green) leave the spinal cord and pass directly to the **medulla of the adrenal gland**, which can be likened to a modified sympathetic ganglion. This is the third area where there is a synaptic connection. The cells of the adrenal medulla (illustrated in the magnified area) are innervated by the preganglionic sympathetic fibers.

Page 15/23

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system

This section will take you through a step-by-step, illustrated explanation of the structure of the sympathetic nervous system. There are a total of 7 screens. You need to work through all 7 screens, and return to this screen to continue.

The content, displayed schematically here, is explained in three sections:

- The origins of the efferent fibres
- The ganglia and synaptic connections
- The effector organs

Follow the instructions on the screen. Explore the content on your own as well. Rolling the mouse over the images will often display more information.

Click on the white text below to get to the different sections. When you have completed a section you will see the word 'Completed' under the relevant section. Once all are complete you can use the forward button to move on in the lesson

Site Map

Page 12/23

APPENDIX J: MULTIMEDIA INTERVENTION: WHOLE VIEW VERSUS PARTS VIEW

Animation version – Screen 18

Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

Summary of the Autonomic Nerve Supply and effects on visceral organs

A. Organs with dual supply		
Organ	Sympathetic effects	Parasympathetic effects
Iris	Pupil dilation	Pupil constriction
Bronchioles	Dilation	Constriction
Heart	Rate ↑	Rate ↓
Digestive system	Motility ↓	Motility ↑
Sphincters of digestive tract	Contraction	Relaxation
Urinary bladder	Relaxation of wall → filling	Contraction of wall → emptying
B. Organs with mainly sympathetic supply		
Organ	Absence of sympathetic activity	Increased sympathetic activity
Blood vessels of skin	Vasodilation	Vasoconstriction
Sweat glands	None	Secretion
C. Organs with mainly parasympathetic supply		
Salivary glands	Secretion of large amounts of watery enzyme-rich saliva. (Stimulation of sympathetic fibres produce small amounts of a thick, viscous secretion)	
Stomach glands	Secretion of enzyme-rich gastric juice	
Pancreas glands	Secretion of enzyme-rich pancreatic juices	
Sex organs	Vasodilation in erectile tissue → erection	

◀ Site Map ▶ Page 18/19

All information presented in one table

Static images & text version – Screen 19

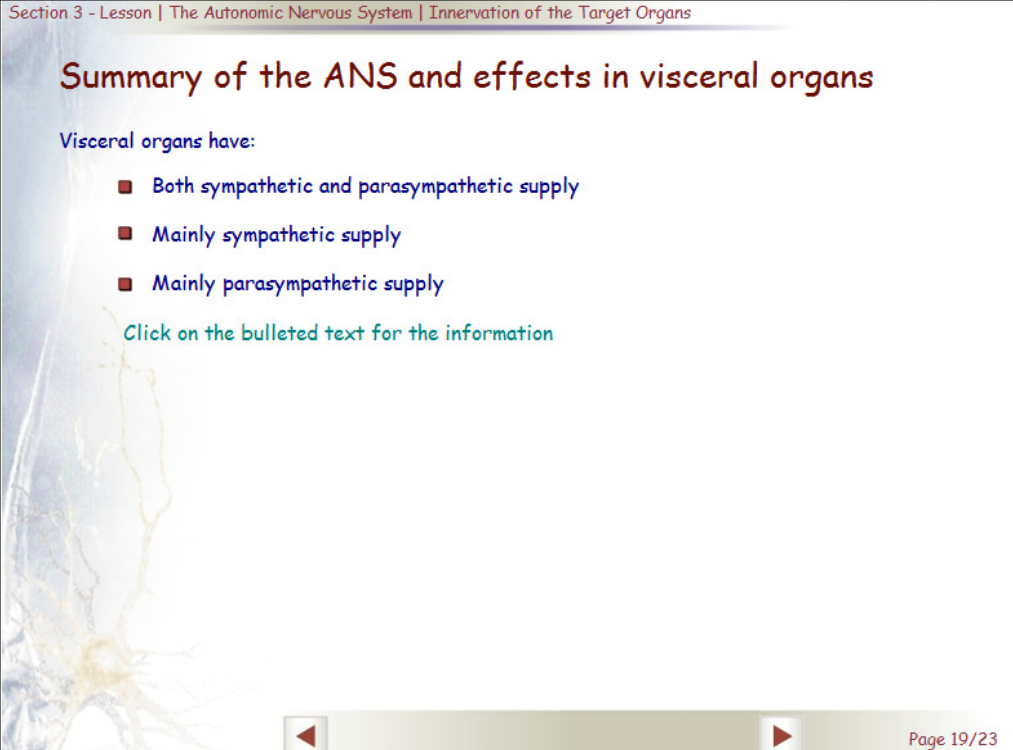
Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

