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.

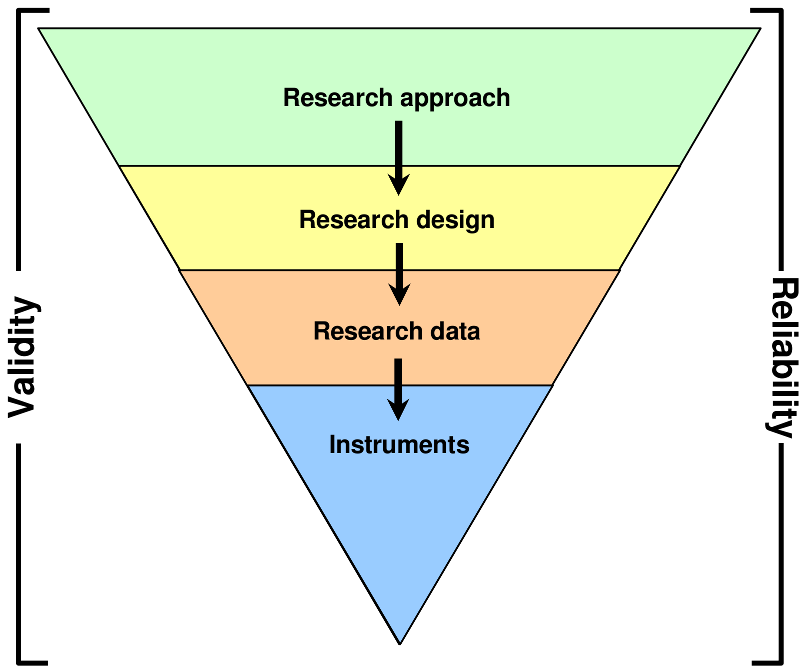


Figure 3.1: A visual representation of the research methodology

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?

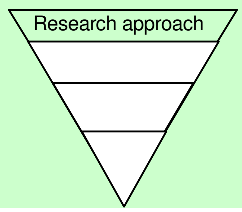
i. What were the cognitive styles of the participants who took part in the study?

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?

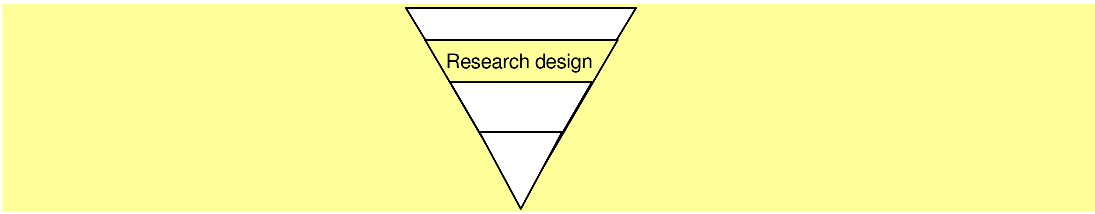


3.5 The research approach

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.

