UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Factors in the measurement of cognitive load of multimedia learning

by

M E Smith

submitted in partial fulfilment of the requirements of the degree

Magister Educationis

in

Computer-Integrated Education

in the

Department of Curriculum Studies

Faculty of Education

University of Pretoria

Supervisor: Mrs A Strehler

Co-supervisor: Prof. Dr J.C. Cronjé

2007

Abstract

Candidate: M E Smith

Supervisors: Mrs. A Strehler, Prof. Dr J.C. Cronjé

Department: Curriculum Studies

Faculty: Education

Degree: MA (Computer Integrated Education)

Title: Factors in the measurement of cognitive load of multimedia learning

In this study, the author investigated factors that influence the direct measurement of cognitive load using the dual-task method. The dual-task method is an unambiguous and objective technique to measure cognitive load. The primary task was to master content in a lesson about the Autonomic Nervous System. The secondary task was to respond to a symbol that changed colour by pressing the Enter key. The time between the symbol changing colour and the response of the student was measured. Two versions of the multimedia program tested the influence of the presentation format and instructional strategy on cognitive load. Each version of the program was further subdivided into four lessons, which were used to test the influence of the position of the secondary task on the cognitive load. All the data was collected electronically. The statistical analysis revealed that the position of the secondary task does not influence cognitive load ($F(1, 2661) = 3.25$, $p = 0.071$). The presentation format and instructional strategy used in this study however did result in a significant difference between the cognitive load of the two versions. The mean cognitive load of the version using animation was 6.408 and that of the version using predominantly static images and text was 5.684. This difference was found to be highly significant ($F(1, 2661) = 52.39$, $p < .0001$). It was concluded that using animation to present content required more mental effort by participants than using images and text to present the same content.

Acknowledgements

My Skepper	Aan U die eer en dank
My supervisor, Anne Strehler	Not only have you been my supervisor but also became my mentor and friend. Thank you for your time and knowledge that you shared with me.
My promoter, Prof Johannes Cronje	Thank you for being inspiring and kind.
Dr Mike van der Linde, Department of Statistics, University of Pretoria.	Thank you for tremendous patience, guidelines and excellent service.
Colleagues – Epi-Use Learning	Your support and encouragement are much appreciated.
Familie en vriende	Sonder julle begrip en ondersteuning sou dit soveel moeiliker gewees het.
Janna en seuns	Dankie julle drie vir liefde, begrip, geduld en opofferings.
Francois en Brenda	Dankie vir al die liefde, aanmoediging, bemoediging en inspirasie.
My meisiekind, Este	Dankie vir kritiek, insette, proeflees, kritiese lees, taalversorging, begrip en geen simpatie! Ek weet met hoeveel liefde dit gedoen is.

Contents

List of tables

List of figures

Chapter 1: Background

1. Introduction

In this dissertation a study is described that used quantitative means to investigate selected factors that influence the measurement of cognitive load. The study measured the cognitive load of a multimedia program using the dual task method.

In this chapter I

- discuss the purpose and rationale of the study,
- state the research question,
- clarify concepts used in the study,
- briefly describe the methodology, including the limitations and assumptions of the study, and
- provide an outline of this dissertation.

2. Purpose

Although cognitive load is well described as a theoretical construct (Brünken, Plass & Leutner, 2003), the literature describes several techniques for measuring cognitive load, particularly in multimedia instruction (Brünken et al., 2003; Paas, Tuovinen, Tabbers & Van Gerven, 2003). A review of the literature seems to indicate that the direct measurement technique is seldom used. However, the literature also suggests that this technique might be more accurate than the subjective measures of cognitive load (Mayer & Moreno, 2003).

The purpose of this study is to determine whether selected factors influence the measurement of cognitive load of multimedia learning material, using a direct measurement technique.

This research forms part of a wider study where the researcher investigated the relationship between multimedia learning, cognitive load and cognitive style. In the wider study, the researcher gave the students two different formats of the same content.

The aim of my study was to measure the cognitive load of the different learning formats.

I was initially only going to determine the cognitive load of two multimedia programs, using the dual task approach. In piloting the study, I discovered that the colour of the symbol used in the secondary task seemed to influence the visibility and therefore also the reaction time to the secondary task. The question "Are there other factors that also influence the direct measurement using the dual task approach?" became relevant.

These factors include the presentation format of the learning content as well as the design, position and frequency of the secondary task. I then decided to explore the influence of the position of the secondary task on the measurement of cognitive load.

3. Rationale for the study

Even though research on cognitive load has been done for the past thirty years (Antonenko, 2005) and there are well-defined principles for instructional design with the goal of reducing cognitive load, the measurement of cognitive load in multimedia learning in a direct manner is limited. To date the measurement of cognitive load has primarily been by way of subjective ratings (Sweller & Van Merrienboer, 2005; Paas & van Merrienboer, 1994; Brünken, Plass & Leutner, 2003).

Direct measurement might be a more accurate measurement because external factors, such as opinion, are eliminated to a certain extent. In the direct measurement method the environment in which cognitive load is evaluated can be controlled and manipulated. The accuracy of these measurements could therefore be improved since the measurement instruments measure the cognitive load in terms of reaction time to a secondary task instead of the subjective opinions of learners obtained through questionnaires (Brünken, Plass & Leutner, 2004).

In the wider study, the cognitive load was also measured, using subjective techniques. The results of my study will be used to investigate the correlation between the subjective rating and the direct measurement of cognitive load.

4. Research questions

The research questions for this study are:

1. What is the cognitive load of the different presentation formats of a multimedia program, using the direct measurement technique?
2. How does the position of a secondary task influence the measurement of cognitive load?

5. Concept clarification

In this specific field of research, it might be necessary to clarify some of the terminology upfront. I will define other terminology per chapter as it is used. A few basic concepts are described in this section and are presented in Table 1.1.

Concept	Explanation
Multimedia	Information presented in a variety of formats that can include textual, audio and video.
Multimedia learning	Learning from text and graphics, animations or audiovisual programs.
Multimedia instruction	The presentation of learning content using text and graphics, animations or audiovisual programs.
Cognitive load	The level of effort associated with thinking and reasoning (including perception, memory, language, etc.), when potentially interferies with other cognitive processes.
Cognitive load theory	The assumption that optimum learning occurs in humans when the load on working memory is kept to a minimum to best facilitate the changes in long term memory (Sweller, 1988).

Concept	Explanation
Subjective measurement of cognitive load	This technique is based on the method where a rating scale is used and respondents are expected to introspect on their cognitive process and rate the amount of mental effort used to understand specific instructions.
Direct measurement of cognitive load	This technique measures the cognitive load without expecting the respondents to understand or evaluate the cognitive process. Direct reactions or physical changes are monitored. Psycho-physiological methods can be used such as the measurement of the cardiovascular state or the pupil diameter that changes according to the mental effort in the cognitive process. Another method in the direct measurement of cognitive load is the dual task method (Schultheis & Jameson, 2004).
Primary task	The primary task is to achieve the outcome of the learning program. In order to achieve these learning outcomes the learner must acquire specific knowledge and/or master certain skills during the learning instruction. It is possible to assess the outcome of the primary task.
Secondary task	This additional task is added to the learning intervention and is unrelated to the learning outcome. It cannot be assessed in itself. However, certain actions can be measured, similar to the reaction time on a symbol that is displayed, or the physiological changes of the learner like the pupil size.

Table 1.1: Concept clarification

6. A broad overview of the research design

This study used a quantitative approach, with an experimental design. The study comprised of a pilot study, which was done in order to test the design of the study and the data collection process, and a main study.

The unit of analysis was students from the Faculty of Health Sciences of the University of Pretoria. A convenience sampling method was used to identify the respondents for both the pilot and the main study. Respondents for the pilot study were 38 second year physiotherapy students. Respondents for the main study were 240 second year dental and medical students.

Four different multimedia programs were developed to measure the cognitive load and test the influence of the position of the secondary task on the measurement of cognitive load. A computer program was used to randomly assign the respondents to one of these four programs.