Summary of the ANS and effects in visceral organs

Visceral organs have:

- Both sympathetic and parasympathetic supply
- Mainly sympathetic supply
- Mainly parasympathetic supply

Click on the bulleted text for the information



Page 19/23

Information accessed by clicking on bulleted text

Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

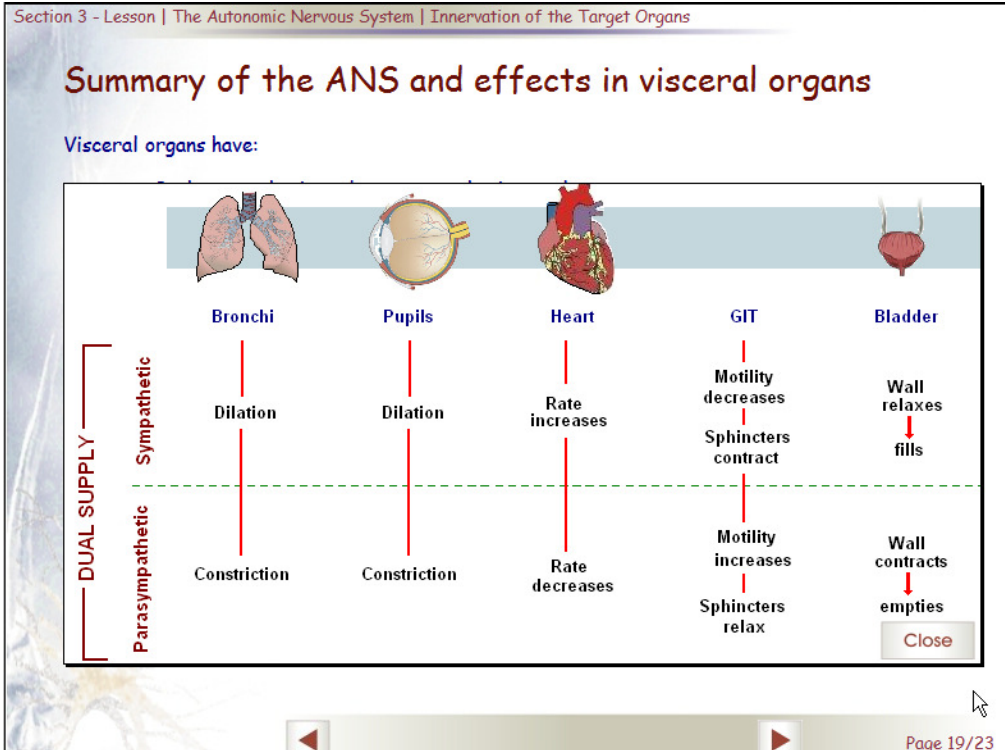
Summary of the ANS and effects in visceral organs

Visceral organs have:

	Bronchi	Pupils	Heart	GIT	Bladder
Sympathetic	Dilation	Dilation	Rate increases	Motility decreases Sphincters contract	Wall relaxes fills
Parasympathetic	Constriction	Constriction	Rate decreases	Motility increases Sphincters relax	Wall contracts empties

DUAL SUPPLY

Close



Page 19/23

Information provided in smaller chunks with text and visual material

Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

Summary of the ANS and effects in visceral organs

Visceral organs have:

- Both sympathetic and parasympathetic supply
- Mainly sympathetic
- Mainly parasympathetic

Click on the diagram to view the effects of sympathetic and parasympathetic activity on visceral organs.