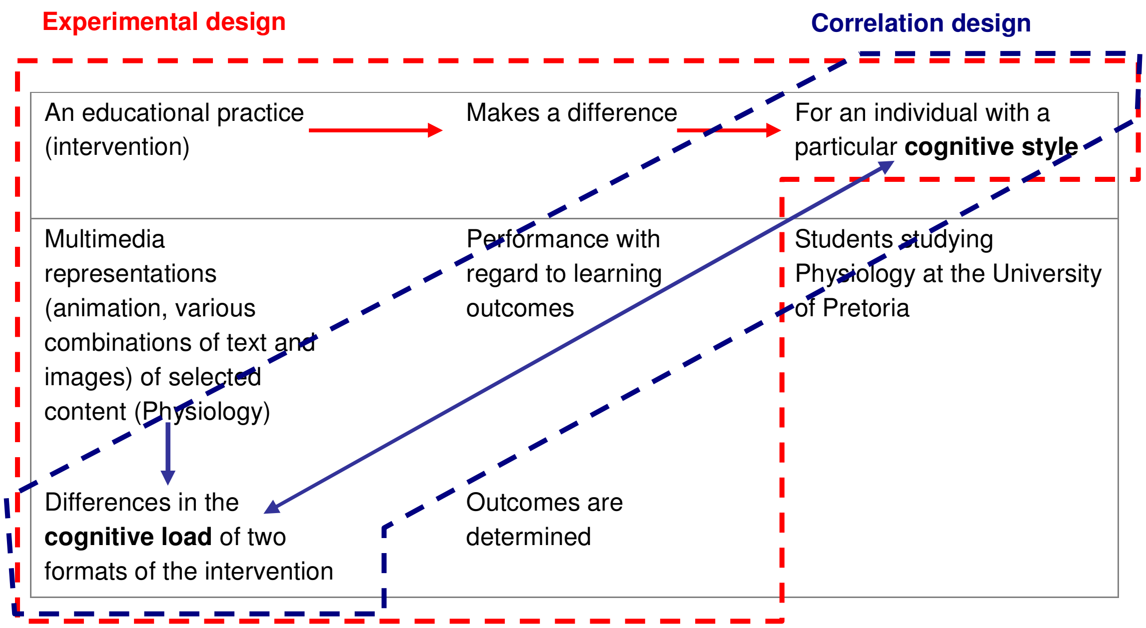


Figure 3.2: Graphic representation of the issues that define this research design

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: • 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 *a priori* – 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.

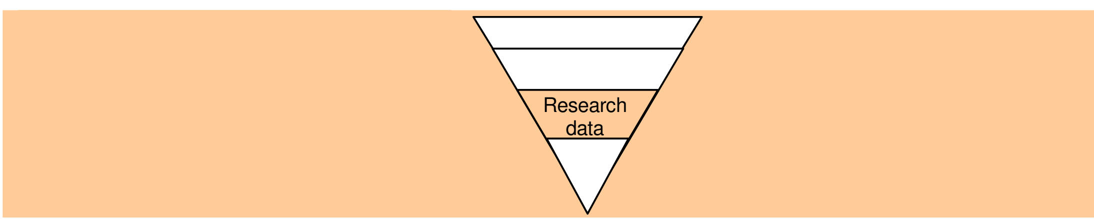
Convenience selection was thought to be appropriate for this study for the following reasons:

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:

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

Self-report of cognitive load for each version:

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

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:

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

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

Direct measure of cognitive load for each version:

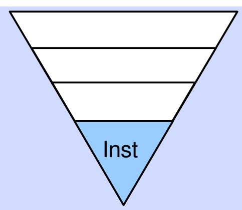
$$\text{DM1} = \text{MEAN}\ (MEANCL_1, MEANCL_2, \ldots) \text{ and } \text{DM2} = \text{MEAN}\ (MEANCL_1, MEANCL_2, \ldots)$$

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. Tere 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.

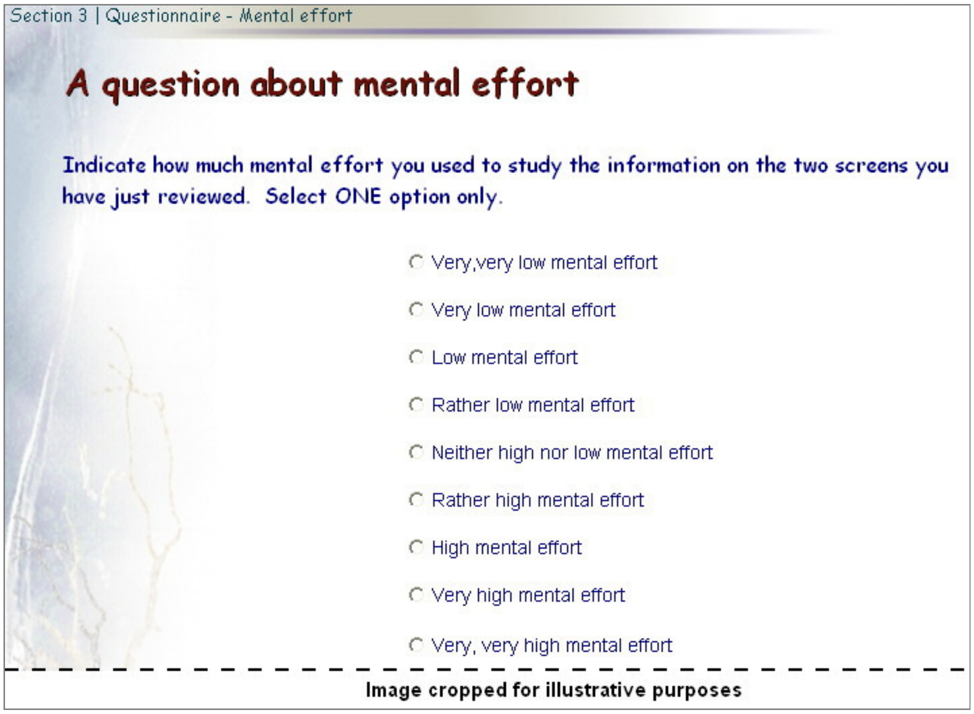


Figure 3.3: Screen format of self-reporting a rating for mental effort

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)