The differences in the programs used in this experiment are indicated in Table 1.2.

Program	Presentation format	Position of the secondary task
1.1	Graphics, text, audio and animations	Bottom right
1.2	Graphics, text, audio and animations	Top right
2.1	Text and static images only	Bottom right
2.2	Text and static images only	Top right

Table 1.2: Differences in the programs

The data collection instrument was built into the programs. The programming was done in such a way that the following data was recorded in an external log file:

- time the lesson was accessed,
- time the lesson was exited,
- time a screen was entered,
- time the learner left a screen,
- the trigger time of the secondary task and
- the response time of the learner to the secondary task.

The data was written out to an .ini file, together with demographic data, style analysis results, pre- and posttest results and the subjective measurement of the cognitive load according to a nine-point rating scale. All of the latter were for the wider study.

Chapter 3 describes the programs, data collecting and direct measurement technique in more detail.

7. Limitations to the study

The size of the different computer laboratories necessitated that the study run across several sessions as no laboratory could accommodate more than 70 students in one session. It would have been a limitation to the study if the students discussed the study amongst each other before every one has completed it. This might have contaminated the data. A solution was to run the sessions simultaneously in different laboratories.

Other limitations to the study are the large group tested at once and that no qualitative data was recorded for the experiences.

8. Structure of the dissertation

This dissertation is structured as follows:

Chapter 1:	Background	Introduction Purpose Rationale for the study Research questions Concept clarification A broad overview of the research design Limitations to the study Structure of the dissertation Summary
Chapter 2:	Literature review	Introduction Clarification of concepts Existing theories and concepts Existing research
Chapter 3:	Methodology	Introduction Clarifying the concepts Planning the research Conducting the study Summary
Chapter 4:	Data analysis and findings	Introduction Clarifying the concepts Statistical analysis Results of the experiment Methodological reflections Summary
Chapter 5:	Closure	Introduction Synopsis of the study Discussion Conclusion Recommendations for further studies

9. Summary

This was a quantitative study to determine the cognitive load of a learning interventions that used different presentation formats. The research questions were to determine what factors influence the measurement of the cognitive load while using the dual task method. The learning intervention consisted of an interactive computer program. Data was collected electronically in the format of .ini files.

Chapter 2 provides a critical review of the relevant research literature.

Chapter 2: Literature review

1. Introduction

This review summarises existing research and emergent practice in direct measurement of cognitive load using the dual task approach.

The review considers the research fields of:

- multimedia learning,
- cognitive load (CL),
- cognitive load theory (CLT),
- measurement of cognitive load
  - direct and
  - self-report measures.

2. Clarification of concepts

In this area of research there are concepts that the reader should understand as defined and explained in Table 2.1.

Concept	Explanation
Sensory memory	This memory deals with stimuli from our senses like sight, smell, taste and touch and is extinguished quickly (Cooper, 1998).
Working memory	This refers to a limited part of the memory that is used to direct attention or process information (Cooper, 1998).
Long term memory	It is a body of knowledge and skills stored in a relatively permanent accessible form in an unlimited space (Cooper, 1998).
Learning	Learning can be defined as the structure and storage of knowledge and skills in the long term memory in such a way that it can be re-called and applied later on demand (Cooper, 1998).

Concept	Explanation
Schemas	A schema is defined as a mental construct permitting problem solvers to categorise problems according to solution modes (Sweller, Chandler, Tierney & Cooper, 1990). It refers to the mental encoding of experiences or the cognitive structure utilized to make sense of the world.
Experts	Experts are people who have built up an extensive set of schemas with regard to a particular area and can carry out certain actions with a high degree of automation. They are familiar with solutions and can apply these without much mental effort (Cooper, 1998).
Novices	Novices have relatively few schemas and do not recognise a solution, or if they do, they have difficulty in applying it (Cooper, 1998).
Mental model	Mental models are thought of as organised knowledge structures that humans process in order to describe events in their environment, make sense of them and predict future events (Neumann, Badke-Schaub & Lauche, 2004).
Mental load	Mental load is the aspect of cognitive load that originates from the interaction between task and subject characteristics. Mental load can be determined on the basis of our current knowledge about task and subject characteristics. As such, it provides an indication of the expected cognitive capacity demands and can be considered an estimate of the cognitive load. This is imposed by the task of environmental demands (Paas & Van Merrienboer, 1994).
Mental effort	Mental effort refers to the amount of non-automatic mental elaborations necessary to solve a problem. It can also refer to resources that are actually allocated to accommodate the task demands (Paas & van Merrienboer, 1994).
Intrinsic cognitive load	This is defined by the difficulty level of the content and cannot be influenced by the instructional design. It is the load on memory required by the thinking task at hand (Chipperfield, 2004).
Extraneous cognitive load	This load is due to the instructional design used to present the content (Kirschner, 2002).
Germane cognitive load	This load refers to the cognitive capacity needed for active knowledge construction or schema integration (Brünken, Plass & Leutner, 2004).

Table 2. 1: Concept clarification

3. Existing theories and concepts

3.1 Multimedia

In this section the most relevant theories of multimedia learning for this study will be discussed.

3.1.1 Defining multimedia

The idea of using multimedia for education is not new. Originally the term multimedia was used by institutions that provided distance learning courses (Oliver, 1998). These institutions delivered content via a combination of text, TV, telephone, audio cassette and the radio. Later developments included using computers to bring the media together. The advent of high resolution screens and sound and video playback facilities for computers has resulted in the increased use of multimedia applications for education.

With time and progression in technology the definition of multimedia has been extended. Richard Mayer (2001) defines multimedia as modern presentation modes (e.g. text and pictures) and different modalities (e.g. visual, auditory) that are presented by an integrated technical system e.g. computer and internet.

Mayer and Moreno (2003) define multimedia learning as learning from words and pictures and multimedia instruction as **presenting** words and pictures.

On the website of The Learning Technology Development Unit (LTDU), Andrew Oliver (1998) describes the elements of multimedia as follows:

1. **Text**: This is the basis of most applications. It is the on-screen display of words. The use of different styles, fonts and colours can emphasise specific issues in the instruction.
2. **Images**: Seeing a picture of an object has more impact than merely reading about it. Examples include conventional artwork, computer-generated artwork, photographs or captured video frames.

3. **Movies**: You can present information which is normally outside the scope of the ordinary classroom, such as medical operations or archaeological excavations.
4. **Animation**: Animations can render a procedure more accurately than a movie. For instance objects that appear blurred in a movie can be represented more clearly.
5. **Sound**: Sound can be used in strategic parts of the program or during a movie to emphasise certain points. This may include speech, audio effects (e.g. applause), ambient sound (e.g. the background sound of the sea etc.) and music.
6. **User control**: There has to be some degree of user control so as to provide learners with the option to leave certain parts of the application and thus prevent boredom. On-screen options should exist for them to visit other areas of the program.

The above media should ideally be combined to produce a seamless application. Typical examples of such integration include

- linking an animation to static in-text diagrams,
- linking a video clip to a descriptive paragraph and
- audio recordings of foreign language pronunciation together with the words and/or pictures.

3.1.2 Theories regarding multimedia

Multimedia provides the opportunity to learners to associate information they get from different presentations and different sensory modalities into a significant experience (Moreno & Mayer, 2000).

Mayer has based the majority of his multimedia work on an integration of Sweller's cognitive load theory (Chandler & Sweller, 1991; Sweller, 1999), Pavio's dual-coding theory (Clark & Paivio, 1991; Paivio, 1986), and Baddeley's working memory model (Baddeley, 1992; Baddeley& Logie, 1999). Mayer focuses on the auditory/verbal channel and visual/pictorial channel, and then defines multimedia as the presentation of material using both words and pictures. He motivates the use of this narrow definition of just two forms, verbal and pictorial, because the research base in cognitive psychology is most relevant to this distinction.

Mayer conceptualises this theory of multimedia visually as seen in Figure 2.1.