	Blood vessels	Sweat glands
MAINLY Sympathetic Absence of sympathetic activity	Vasodilate	No secretion
MAINLY Parasympathetic Increased sympathetic activity	Vasoconstrict	Secretion

Close

Page 19/23

Second pop-up for screen 19

Section 3 - Lesson | The Autonomic Nervous System | Innervation of the Target Organs

Summary of the ANS and effects in visceral organs

Visceral organs have:

- Both sympathetic and parasympathetic supply
- Mainly sympathetic
- Mainly parasympathetic

Click on the diagram to view the effects of sympathetic and parasympathetic activity on visceral organs.

	Salivary glands	Stomach glands	Pancreas	Sex organs
MAINLY Sympathetic	Secrete large amounts of watery enzyme-rich saliva	Secretion of enzyme-rich gastric juice	Secretion of enzyme-rich pancreatic juices	Vasodilation in erectile tissue Erection
MAINLY Parasympathetic	Stimulation of sympathetic fibres produces small amounts of a thick, viscous secretion			

Close

Page 19/23

Third pop-up for screen 19

APPENDIX K: A SUMMARY OF THE MULTIMEDIA AND THE INTEGRATION OF THE RESEARCH INSTRUMENTS


Instrument	Section in multimedia	Appendix	It will measure...	Development	No of times administered	Source	Alignment with research sub-question / Purpose of this instrument
Cognitive Style Analysis (CSA)	Not included	Not included	Cognitive style	Existing instrument	Once	License to use CSA purchased from Learning & Training Technology, UK	What are the cognitive styles of the participants taking part in the study? What is the relationship between cognitive style and cognitive load when learning with multimedia?
Practice secondary task	1	Appendix G		Self-developed	Once – at beginning of multimedia	Not applicable	This was included to expose the participants to the direct measurement technique for measuring cognitive load. This was an attempt to control the potential extraneous cognitive load caused by this secondary task.
Demographic questionnaire	2	Appendix M	Demographic profile of sample	Self-developed	Once	Not applicable	Provides data for other potential covariates.
Self-report rating scale	2	Appendix N	Self-rating of Verbaliser-Imager dimension of cognitive style	Modified an existing instrument	Once	Adaptation of Mayer's instrument (Mayer & Massa, 2003)	What are the cognitive styles of the participants taking part in the study? The data collected will be used to determine the correlation between self-report measures of the style dimension and the measure obtained using Riding's Cognitive Style Analysis.

Instrument	Section in multimedia	Appendix	It will measure...	Development	No of times administered	Source	Alignment with research sub-question / Purpose of this instrument
Pretest	3	Appendix E	Prior knowledge of learning outcomes. First step in assessing extent of the learning gain.	Self-developed	Once	Physiology textbooks Validated by subject matter experts for content validity	How do participants with different cognitive styles perform when using the same content with different cognitive load?
Self-rating questionnaire	4	See Figure 3.3 on page 137	Self-report rating of mental load	Embedded at selected points in the program. Existing instrument	Version 1 Five (5) times Version 2 Six (6) times	Developed by Paas (1994)	What is the correlation between the participant's self-report of cognitive load and the direct measure of the cognitive load of the content?
Direct measurement technique	4	See Figure 3.9 and 3.10 on page 162	Cognitive load	Instrument embedded at selected points in multimedia program. Protocol modified for Smith's study (2007). Basic principles not changed.	Version 1 Eleven (11) times. Version 2 Thirteen (13) times.	Method described by Brünken, Plass & Leutner (2003).	Which presentation format was instructionally more efficient? To what extent do the presentation formats influence cognitive load? What is the relationship between cognitive style and cognitive load when learning with multimedia?
Posttest - Section 1	5	Appendix F	Knowledge and achievement of learning outcomes.	Self-developed	Once	Same test as pretest	How do participants with different cognitive styles perform when using the same content with different cognitive load?
Posttest - Section 2	Not included in multimedia		Test ability to apply knowledge.	Self-developed - pencil and paper test	Once	Validated by subject matter experts for content validity	Which presentation format was instructionally more efficient?