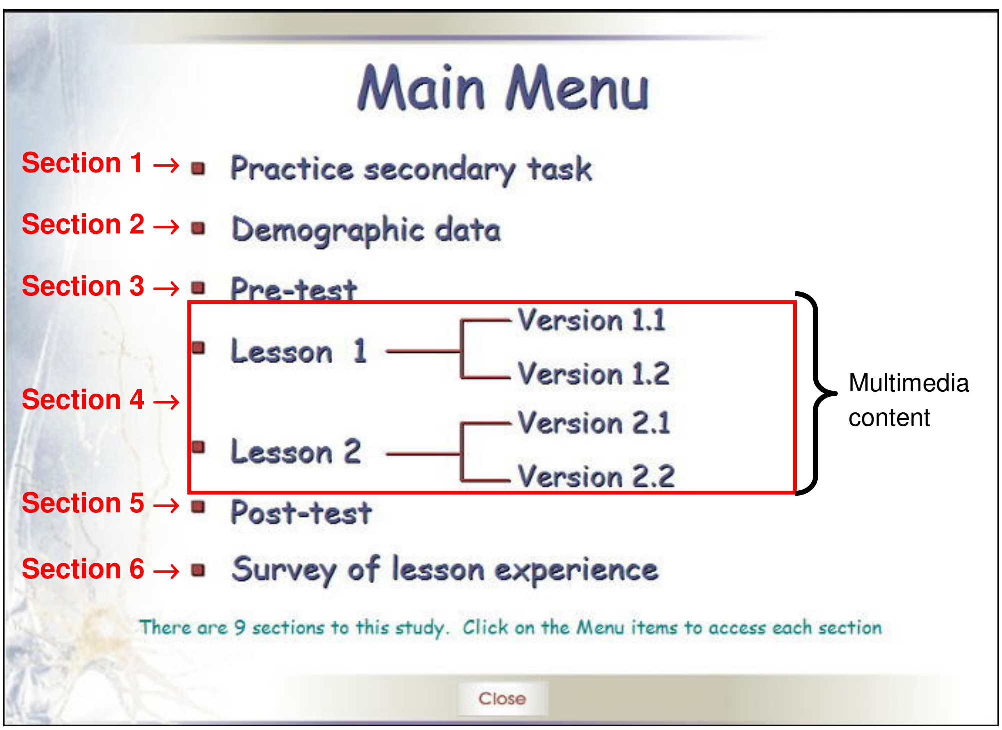


Figure 3.4: Menu for multimedia intervention used in the second pilot study

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.