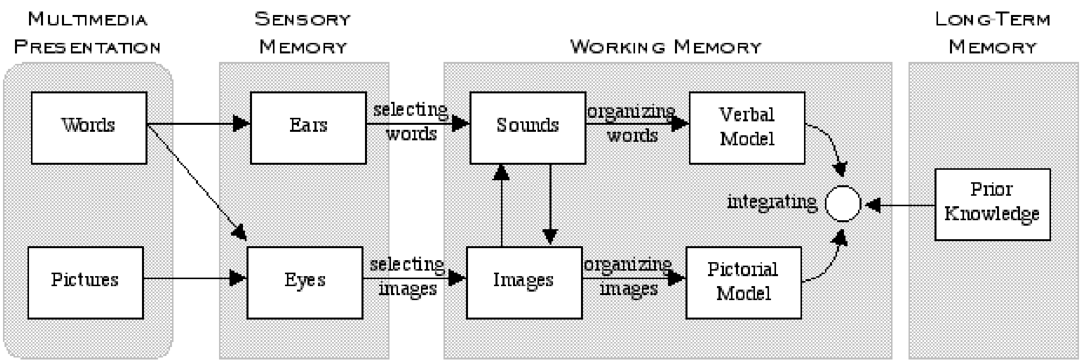


Figure 2. 1: Multimedia Model (Mayer, 2001)

This model is based on the following assumptions (Mayer, 2001):

- There are distinct information processing channels in the memory that processes visual and auditory experiences/information.
- Each information processing channel is limited in its ability to process experience/information.
- Processing experience/information in channels is an active cognitive process designed to construct coherent mental representations.

An explanation of how the model works include five steps. First the learner selects relevant words for processing in the verbal working memory and then the learner selects relevant images for processing in the visual working memory. Thereafter the learner organises selected words into a verbal mental model and selected images into a visual mental model. Lastly the learner integrates verbal and visual representations as well as prior knowledge into a schema that is stored in the long term memory (Mayer, 2001).

To summarise the work done by Mayer over a period of 15 years the following design practices for multimedia emerged:

- Learners learn better from words and pictures than from words alone.

- When corresponding words and pictures are presented near rather than far from each other and simultaneously rather than successively on the page or screen, learning improves.
- The progress in learning tends to be better when extraneous words, pictures, and sounds are excluded rather than included in the multimedia program.
- Learners perform better when they have learnt from animation and narration than when they have learnt from animation, narration and on-screen text.

In this study, I will look at multimedia formats that include text and static graphics as one format and text, animation and audio as the other format. I will measure the cognitive load of the different formats.

3.2 Defining cognitive load

Sweller (2003) describe cognitive load as the level of “mental energy” required to process a given amount of information. It refers to the total amount of mental activity imposed on the working memory at an instance in time. Working memory is the stage of memory where information is stored for a short period prior to either being forgotten or transferred to long term memory. Long term memory refers to the relatively permanent memory.

We experience cognitive load because of the limitations of the working memory.

In this section I describe the types of cognitive load and the different memory influenced by cognitive load.

Cognitive load refers to a multidimensional construct representing the load that performing a particular task imposes on the learner’s cognitive system (Paas & van Merrienboer, 1994). Paas and van Merrienboer (1994) present a model wherein the construct reflects the interaction between task and learner characteristics and another element reflecting the measurable concepts of mental load, mental effort, and performance. As such, the amount of cognitive load, measured at a given time, is a way of assessing the level of information being manipulated in working memory (Paas, Tuovinen, Tabbers & Van Gerven, 2003).

Cognitive load can be intrinsic or extrinsic. Extrinsic cognitive load is further divided into extraneous cognitive load and germane cognitive load. Sweller (1988) differentiates between intrinsic, germane, and extraneous cognitive load.

3.2.1 Intrinsic cognitive

Intrinsic cognitive load refers to the inherent difficulty of the content. Intrinsic cognitive load refers to the number of elements that are integrated into the content schema and therefore have to be processed simultaneously. The number of elements that need to be attended to simultaneously determines intrinsic cognitive load. The higher the number of elements that interact with each other, the higher the intrinsic load is said to be. It depends on the complexity of the content to be learned and the learner's degree of prior knowledge.

3.2.2 Extraneous cognitive load

Extraneous cognitive load refers to the load imposed by the instructional design (an external factor). Inefficient instructional designs adds unnecessary load. For example, an audio-visual presentation format usually has lower extraneous load than a visual plus text format, because in the former case, working memory has less information to process. In the first case only one channel in the memory is used to process the information. In the second scenario two channels, one for the text and one for the visual presentation are used (Mayer, 2001).

3.2.3 Germane cognitive load

Germane cognitive load refers to the degree of effort involved in processing, internal organisation, integration and schema construction of information. Germane load is sometimes associated with motivation and interest.

According to Sweller (1994), intrinsic load is unchangeable, whereas the instructional designer can manipulate extraneous and germane load.

3.2.4 Working memory

Working memory is used to process information and create schemas in the long term memory. The working memory has a limited capacity and is affected by all the different types of cognitive load. The aim of instructional design is to keep the load as low as possible so that there is free working memory that can be used for creating information that will be transferred to the long term memory by way of schemas.

This is illustrated in the Figure 2.2

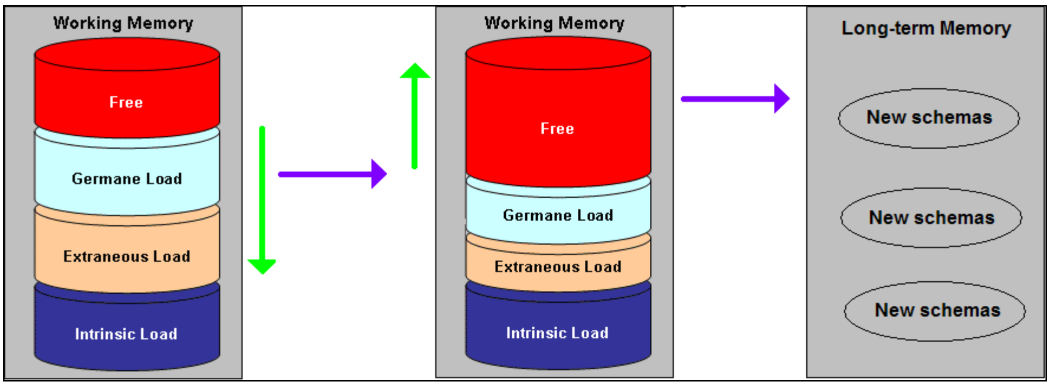


Figure 2. 2: Capacity of the working memory

3.2.5 Schemas

Schemas are defined by Sweller and Cooper (1985, p60) as

"...mental constructs that allow patterns of configuration to be recognised as belonging to a previously learned category and which specify what moves are appropriate for that category."

Sweller found that learning will only occur if a connection is made to a schematic structure in the long-term memory. Otherwise, the learner will forget the material (Sweller & Cooper, 1985).

3.2.6 Summary

Instructional design directly affects extrinsic cognitive load. This load can be categorised according to whether it will be beneficial for schema construction (germane load) or not (extraneous cognitive load). If the instructional design of the content is good (low extraneous cognitive load), cognitive resources are kept available and can be used to enhance learning through schema construction (germane cognitive load). On the other hand, implementing instructional design techniques that requires the learner to get involved in activities that are not directed to construct schemas in the long term memory will increase extraneous cognitive load (Sweller, 1994).

Excellent instructional design will optimise cognitive load by decreasing extraneous cognitive load and increasing germane cognitive load.

3.3 Cognitive load theory

This theory originated out of work done by Sweller. He defines cognitive load theory (CLT) as both a theory of cognition and learning and an instructional design model (Sweller, 1994).

Cognitive load theory describes how a person's cognitive architecture has implications for learning and therefore for instructional design. The theory has influenced instructional design and provides a conceptual framework for instructional designers to limit and maybe control the circumstances that create unnecessary cognitive load in learning materials.

In 1956 Miller conducted research and came to the conclusion that the working memory can only process seven pieces of information at the same time (Miller, 1956). Sweller built his research on this empirical evidence and eventually defined the cognitive load theory (CLT) as an information processing model of cognition. He emphasised the inherent limitations of the working memory and used schemas as the relevant building blocks for instructional materials.