APPENDIX L: ELECTRONIC QUESTIONNAIRE TO COLLECT DEMOGRAPHIC DATA

Section 1 | Questionnaire - Personal Data

I need some personal information

Please complete all the fields and click on the  button when you are finished. Use the TAB key to move from field to field. Select the most appropriate options from the drop-down lists.

Personal information

Age (enter only the number e.g.24) :

Gender:

Which cultural group best describes you?:
(Select the option from the drop down list)

Information about your course


What programme are you registered for (e.g. MBChB):

What is your year of study?(enter only the number e.g. 3) :

Page 1/4

Section 1 | Questionnaire - Personal Data

I need some personal information

Please complete all the fields and click on the  button when you are finished. Use the TAB key to move from field to field. Select the most appropriate options from the drop-down lists.

Information about the lesson

Have you previously studied the physiology of the Autonomic Nervous System? (Y for Yes, N for No)

Rate your own knowledge and understanding of the physiology of the Autonomic Nervous System

I think I know & understand
(select the best option from the drop-down list)

Information about your language skills:

English is my:
(select the best option from the drop down list)

 language.

Page 2/4

APPENDIX M: ALLOCATION OF GROUPS FOR THE MAIN STUDY

	Animation version 10:30 – 12:30		Static images & text version 12:30 – 14:30	
Session 1	HWS Lab Lesson 1 Group 1 Sample size: n = 17	HWS Lab Lesson 2 Group 16 Sample size: n = 17	HWS Lab Lesson 3 Group 6 Sample size: n = 17	HWS Lab Lesson 4 Group 9 Sample size: n = 16
	BMW Lab Lesson 1 Group 2 Sample size: n = 16	BMW Lab Lesson 2 Group 15 Sample size: n = 16	BMW Lab Lesson 3 Group 5 Sample size: n = 16	BMW Lab Lesson 4 Group 10 Sample size: n = 16
Session 2	HWS Lab Lesson 1 Group 3 Sample size: n = 16	HWS Lab Lesson 2 Group 14 Sample size: n = 16	HWS Lab Lesson 3 Group 7 Sample size: n = 16	HWS Lab Lesson 4 Group 12 Sample size: n = 16
	BMW Lab Lesson 1 Group 4 Sample size: n = 16	BMW Lab Lesson 2 Group 13 Sample size: n = 17	BMW Lab Lesson 3 Group 8 Sample size: n = 16	BMW Lab Lesson 4 Group 11 Sample size: n = 17

**APPENDIX N: EXAMPLE OF DATA FOR A SINGLE PARTICIPANT - WRITTEN
OUT TO AN .INI FILE**

[Bookmark]

Last screen=\latt04

[Accessed Lesson Date]

Date=02/27/06

[Accessed Lesson Time]

Time=10:59:46

[Exited Lesson Time]

Time=11:54:57

[XXXXXXXX] - Student number removed to protect identity of participant

02/27/06

Student No V1=**XXXXXXXX - Student number removed to protect identity of participant**

[Demographic Data2]

Version V2=1

[Demographic Data3]

Age V3=18

[Demographic Data4]

Gender V4=2

[Demographic Data5]

Culture V5=3

[Course Detail 1]