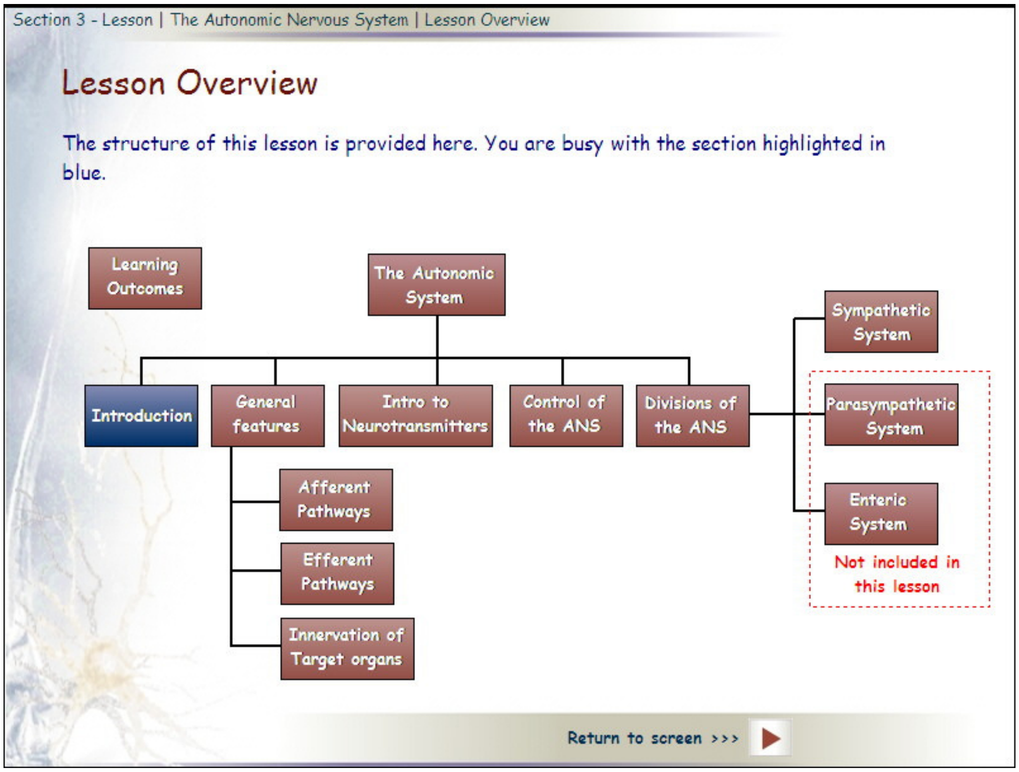


Figure 3.5: A visual representation of the structure of the multimedia program

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.

<table>
<tr><th>Section</th><th>Scope of the content</th><th colspan="2">No of screens</th><th colspan="2">Instructional strategy and media used</th></tr>
<tr><td></td><td></td><td>Vers 1</td><td>Vers 2</td><td>Version 1</td><td>Version 2</td></tr>
<tr><td>Introduction</td><td>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.</td><td>6</td><td>6</td><td colspan="2">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.</td></tr>
<tr><td rowspan="4">General features – 3 sections</td><td>Afferent pathways: Describes the structure and function of pathways carrying stimuli from the environment to the spinal cord and cortex.</td><td>3</td><td>3</td><td colspan="2">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.</td></tr>
<tr><td>Efferent pathways: Describes the structure of the pathway and its components.</td><td>1</td><td>1</td><td colspan="2">Efferent pathways :The instruction was tutorial in style. It made use of text and a static image.</td></tr>
<tr><td rowspan="2">Innervation of the target organs: Describes how this works</td><td rowspan="2">3</td><td rowspan="2">3</td><td colspan="2">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.</td></tr>
<tr><td>Version 1: The third screen summarised the information in one large table. This is illustrated in Appendix J, page 389</td><td>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</td></tr>
<tr><td>Introduction to Neuro-transmitters</td><td>Basic information about the functions of neurotransmitters. Also presented in the face-to-face presentation of the course.</td><td>1</td><td>1</td><td colspan="2">A narrated animation was used to explain this content in both versions.</td></tr>
<tr><td>Control of the ANS</td><td>Basic control mechanisms are discussed, with particular emphasis on the Sympathetic Nervous System (SNS).</td><td>1</td><td>1</td><td colspan="2">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.</td></tr>
</table>

Section	Scope of the content	No of screens		Instructional strategy and media used	
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.	**Vers 1**	**Vers 2**	**Version 1.**	**Version 2**
		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 3rd, 4th 5th and 6th 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 the 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.

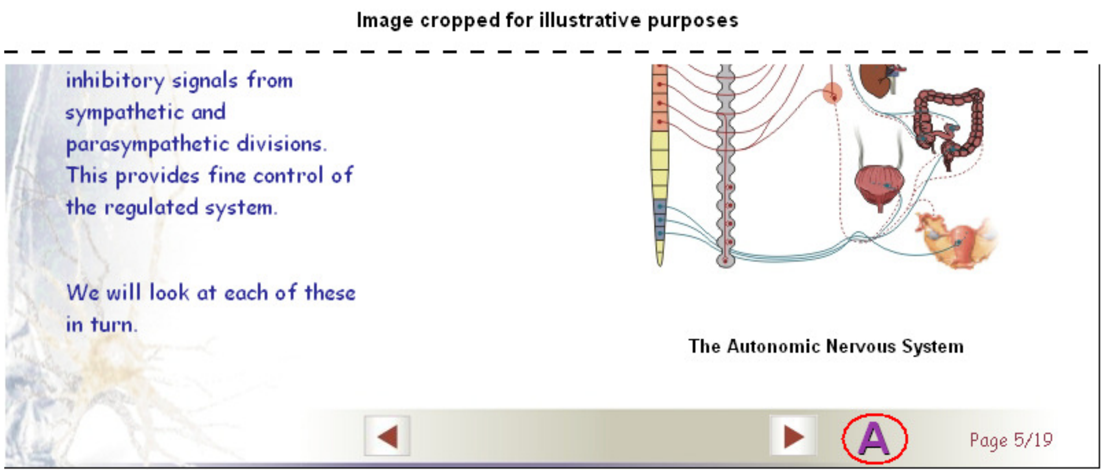


Figure 3.6: Display of secondary task in the program used in pilot study 1

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.