Cooper, one of Swellers co-workers, presents CLT as follows (Cooper, 1998, p1):

> *It describes learning structures in terms of an information processing system involving long term memory, which effectively stores all of our knowledge and skills on a more-or-less permanent basis and working memory, which performs the intellectual tasks associated with consciousness. Information may only be stored in long term memory after first being attended to, and processed by, working memory.*

Cooper of the School of Education Studies, the University of New South Wales, Sydney, NSW 2052, Australia, did research into cognitive load theory and instructional design in 1998. He came to the conclusion that the working memory is extremely limited in both capacity and duration and, under certain circumstances, will hold learning back. The basic belief of cognitive load theory is that the limitations of the working memory should be taken into consideration when the form and quality of the instructional design is raised (Cooper, 1998). For a more detailed discussion on this article refer to paragraph 4.2.

3.3.1 Experts and Novices

According to the learning process, when something has been learned, it implies that it has been successfully encoded into long term memory and can later be recalled on demand (Cooper, 1998).

Cooper is of opinion that the only two distinguishing features of expertise are:

- the expansive schemas (information networks) that experts hold, and
- the high level of automation (ability to perform tasks without concentrating) that experts exhibit.

This investigation into the differences between experts and novices initiated the development of the cognitive load theory (Cooper, 1998). Experts have power that is derived through advanced knowledge or experience in a particular subject. Novices, on the other hand, are new to the subject.

Schemas and automation appear to explain all other expert/ novice differences (Cooper, 1998). This is supported by Gick and Holyoak who say that the essential strength of expert performance lies in schemas (Gick & Holyoak, 1983).

Schemas enable experts to categorise problems and consequently solve them, but novices do not possess schemas and are therefore unable to categorise or solve problems. Novices have no alternative other than to engage in general search techniques such as trial-and-error, or means-ends analysis (Chi, Glaser & Rees, 1982; Larkin, McDermott, Simon & Simon, 1980).

3.3.2 Cognitive load and instructional design

Sweller (1994) is of the opinion that most often learning is unsuccessful because the instructions are unclear, difficult to understand or misleading. The learner's attention is focused on the instruction and not on the learning.

It is usually assumed in the cognitive load theory that intrinsic load cannot be changed by instructional design. However, this assumption is controversial (Van Merrienboer, Kirschner & Kester, 2003). Different instructional design techniques have been proposed by Van Merrienboer, Kirschner and Kester that tend to reduce intrinsic cognitive load associated with learning materials. Techniques like scaffolding, simple-to-complex sequencing and worked-out examples can structure the cognition process whilst presentation format might influence the cognitive load.

According to Moreno and Mayer (2002) current studies are based on the assumptions that multiple representations help learning and cognitive load hurts learning. They furthermore state that multiple representations are easier to integrate with prior knowledge and less likely to overload the working memory. Nevertheless it is my opinion that the validity of the statement that multimedia learning strengthens learning while cognitive load burdens learning can only be tested if cognitive load is measured.

3.4 Measurement of cognitive load

With regard to the measurement of cognitive load there exist basically two different techniques- those that measure the load subjectively and the direct measurement of cognitive load. Subjective measurements expect the learner to analyse the cognitive process and evaluate the cognitive load, whilst during direct measurement the learner unconsciously reacts to the cognitive load.

Brünken, Plass and Leutner (2003) indicate that there are two major means to measure cognitive load namely objectivity and casual relations. Objectivity refers to whether the data collecting method is objective or subjective. Direct or indirect measurement refers to the casual relation between the method and the participants. These methods of measuring cognitive load are summarised in Table 2.2 (Brünken, Plass & Leutner, 2003).

	Indirect	Direct
Subjective	• Perception of invested mental effort. • Post treatment questionnaires to report the amount of mental effort (not related to cognitive load).	• Rating of difficulty of material (relates directly to the cognitive load imposed).
Objective	• Analyse performance outcomes measures. • Analysis of behavioural patterns. • Physiological measures such as heart rate and pupil dilation.	• Neuro-imaging techniques that measures brain activity (not inclusive of the complete cognitive process). • Dual task paradigm - Secondary task is added to induce the memory load. Performance in primary task is measured. - Use secondary task to measure memory load. Performance in secondary task is measured.

Table 2. 2: Classification of methods for measuring cognitive load

This study used the direct measurement technique as described by Brünken, Plass and Leutner (2003) namely objective measurement using dual task performance.

There are several benefits in using this method. Firstly is that the multimedia learning (primary task) and the secondary task are attended to at the same time. This means that the load is measured at the exact time and point that it is presented to the learner, whereas other subjective measures can only be applied after the learning event. Secondly, research on working memory has shown that there are different secondary tasks that are linked to different process steps of information processing, such as perception, pre-processing in one of the slave systems, or information integration (Baddeley, 1986) as cited by Brünken, Plass and Leutner (2003). These different tasks make it possible to identify in which of the process steps the cognitive load is imposed.

Another benefit of the design used by Brünken, Plass and Leutner (2003) is that the dual task analysis is carried out within specific subject designs. Therefore the measurement of load is induced by different design variants of multimedia instruction for the same learner and makes the load measurement independent from individual differences, such as abilities, interest, or prior knowledge, that are known to affect learning outcomes in-between subjects designs.

The weaknesses of the dual task method lie in the methodological and technical challenges that have to be taken into consideration (Lansman & Hunt,1982). Paas et al. (1994, as cited by Brünken, Plass & Leutner, 2003) elaborate on the methodological weaknesses and technical challenges as identified by Lansman and Hunt (1982). They claim that:

- the secondary task has to require the same cognitive resources as the primary task; otherwise, secondary task performance will be independent of primary task performance.
- the performance measure for the secondary task has to be reliable and valid.
- the secondary task has to be so simple that it does not suppress simultaneous learning processes.
- the secondary task has to be able to consume flexibly all of the available free cognitive capacity.

Furthermore they found that:

- to measure reaction time fulfils these requirements.
- this method has been used in studies before (Verwey & Veltman, 1996 and Wickens,1984).

In this study I use the same principle for load measurement where the learner has to react to a specific signal as soon as possible during simple continuous task monitoring.

4. Existing research

The theories and concepts discussed above have been used as the theoretical framework in a number of research studies. The method, conclusions and limitations of some of the relevant studies on multimedia, cognitive load theory and the measurement of cognitive load are further described below.

4.1 Multimedia

4.1.1 How do we learn from multimedia

Van Gerven, Paas, Van Merrienboer, Hendriks and Schmidt (2003) did a study on the efficiency of multimedia learning into old age.

The main question posed was whether or not multimedia-based worked examples add to the efficiency of skill training for elderly learners. The researchers had two hypotheses. The first one stated that worked examples presented in a multimedia format are more efficient than both conventional problems and uni-modally presented worked examples, in that at least an equal level of performance can be obtained with less effort. The second hypothesis followed a framework set by Van Gerven et al. (2000) and indicated that elderly learners will profit more from multimedia-based worked examples. The purpose of the study was to investigate the impact of age and instructional format on training effectiveness. They expected to find that the use of worked examples presented using animation would be a more effective way of learning.

In this experimental study 60 secondary school students with mean age = 15.98 years, and sixty elderly persons with mean age = 64.48 years participated. Participants of both age groups were trained in either a conventional, a uni-modal, or a multimedia condition. Subsequently, they had to solve a series of test problems.

A nine-point symmetrical category scale was used as a subjective cognitive load (SCL) measure (Paas, Van Merrienboer, & Adam, 1994).

Participants studying the worked examples of both formats used considerably less mental effort than those who solved the conventional problems. In addition the research showed the beneficial effect that multimedia learning had on the elderly.

4.1.2 The influence of instructional design on cognitive load

Mayer and Moreno (2003), researched ways of reducing cognitive load in multimedia learning.

The purpose of the research was to examine the conditions under which the use of words and pictures promote meaningful learning.