Programme V6=MBCHB

[Course Detail 2]

Year of study V7=2

[Lesson Detail]

Prior Know V8=0

[Self rating]

Self rating V9=2

[Language]

Language V10=

Language V10=2

[SBCSQ1]

SBCSQ1 V11=5

[SBCSQ2]

SBCSQ2 V12=4

[SBCSQ3]

SBCSQ3 V13=4

[SBCSQ4]

SBCSQ4 V14=4

[SBCSQ5]
SBCSQ5 V15=5

[SBCSQ6]
SBCSQ6 V16=5

[SBCSQ7]
SBCSQ7 V17=3

[ScreenData -Screen 1]
Screen No=Using this multimedia

[ScreenData -Screen 1_2]
Date in=02/27/06

[ScreenData -Screen 1_3]
Time in=10:44:23

[ScreenData -Screen 1_4]
Time out=10:45:50

[ScreenData -Screen 2]
Screen No=Using this multimedia 2

[ScreenData -Screen 2_2]
Date in=02/27/06

[ScreenData -Screen 2_3]
Time in=10:45:50

[ScreenData -Screen 2_4]
Time out=10:47:33

[Pretest Q1]
V18=0

[Pretest Q2.1]
V19=1

[Pretest Q2.2]
V20=1

[Pretest Q3.1]
V21=1

[Pretest Q3.2]
V22=1

[Pretest Q3.3]
V23=1

[Pretest Q3.4]
V24=1

[Pretest Q3.5]
V25=1

[Pretest Q3.6]
V26=1

[Pretest Q4]

V27=1

[Pretest Q5]
V28=1

[Pretest Q6]
V29=0

[Pretest Q7.1]
V30=0

[Pretest Q7.2]
V31=0

[Pretest Q7.3]
V32=1

[Pretest Q7.4]
V33=0

[Pretest Q7.5]
V34=0

[Pretest Q8.1]
V35=1

[Pretest Q8.2]
V36=1

[Pretest Q8.3]
V37=1

[Pretest Q8.4]
V38=0

[Pretest Q9]
V39=0

[Pretest Total]
Total=

[CL01]
Cognitive load V40=7

[CL02]
Cognitive load V41=8

[CL03]
Cognitive load V42=6

[CL04]
Cognitive load V43=.