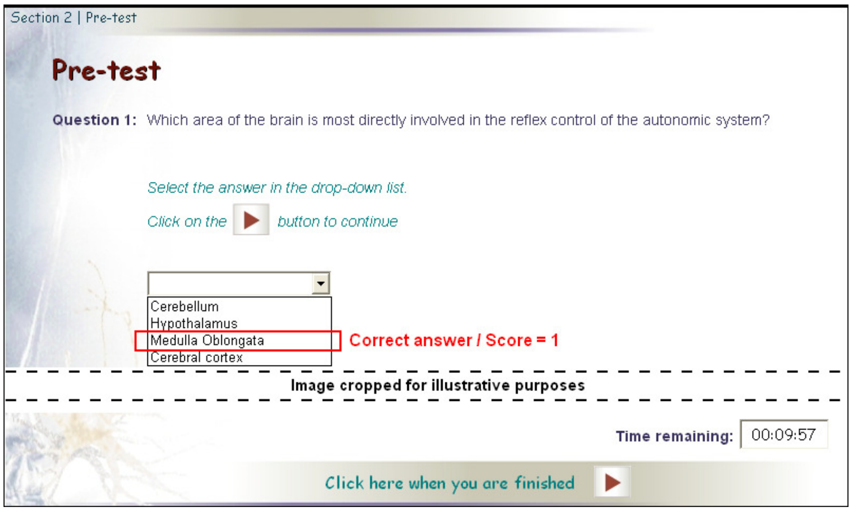


Figure 3.7: How remaining time was displayed on screen in the pretest

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

Image cropped for illustrative purposes

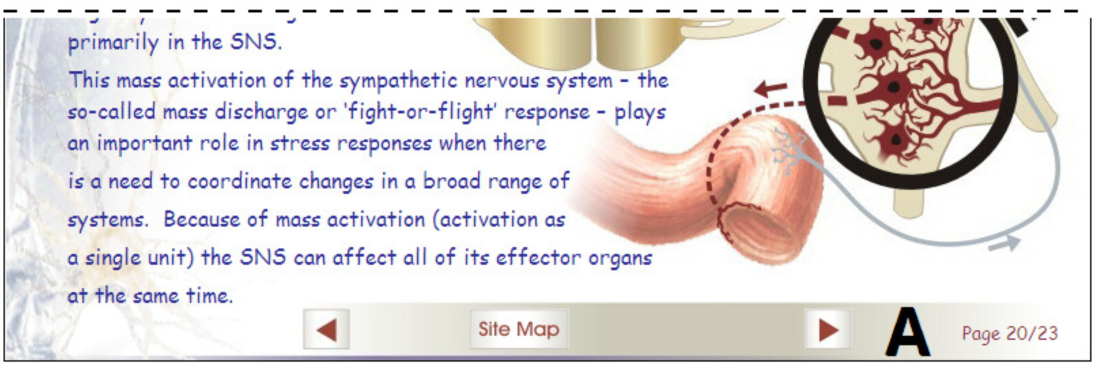


Figure 3.9: Placement of the secondary task in Programs 1 and 3

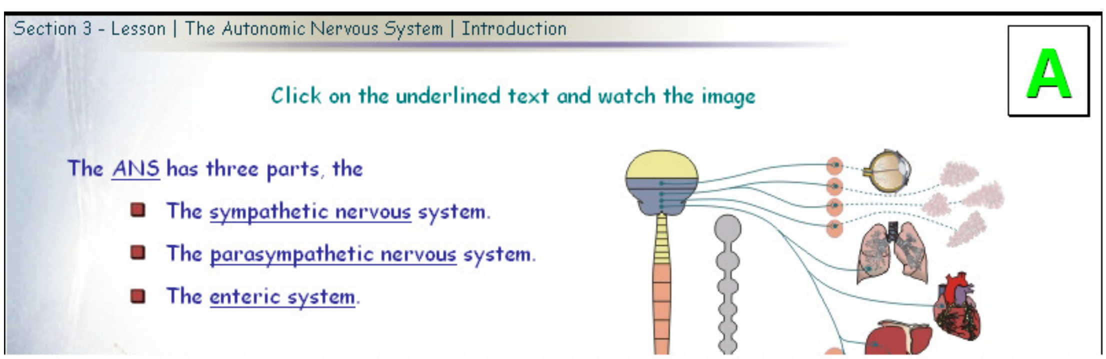


Image cropped for illustrative purposes

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 as 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.

----- o O o -----