The researchers based their approach on the idea that the best way to understand how people learn is to test theory-based predictions in the context of student learning scenarios. They define multimedia as learning from text and images, using two modalities - auditory and visual.

They distinguish between three kinds of cognitive demands made on the learner when they study:

- essential processing is the cognitive process of making sense of the material,
- incidental processing is the cognitive process that plays a role in processing information which is not essential to the learning task but which has been built into the design of the learning material and
- representational holding which is the cognitive process which requires that the learner hold some information in working memory while other essential information, often related, is being processed.

The researchers described different scenarios for cognitive overload, with a comprehensive exposition of the problems and possible solutions summarised in Table 2.3.

Over load scenario	Solution
Off-Loading: when one channel is overloaded with essential processing demands.	One solution to this problem of off-loading is to present words as narration. In this way, the words are processed, at least initially, in the verbal channel.
Segmenting and Pre-training: when both channels are overloaded with essential processing demands in the working memory.	A solution to segmenting is to allow some time between successive segments and pre-training, learners receives prior instruction.
Weeding and signalling: when the system is overloaded by incidental processing demands due to extraneous material.	A solution to weeding and signalling is to eliminate interesting but extraneous material, provide cues to learner on how to select and organize material.
Aligning and eliminating redundancy: when the system is overloaded by incidental processing demands attributable to how the essential material is presented.	Aligning words and pictures would be a solution to eliminate redundancy.
Synchronizing and individualizing: when the system is overloaded by the need to hold information in working memory.	Best solution would be to synchronize the presentation of corresponding visual and auditory material.

Table 2. 3: Scenarios and solutions for cognitive overload

They came to the conclusion that cognitive load can be reduced by re-distributing essential processing across more than one sensory channel or reducing the amount of incidental and informational holding needed. This work was based on 12 years of research. I will discuss two of the studies.

4.1.3 Aids to multimedia learning

Mayer and Moreno (2003), researched the techniques that aid multimedia learning. They based their researched on the premise in the cognitive load theory that instructional messages should be designed in ways that minimise the chances of overloading the learner's cognitive system.

Moreno and Mayer (2000) also propose that multimedia learning involves three cognitive processes namely selecting, organising and integrating.

- Selecting is the process where verbal information is processed as text and visual information is processed as images.
- Organising is the process where the verbal base and the image base are applied to the yet to be learned concept.
- Integrating is the process where the learner builds connections between the two.

They were guided in their research by a cognitive theory of multimedia learning which draws on dual coding theory, cognitive load theory, and constructivist learning theory. From dual coding theory they took the idea that visual and verbal materials are processed in different processing systems (Clark & Paivio, 1991; Paivio, 1986).

From cognitive load theory they concluded that presenting too many elements that need to be processed in visual or verbal working (i.e., too many words or too complex a picture) can lead to overload where some of the elements are then not processed. From constructivist learning theory they reach the conclusion that cognitive construction depends on the cognitive processing of the learner during learning. Mayer and Moreno summarised some of their empirical research on multimedia learning that supports and clarifies the basic tenets of cognitive load theory that working memory can process only a few elements at any one time.

They also proposed a series of design principles that are sensitive to cognitive load issues as summarised in Table 2.4 (Mayer & Moreno, 2003).

Type of aid	Description
Multimedia aids	Use narration and animation rather than narration alone.
Contiguity aids	Present corresponding narration and animation simultaneously rather than successively.
Coherence aids	Eliminate unneeded words and sounds.
Modality aids	Present words and narration rather than on screen text only.
Redundancy aids	Present narration and animation rather than narration, animation and on-screen text.

Table 2. 4: Aids to computer-based multimedia learning

The design of the multimedia program used in this study was based on these principles.

4.2 Cognitive load and cognitive load theory

In the previous article that I discussed, Mayer and Moreno (2003) proposed design principles that are sensitive to cognitive load. In 1998 Cooper did research on some of the literature available on the subject and then outlined the basic principles of the cognitive load theory. He provided examples of the instructional design strategies generated by cognitive load theory. His work is discussed next.

4.2.1 Research into cognitive load and cognitive load theory and instructional design at UNSW

Although there is no definite research question, the author supplied information on and investigated the cognitive load theory and how it correlated with current knowledge regarding memory, thought, learning and problem solving.

He first described the concepts out of the literature concerning memory as described in paragraph 3.2.3. Then he looked into the concept of learning and described learning, the process of learning and what novices need to learn to become an expert. He came to the conclusion that learning happens when schemas are built. This refers to the learning process where information must first be attended to, and processed by working memory before it can be encoded into schemas.

His research on the cognitive load theory provided information regarding reasons why some material is difficult to learn and how to apply the cognitive load theory to instructional design. He used examples out of previous research to eventually discuss the effects generated by cognitive load theory and then described the benefits for learning.

Cooper came to the conclusion that cognitive load theory views the limitations of working memory to be the primary obstruction to learning. He claimed that by reducing total cognitive load imposed by a body of to-be-learned information, it increased the portion of working memory which is then available to attend to the learning process. To achieve this, extraneous cognitive load levels should be reduced through instructional design. He summarised his findings by comparing standard practice and the cognitive load generated effect, as can be seen in Table 2.5. He found that the effects generated by cognitive load theory often "fly in the face" (as he puts it) of standard practices. This observation attests to the strength of the theory.

Standard practice	Cognitive load generated effect
Specify the goal of a problem so that learners know what they have to find.	*The goal free effect* Use goal free problems
Learners need to repeat because 'practice makes perfect'.	*The worked example effect* Learners study worked examples and problem solving is used to test if learning has been effective.
Instructional material that requires both textual and graphical sources of instruction should be clear and simple and the text and graphics should be located separately.	*The split attention effect* If both text and graphics are used, instructional design should integrate the text into the graphic in such a way that the relationship between textual and graphical components is clearly defined.
Similar to be learned information should be presented using identical media format to ensure consistency in the instructional presentation	*The modality effect* Use multimedia to present some content visually and others auditorily.
The same information should be presented in several different ways at the same time.	*The redundancy effect* Simultaneous presentations of redundant content must be avoided.

Table 2. 5: Standard practise against cognitive load generated effect

The effects correspond with the principles as defined by Mayer and Moreno (2003).

Cooper suggested that the effects generated by cognitive load theory should be viewed as "rules of thumb" rather than absolute "laws of instruction". He concluded that in the end, according to cognitive load theory, there will always be the need to reduce total cognitive load, and the need to maximise cognitive resources available to be utilised in the learning process. If for some reason cognitive load increases rather than decreases, then learning will be inhibited.

These effects were taken into consideration for the design of this study, in an attempt to control extraneous cognitive load.

4.3 Measuring cognitive load

There are a limited number of studies that measures cognitive load directly. I found the work of Brünken, Plass and Leutner (Brünken &Leuthner, 2001; Brünken, Steinbacher, Plass & Leuthner, 2002; Brünken, Plass & Leutner, 2003 & 2004) relevant to my study and discuss two of them in detail.

4.3.1 Direct measurement of cognitive load in multimedia learning

The measuring of cognitive load in multimedia learning was examined by Brünken, Plass and Leutner in 2003. They distinguished between direct and indirect performance measures; subjective ratings; and behavioural, physiological, and neuro-imaging measures. This research introduced a new direct measure of cognitive load in multimedia learning. The aim of the study was to determine whether cognitive load could be measured using the dual task approach.

The measurement method involved using a continuous visual monitoring task as the secondary task. This allowed the measurement of cognitive load in the visual system.

Two experiments were conducted in 2001 to demonstrate the feasibility of this approach. The content for the first experiment was a multimedia lesson on the cardiovascular system and for the second experiment they used a multimedia travel guide.

The secondary task was a simple visual-monitoring task which required the learners to respond every time the task happened by hitting the spacebar.

The secondary task was placed above the primary task frame. Within this small frame, a single letter was displayed. Occasionally, the letter's colour changed from black to red, indicating a response request. Software recorded all the reaction times. The experiments included three experimental conditions and participants were randomly assigned to an experimental condition. These conditions were:

- a single-task condition with the secondary task alone,
- a dual task condition with visual-only learning material as primary task and
- a dual task condition with audiovisual learning material.