[CL05]
Cognitive load V44=8

[CL06]
Cognitive load V45=not tested

[CL07]
Cognitive load V46=not tested

[Trigger1_1]
Trigger1=11:01:0411:03:56

[Trigger1_2]
Trigger2=11:01:1411:04:06

[Trigger1_3]
Trigger3=11:01:2411:04:16

[Trigger1_4]
Trigger4=11:01:3411:04:26

[Trigger2_1]
Trigger1=11:03:4811:05:28

[Trigger2_2]
Trigger2=11:05:38

[Trigger2_3]
Trigger3=11:05:48

[Trigger3_1]
Trigger1=11:12:15

[Trigger3_2]
Trigger2=11:12:25

[Trigger3_3]
Trigger3=11:12:35

[Trigger3_4]
Trigger4=11:12:44

[Trigger4_1]
Trigger1=11:16:48

[Trigger4_2]
Trigger2=11:16:58

[Trigger4_3]
Trigger3=11:17:08

[Trigger5_1]
Trigger1=11:19:22

[Trigger5_2]
Trigger2=11:19:32

[Trigger5_3]
Trigger3=11:19:42

[Trigger6_1]
Trigger1=11:22:03

[Trigger6_2]
Trigger2=11:22:13

[Trigger6_3]
Trigger3=11:22:23

[Trigger6_4]
Trigger4=11:22:33

[Trigger6_5]
Trigger5=11:22:43

[Trigger6_6]
Trigger6=11:22:53

[Trigger6_7]
Trigger7=11:23:03

[Trigger6_8]
Trigger8=11:23:13

[Trigger7_1]
Trigger1=11:31:09

[Trigger7_2]
Trigger2=11:31:19

[Trigger7_3]
Trigger3=11:31:29

[Trigger7_4]
Trigger4=11:31:39

[Trigger7_5]
Trigger5=11:31:49

[Trigger8_1]
Trigger1=11:32:52

[Trigger8_2]
Trigger2=11:33:02

[Trigger8_3]
Trigger3=11:33:12

[Trigger8_4]
Trigger4=11:33:22

[Trigger9_1]
Trigger1=11:46:12

[Trigger9_2]
Trigger2=11:46:22

[Trigger9_3]
Trigger3=11:46:32

[Trigger9_4]
Trigger4=11:46:42

[Trigger9_5]
Trigger5=11:46:53

[Trigger9_6]
Trigger6=11:46:52

[Trigger10_1]

Trigger1=11:48:0611:49:43

[Trigger10_2]
Trigger2=11:49:53

[Trigger10_3]
Trigger3=11:50:03

[Trigger10_4]
Trigger4=11:50:13

[Trigger10_5]
Trigger5=11:50:23

[Trigger10_6]
Trigger6=11:50:33

[Trigger10_7]
Trigger7=11:50:43

[Trigger10_8]
Trigger8=11:50:53

[Meta Cog 1]
V47=3

[Meta Cog 2]
V48=.

[Meta Cog 3]
V49=19

[Posttest Q1]
V50=0

[Posttest Q2.1]
V51=1

[Posttest Q2.2]
V52=1

[Posttest Q3.1]
V53=1

[Posttest Q3.2]
V54=1

[Posttest Q3.3]
V55=1

[Posttest Q3.4]
V56=1

[Posttest Q3.5]
V57=1

[Posttest Q3.6]
V58=1

[Posttest Q4]
V59=0

[Posttest Q5]
V60=0

[Posttest Q6]
V61=1

[Posttest Q7.1]
V62=1

[Posttest Q7.2]
V63=1

[Posttest Q7.3]
V64=1

[Posttest Q7.4]
V65=1

[Posttest Q7.5]
V66=0

[Posttest Q8.1]
V67=1

[Posttest Q8.2]
V68=1

[Posttest Q8.3]
V69=1

[Posttest Q8.4]
V70=0

[Posttest Q9]
V71=1

[Posttest Open1]
V72=.

[Posttest Open 2]
V73=.

[Survey S1.1]
V74=4

[Survey S1.2]
V75=5

[Survey S1.3]
V76=4

[Survey S1.4]
V77=5

[Survey S1.5]
V78=5

[Survey S1.6]
V79=5

[Survey S1.7]
V80=5

[Survey S2.1]
V81=5

[Survey S2.2]
V82=4

[Survey S2.3]
V83=6

[Survey S2.4]
V84=5

[Survey S2.5]
V85=5

[Survey S2.6]
V86=5

[Survey S2.7]
V87=4

[Survey S3.1]
V88=5

[Survey S3.2]
V89=5

[Survey S3.3]
V90=5

[Survey S3.4]
V91=5

[Survey S4.1]
V92=5

[Survey S4.2]
V93=3

[Survey S4.3]
V94=3

[Survey S4.4]
V95=3

[Survey S4.5]
V96=2

[Survey S4.6]
V97=5

[Survey S4.7]
V98=4

[Trigger]
Trigger1=10:41:18
Trigger2=10:41:28
Trigger3=10:41:38

Trigger4=10:41:48

[sScreen01]
10:59:51=Exited

[sScreen02]
10:59:51=Accessed
11:00:35=Exited

[sScreen03]
11:00:35=Accessed
11:01:00=Exited

[sScreen04]
11:01:00=Accessed
11:03:44=Exited
11:03:52=Accessed
11:05:24=Exited

[Responses1_2]
11:01:14=Hit space bar
11:04:08=Hit space bar

[Responses1_3]
11:01:25=Hit space bar
11:04:17=Hit space bar
11:04:18=Hit space bar