Within each condition, repeated measures of reaction time were taken at random intervals.

The results of the two experiments demonstrated the feasibility of this approach. The performance on the secondary task was significantly faster in the single task condition than in the dual task conditions. Comparing the dual task conditions against each other the reaction times were significantly faster for the audiovisual condition than for the visual-only condition. This was in line with the cognitive load theory and was expected.

Their results indicated that the dual task method was an effective way to measure cognitive load, although there are many other variables which influence the learning outcomes that should be taken into consideration.

The dual task method is also applied in my study.

4.3.2 Assessment of cognitive load in multimedia learning with the dual task methodology: auditory load and modality effects

The purpose of this study by Brünken, Plass and Leutner (2004) was to investigate whether or not the limited capacity assumption of cognitive load theory (CLT) or cognitive theory and multimedia learning (CTML) could be validated with a more direct measurement of resource demands in multimedia learning. They wanted to explore auditory load in the modality effect and to examine a closely related instructional design.

The main question was whether the limited capacity assumption of CLT and CTML, which has primarily been studied using learning outcome measures, could be validated with a more direct measurement of resource demands in multimedia learning. Sub-questions are:

- Would the audiovisual presentation of verbal and pictorial learning materials lead to a higher demand on phonological cognitive capacities than the visual-only presentation of the same material?
- Would adding seductive background music to an audiovisual information presentation increase the phonological cognitive load?

The researchers used cognitive load theory and cognitive theory of multimedia learning as a framework to conducted two within-subject experiments with 10 participants each.

They used the dual task methodology in order to achieve a direct measurement of cognitive load in the phonological system. They conducted two experiments that used the same experimental design, same secondary task but different learning contents of the multimedia learning system as primary task.

Experiment 1

Ten female students, enrolled for the undergraduate (BA) program at Erfurt University with a mean age of 20.9 years (SD = 1.45) participated in the study. The design of the experiment was a within subject of all the participants and the blood circulation system was chosen as topic for the program.

The learning content was classified according to the amount of auditory information of the primary task and had three levels:

- no auditory information,
- background music only and
- background music and narration.

Several repeated measures of the secondary task performance were taken within each level of this independent variable.

The dependent variable was the performance on the secondary task. This variable was measured as a reaction time. The primary task was to gain knowledge on the subject and was measured as a control variable using a pre-test and a post-test.

A program called WinRT (Brünken et al. 2002), was used to detect the response to a simple auditory stimulus. A single tone was presented to the learner at random intervals of 5 to 10 seconds. The learners were instructed to press the space bar on their computer keyboard as soon as they detected the tone. The computer program automatically recorded the lag time between the presentation of the tone and the learners' reaction.

The results of the experiment were in line with the CTML and CLT concerning the modality effect. In short the results show that secondary task performance decreased when in addition to the auditory secondary task, background music and narration had to be processed simultaneously, but not when only background music had to be processed.

Experiment 2

Brünken, Plass and Leutner duplicated the results of experiment 1, by conducting a second investigation where they used the same experimental design, the same secondary task and comparable participants as in experiment 1, but different learning materials.

For this experiment they used 10 female students, enrolled for the undergraduate (BA) program at Erfurt University, with a mean age of 25.6 years (SD = 4.57).

The cognitive load in this experiment was also measured using the dual task method and measured reaction time to a secondary task. The primary task was a multimedia tourist guide containing verbal and pictorial information about the historic city of Florence, Italy. The secondary task was an auditory stimulus that the participants had to react to by pressing the space bar on the keyboard. The difference between the time the stimulus was activated and the time the student reacted was recorded electronically.

Experiment 2 replicated the findings from experiment 1, showing that the performance in an auditory secondary task decreases when the primary task requires the simultaneous processing of verbal information and background music, but does not decrease when, simultaneous to the secondary task, only background music is presented in the primary task, and the verbal information is presented visually.

The experiments support the theory that the auditory presentation of verbal information requires specific cognitive resources and therefore decreases the available cognitive resources for actual learning.

Up to now this effect was only derived from an analysis of learning outcomes. This research demonstrated that a more direct method of measuring the demand on cognitive capacity can be used to underpin the CLT. In the wider study that this study is part of, the cognitive load and the learning outcomes are measured.

4.3.3 Conclusion

These studies on measuring cognitive load research in multimedia learning, show that the challenges of the dual task method lie in the methodological and technical issues that have to be taken into consideration (Brünken, Plass & Leutner, 2004).

A descriptive summary of the challenges and how they were addressed in this study is described in Table 2.6.

Challenge	Addressed in this study
The secondary task has to require the same cognitive resources as the primary task; otherwise, secondary task performance will be independent of primary task performance.	The secondary task is integrated into the primary task.
The performance measure for the secondary task has to be reliable and valid.	Reaction time is used as the proven performance measure for the secondary task. (Verwey & Veltman, 1996, Wickens, 1984).
The secondary task has to be so simple that it does not suppress simultaneous learning processes.	The participants only have to press enter on the keyboard when they notice that a symbol change colour.
The secondary task has to be able to consume flexibly all of the available free cognitive capacity.	

Table 2. 6: Challenges of the dual task method

The literature shows different subjective ways to measure cognitive load for example ratings and questionnaires. The dual task approach seems to be the most promising for the direct measurement of extraneous cognitive load where different multimedia instruction induces this extraneous cognitive load in the working memory (Brünken, Plass & Leutner, 2003). The dual task method is used in this study where multimedia and the influence of instructional design and cognitive styles is under surveillance for the wider study. All the studies done by Brünken et al (2001 -2003), used a sample size in the range 10 – 32 in a controlled environment. The literature did not provide studies with extensively larger samples. A strength of this study is the sample size: a sample of 238 participants in an authentic learning environment was used.

The research methodology and design used for this study is described in the next chapter.

Chapter 3: Research methodology

1. Introduction

In this chapter I discuss the research design and methodology. The chapter is divided into two sections as illustrated in Figure 3.1.

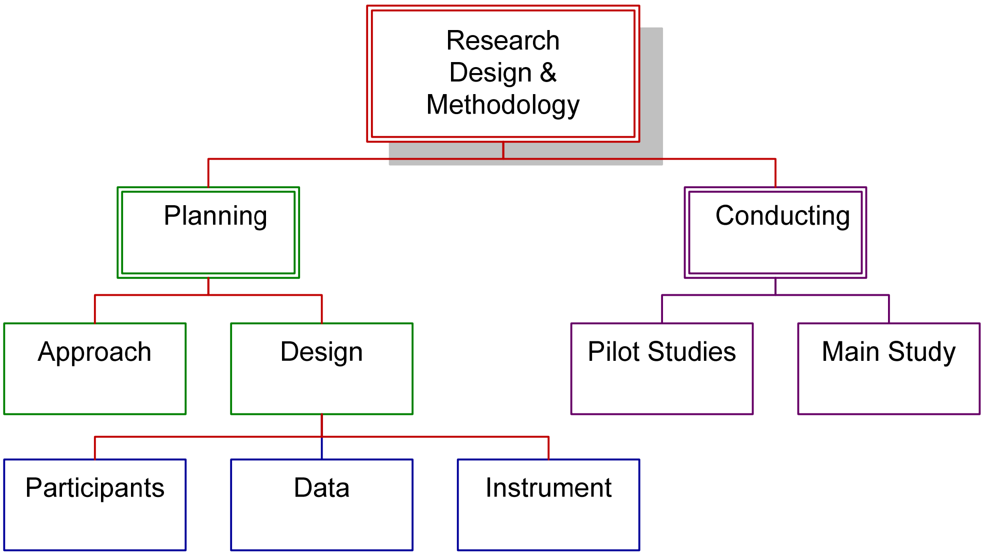


Figure 3. 1: Structure of Chapter 3

The section 'Planning' includes a description of the learning intervention used in the study. The way in which the study was conducted is explained in the section 'Conducting'.

2. Planning the research

2.1 Purpose of the study

The purpose of this study was twofold: firstly, the study measured (using the dual task approach) the cognitive load of two multimedia programs that were developed to present the same content, and secondly the study explored one factor that could influence this measurement of cognitive load.

2.2 Research questions

The research questions are:

1. How does the position of the secondary task influence the measurement of cognitive load?
2. What is the cognitive load of the different presentation formats of a multimedia program, measured using the direct measurement technique?

A quantitative approach, using an experimental design, was used in this study. The study comprised of two pilot studies and the main study. The pilot studies were done in order to test the design of the study and the data collection process.

2.3 Research approach

There are two major approaches to research, quantitative and qualitative (Trochim, 2005).

The real difference between qualitative research and quantitative research is not so much the method, but the researcher's approach. The fundamental characteristics of the approaches and how they apply to my study are summarised in Table 3.1.

Quantitative	This study	Qualitative	This study
Systematic design	√	Systematic design	√
Objective approach	√	Subjective approach	
Deductive logic	√	Inductive logic	
Generalise-able intervention	√	Not Generalise-able intervention	
Numeric data	√	Words	

Table 3.1: Quantitative vs Qualitative Research

This study meets the conditions for a quantitative approach. These conditions are described in Table 3.2.

Quantitative conditions	Relevance to this study
The correlation between variables can be quantified.	The variables are trigger time and response time and the difference is calculated in time units.
The relationship between variables is explored.	The relationship between the dependent variable, the cognitive load and independent variables such as the presentation format of the content and the position of the secondary task is determined.
The research is confirmatory rather than exploratory.	Numerical data from earlier research is available and this study elaborates on that.
The research measures a trend.	The trend of participants' response to the cognitive load of two different learning formats is measured.
There is no ambiguity about the concepts being measured.	The concepts have been clearly defined.
The concept is measured on a ratio or ordinal scale.	The cognitive load of the different multimedia formats was measured in time units using a direct measurement technique, measuring reaction time. The numeric data was analysed using statistical procedures.

Table 3.2: Applicable quantitative conditions

Advantages of using quantitative research as applied in this study are summarised in Table 3.3 (Trochim, 2005).

Advantage	Applied in this study
Reliability: the underlying principle of quantitative research is that the results are an accurate representation of the population being studied.	The Intervention can be presented to any sample of the population group and it can be determined if there is variation in the results.
Large-Scale: to provide increased confidence that the results are not an anomaly, quantitative studies use considerably more participants than qualitative research methods do. The results are applicable to the population.	A larger group of participants was used than has been used in similar research.
Consistency: quantitative studies necessitate the standardization of the data collecting method and include respondents from the entire target population. It involves the use of a specific designed instrument with structured questions.	The data collecting method is standardised and stable.
In-depth analysis: this is possible because an array of statistical techniques can be applied to quantitative data. Reviewing simple averages or frequency distributions is often enough to provide valuable insight into the nature of a population.	This study used both descriptive and inferential techniques to answer the research questions. The Means procedure and General Linear Method procedure are used to explore the relationships between the variables.

Table 3.3: Advantages of using quantitative research approach

On the other hand, some of the factors that make quantitative research reliable and easily replicated can be drawbacks, as discussed in Table 3.4.

Disadvantages	Precautions taken in this study
Recruitment: successful random sampling depends on a sufficiently large sample (Thompson, 1999). Difficulty in recruiting subjects to participate in the pilot studies as well as the main study is remarkably common (McMahon, 1994).	The content of the learning intervention is part of their study programme. As such it was possible to use a large sample. The research study was conducted in the student's academic time. The size of the population is sufficient to have an extensive sample group even if only a small percentage participates.
Limited for exploration: as a rule, quantitative research is not suitable for initial learning, or even as a method to develop ideas.	This quantitative research determines the cognitive load per screen using existing techniques. It does not explore new ideas on techniques.
Limited responses: the standardization of the instrument tends to limit testing to predetermined hypotheses. Therefore, some potentially interesting spontaneous or tangential responses may be missed or excluded.	The instrument measures cognitive load using a secondary task and a questionnaire. This study only reports on the cognitive load measured using the secondary task. Although this study misses some spontaneous responses that would have been included with a qualitative approach, the goal was to obtain objective results and spontaneous responses are therefore a justified sacrifice.
Accessibility: quantitative data is an abstraction, and as such can be difficult for some decision makers to relate to. Lack of ability to interpret quantitative research drives researchers to conduct qualitative research instead when more rigorous quantitative methods would be cheaper and more appropriate.	This will remain a challenge.

Table 3.4: Disadvantages and precaution steps taken

2.4 Design

The research design is used to structure the research and includes information on the type of research, participants, measurement instrument, interventions, and methods of assignment. The main purpose of the design is to address the research questions.

There are a number of possible quantitative research designs. It can be difficult to do pure, experimental research, especially in the human sciences, because of the kind of variables and ethical considerations. Therefore adaptations of experimental designs have been developed, for example quasi-experimental and non-experimental designs. Different types of quantitative research design, as described by Trochim (2005), are summarised in Table 3.5.

Research design	Explanation
Randomised or true experimental design.	Random assignment of participants is used.
Quasi-experimental design.	Specific experimental methods are used but subjects are not randomised.
Experimental design.	Random assignment of subjects and the manipulation of variables are used to determine cause and effect.
Non – experimental design.	A study of completed research findings is used.
Descriptive design.	Subjects are usually measured once only and the associations between variables are established.

Table 3.5: Types of research designs

For the purpose of this study, I used a descriptive experimental design.

Figure 3.2 illustrates the research design, methodology and plan for this study.