[Responses1_4]
11:01:34=Hit space bar
11:04:27=Hit space bar

[sScreen05]
11:03:44=Accessed
11:03:52=Exited
11:05:24=Accessed
11:08:05=Exited

[Responses1_1]
11:03:58=Hit space bar

[sScreen06]
11:08:21=Accessed
11:08:27=Exited
11:08:30=Accessed
11:08:38=Exited

[sScreen07]
11:08:27=Accessed
11:08:30=Exited
11:08:38=Accessed
11:12:10=Exited

[sScreen08]
11:12:10=Accessed
11:14:42=Exited

[Responses3_2]
11:12:25=Hit space bar

[Responses3_3]

11:12:36=Hit space bar

[Responses3_4]

11:12:46=Hit space bar

[sScreen09]

11:14:42=Accessed

11:16:43=Exited

[sScreen10]

11:16:44=Accessed

11:19:18=Exited

[sScreen11]

11:19:18=Accessed

11:21:59=Exited

[Responses5_2]

11:19:34=Hit space bar

[Responses5_3]

11:19:44=Hit space bar

[sScreen12v1]

11:21:59=Accessed

11:30:56=Exited

[sScreen13]

11:31:04=Accessed

11:32:42=Exited

[sScreen14]

11:32:48=Accessed

11:37:05=Exited

[Responses8_3]

11:33:14=Hit space bar

[Responses8_4]

11:33:23=Hit space bar

[sScreen15]

11:37:05=Accessed

11:45:37=Exited

[sScreen16]

11:45:37=Accessed

11:46:07=Exited

11:47:18=Accessed

11:47:45=Exited

[sScreen17]

11:46:08=Accessed

11:47:18=Exited

11:47:45=Accessed

11:47:46=Exited

[sScreen18]

11:47:46=Accessed

11:47:55=Exited

11:48:09=Accessed

11:49:39=Exited

[sScreen19]

11:48:02=Accessed

11:48:09=Exited

11:49:39=Accessed

11:54:17=Exited

[Responses10_1]

11:48:08=Hit space bar

[sScreen20]

11:54:57=Exited

[Survey info]

12:02:58=Accessed

12:03:06=Exited

[Survey S1]

12:03:06=Accessed

12:03:46=Exited

[Survey S2]

12:03:46=Accessed

12:04:26=Exited

[Survey S3]

12:04:26=Accessed

12:04:49=Exited

[Survey S4]

12:04:49=Accessed

12:05:34=Exited

[Thanks]

12:05:34=Accessed

12:05:52=Exited

[Screen logs]

(Screen_Log)= \practice \2 \3 \01 \02 \03 \04 \05 \04 \05 \06 \07 \06 \07 \08 \09 \10 \11 \12v1 \13
\14 \15 \16 \17 \16 \17 \18 \19 \18 \19 \end lesson \final surv \exit

APPENDIX O: STUDENT HANDOUT - PARTICIPATION IN THE STUDY

Letter of consent

The relationship between cognitive load, cognitive style and multimedia learning

A study in the use of multimedia in physiology

Dear Participant

You are invited to participate in a research project aimed at exploring the role which cognitive load and cognitive style play in the successful achievement of learning outcomes when using animations as multimedia learning resources within the higher education sector. The study will also investigate the interrelationship between cognitive load, which is influenced by both the nature of the content and the specific design strategies used, and the cognitive style of the person using the multimedia.

Your participation in this research project is voluntary and confidential. You will not be asked to reveal any information that will allow your identity to be established by persons reading the results of the study. At this stage no follow-up interviews are planned. Attached to this letter is a document explaining your role in this research process. It includes the information provided to you during the briefing. The results from this study will be used to improve existing / extend the range of design guidelines for developing multimedia which makes extensive use of animation and images. The results will also inform designers of the extent to which the design should accommodate different cognitive styles.

If you are willing to participate in this study, please sign this letter as a declaration of your consent, i.e. that you participate in this project willingly and that you understand that you may withdraw from the research project at any time. Participation in this phase of the project does not obligate you to participate in follow up individual interviews, however, should you decide to participate in follow-up interviews your participation is still voluntary and you may withdraw at any time. Under no circumstances will the identity of interview participants be made known to the Faculty of Health Sciences or your individual lecturers. The research is done in fulfillment of the requirements for a doctoral degree in the department of Teaching and Training Studies, University of Pretoria.

Participant's signature: Date:

Student No:

Researcher's signature: Date:

Yours sincerely

Anne Strehler

Accessing the work

Open Windows Explorer

Cognitive styles analysis

- ◆ There is a folder on the C: drive – C:/CSA.
- ◆ There are two files in this folder.
- ◆ Click on the file CSA.EXE.
- ◆ Read the instructions carefully.

Before you close the program – call the research assistant to write down your scores

Opening the Lesson

- ◆ There is a folder on the C: drive – C:/Lesson1 OR C:/Lesson2 OR C:/Lesson2 OR C:/Lesson2.
- ◆ There are several files in this folder.
- ◆ Click on the file Lesson1.exe OR Lesson2.exe OR Lesson3.exe OR Lesson4.exe.
- ◆ The program will take a few seconds to open – please be patient.
- ◆ Start as soon as it is open.
- ◆ Read the instructions carefully.

Participation in the study – What the researcher expects from you

1. You have already been assigned to a particular session and the number of the lesson you must do. The way research participants are assigned is VERY important and affects the validity of the data.
2. Please do NOT swop with anyone.
3. Write your student number on each page of the handout.
4. Read all instructions carefully – everything you need to know is displayed ON THE SCREEN.
5. Do BOTH the Cognitive Styles Analysis and the lesson.
6. Please take this session seriously.
7. Answer EVERY question on every screen – including the ones embedded in between the content of the lesson.
8. Work individually – once the session starts PLEASE do NOT confer with your peers or talk to each other.
9. If you need help – ASK.
10. If the program bombs out – log in again – it is bookmarked and you can continue where you left off.
11. STUDY the content on the screen – this is a STUDY session – some of the content is not going to be repeated in your Block 3.
12. Follow ALL the links.

Results of Cognitive Style Analysis

Student Number _____

WA Ratio _____

VI Ratio _____

Scribble page

Use this page to make any notes / mind maps etc while you study the content in the multimedia

Student No:

.....

APPENDIX P: RIDING’S COGNITIVE STYLES ANALYSIS

The CSA provides a score for each dimension in the cognitive style model. The ratios for each style dimension typically range from 0.4 through to 4.0 with a central value around 1.0. But for descriptive convenience, the dimensions may be divided into groupings and given labels (Riding, 2005a).

Riding clearly states:

Since each dimension is a continuum, the labels are used only for descriptive convenience, and are not meant to imply that there are style ‘types’ in any absolute sense. There is no requirement to use the same cut-off points as those given by the CSA, as long as the cut-offs are clearly reported in the research report.

Riding (2005a, page 7).

The cut-off points suggested by Riding (2005a) for the ratios on each dimension are given below.

WHOLIST-ANALYTIC Dimension	>1.35	Analytic Verbaliser	Analytic Bimodal	Analytic Imager
	>1.02 and <=1.35	Intermediate Verbaliser	Intermediate Bimodal	Intermediate Imager
	<=1.02	Wholist Verbaliser	Wholist Bimodal	Wholist Imager
		<=0.98	>0.98 and <=1.09	>1.09
	VERBALISER-IMAGER Dimension			

There are researchers who divide the style only into two categories with the following ranges: Wholists <0.99 and Analytics >1.00; Verbalisers < 0.93 and Imagers > 1.00 (Riding, Glass & Douglas, 1993).