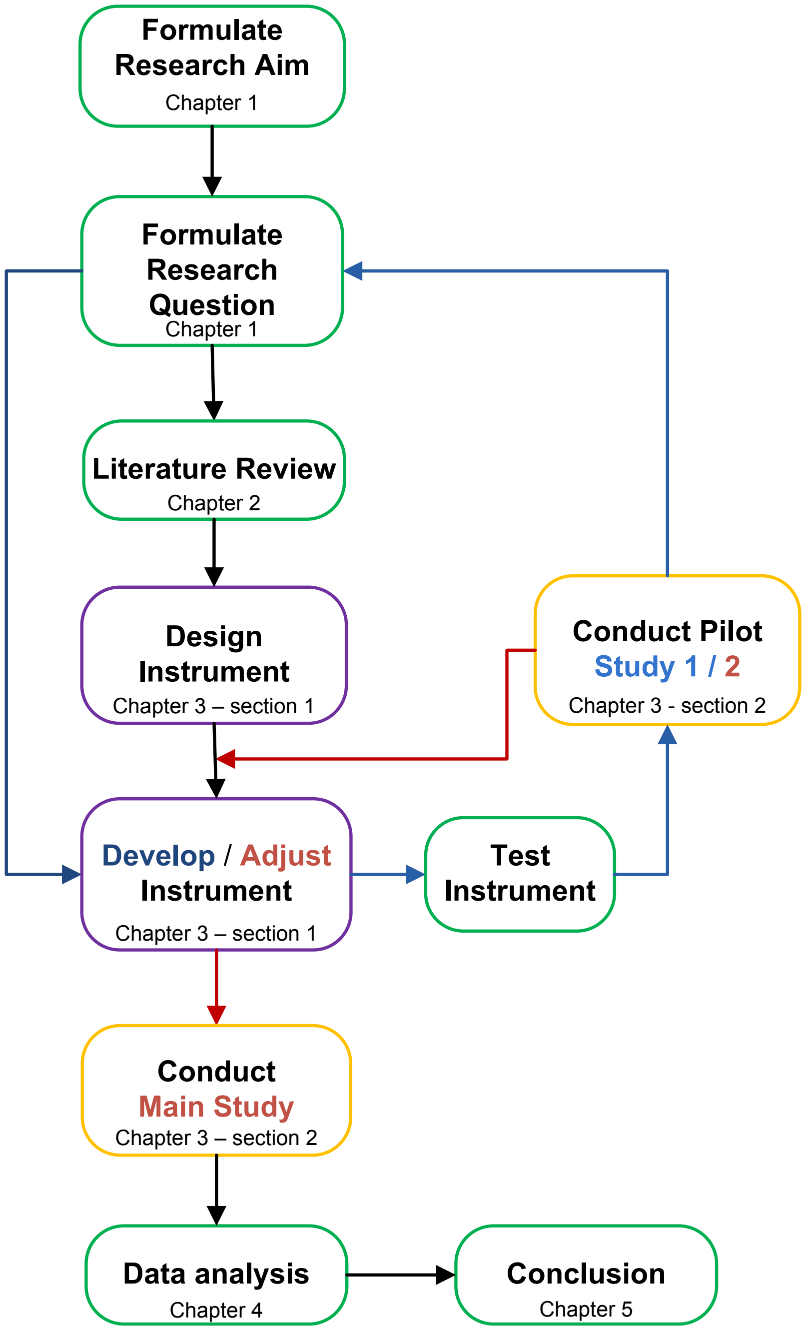


Figure 3. 2: Illustration of the research methodology

2.4.1 Participants

Population

The unit of analysis was students from the Faculty of Health Sciences of the University of Pretoria. The population was made up of all students who take Physiology as part of their study programme.

Sampling

Initially a purposive sampling method was used to select the unit of analysis. The group of students was selected because the content used for the study was part of their programme at the university.

Participants were drawn from the undergraduate students who take Physiology as a subject in their study programme.

The sample for the pilot study was 38 second year physiotherapy students. Because the data did not write to the data files as expected during the first pilot study, a second pilot study was necessary before the main study. For the second pilot study, volunteers were sought from the first year student group at various residences on the Faculty of Education campus of the University of Pretoria.

For the main study a sample of 238 second year medical students and dental students, who are doing the same course in Physiology, was used.

Once the sample was identified, the participants were randomly assigned to one of the four interventions.

2.4.2 Data collection plan

Data collection instruments for quantitative research are developed specifically for the study or can be obtained from another source (Trochim, 2005). In this approach, the data is obtained through a structured research intervention that was developed especially for a wider study. The research instruments were electronic and all data collection was done electronically.

Protocol for measuring cognitive load

A review of the literature indicates that measuring reaction time is a suitable method to determine cognitive load. The idea is based on the fact that the learner has to react to a very simple signal as soon as possible. This secondary task does not hold back the primary task but yet the instantaneous reaction uses up all available cognitive resources (Brünken, Plass & Leutner, 2003).

The secondary task in this study used a symbol, the letter A, which was displayed on the screen. The symbol changed colour at specific time intervals. In this study the first colour change was programmed to occur four seconds after the screen was loaded and every ten seconds thereafter. The number of times this change occurred depended on the design of the screen. For the purpose of description and data collection this change was called the Trigger. Table 3.6 describes the screens where cognitive load was measured directly and the number of triggers per screen.

Program	Version 1 Animations		Version 2 Static images & text	
	Screen	No of triggers	Screen	No of triggers
The content and format were equivalent on these screens.	1 - 3	0	1- 3	0
	4	4	4	4
	5	3	5	3
	6-7	0	6-7	0
	8	4	8	4
	9	0	9	0
	10	3	10	3
	11	3	11	3
	14	4	21	4
	15	0	22	0
	16&17	0 & 6	17 & 18	0 & 6
	17	6	18	6
	19	8	23	8
The content was equivalent but the presentation format was different. Version 1 used one screen to present the content as animation and Version 2 used 5 screens to present the same content as static images and text.	12v1	8	12v2, 13 14 15 16	0 4 3 3 5
The content was equivalent but the presentation format was different, only one screen used to present the content.	13	5	20	5
	18	0	19	0

Table 3.6: Comparison of different formats.

The participant was required to press the ENTER key each time he/she noticed the symbol changing colour. For the purpose of description and data collection this event was called the Response.

The time of each Trigger was recorded electronically and the time of the Response following the Trigger, if this event took place, was recorded.

Because the learner had to attend to the primary and secondary task simultaneously it is possible to measure cognitive load at the very point and time it is induced on the learner.

The secondary task was independent from the content of the learning program. The simplicity of the secondary task makes it a reliable way to measure cognitive load as it should not take up any or very little of the working memory to perform. Previous studies have shown that reaction time is a valid direct measurement method of cognitive load (Brünken, Plass & Leutner, 2004).

Type of data

The data was collected electronically via the multimedia programs. The programming was done in such a way that the following data was recorded in an external log file:

- time the lesson was accessed,
- time the lesson was exited,
- time a screen was entered,
- time the learner left a screen,
- trigger time of the secondary task and
- response time of the learner to the secondary task.

The data was written out to an .ini file, together with demographic data, style analysis results, pre- and post-test results and the subjective measurement of the cognitive load according to a 9-point rating scale. Only the demographic data is used to describe the sample in this study. The rest of the data was used in the wider study.

An example of such an .ini file is broken down and shown in Figures 3.3 and 3.4. Figure 3.3 illustrates the recording of the demographic data.

```
[Demographic Data2]
Version V2=2
[Demographic Data3]
Age V3=21
[Demographic Data4]
Gender V4=1
[Demographic Data5]
Culture V5=1
[Course Detail 1]
Programme V6=BCHD
[Course Detail 2]
Year of study V7=2
[Lesson Detail]
Prior Know V8=1
[Self rating]
Self rating V9=2
[Language]
Language V10=2
```

Figure 3. 3: Variables for demographic data.

Data for age, gender, culture, language and year of study was recorded.

Figure 3.4 illustrates the recording of the trigger times of the secondary task.

```
[Trigger7_1]
Trigger1=13:19:1713:27:22

[Trigger7_2]
Trigger2=13:19:2713:27:32

[Trigger7_3]
Trigger3=13:19:3713:27:42

[Trigger7_4]
Trigger4=13:19:4713:27:52

[Trigger7_5]
Trigger5=

[Trigger8_1]
Trigger1=13:20:38

[Trigger8_2]
Trigger2=13:20:48

[Trigger8_3]
Trigger3=13:20:58

[Trigger8_4]
Trigger4=
```

Figure 3. 4: Trigger times

Figure 3.4 lists the data for screens 17 and 18. The trigger events for these screens were numbered Trigger7 and Trigger8 respectively. The value after the underscore indicates the number of the trigger on the screen. Trigger7_1 is therefore the first trigger for screen 14, Trigger7_2 the second trigger and so forth. Table 3.4 indicates that the trigger should have occurred 6 times on screen 17 and 18. The participant exited the screen before all the triggers could occur. No time was recorded for Trigger 8_4 because the participant exited before the trigger occurred.

If the participant entered the screen more than once the trigger times were recorded in sequence. The example shows that trigger 7_1 was triggered at 13:19:17 and again at 13:27:22 indicating that the participant entered the screen twice.

Note that the time for trigger 7_2 was exactly 10 seconds after 7_1.

The instrument also measured the Response. Figure 3.5 illustrates the section of the .ini file where the responses to the triggers were recorded.

```
[sScreen14]
13:19:13=Accessed
13:20:34=Exited
13:27:18=Accessed
13:27:58=Exited
[Responses7_1]
13:19:19=Hit space bar
13:27:24=Hit space bar
[Responses7_2]
13:19:28=Hit space bar
13:27:33=Hit space bar
13:27:39=Hit space bar
[Responses7_3]
13:19:39=Hit space bar
13:27:45=Hit space bar
[Responses7_4]
13:19:48=Hit space bar
13:27:55=Hit space bar

[sScreen15]
13:20:34=Accessed
13:21:29=Exited
13:27:58=Accessed
13:28:00=Exited
[Responses8_1]
13:20:40=Hit space bar
[Responses8_2]
13:20:49=Hit space bar
[Responses8_3]
13:20:59=Hit space bar
```

Figure 3. 5: Response times

Note that trigger 7_1 was triggered twice and there are two corresponding responses. The data was verified against the time the screen was accessed and exited.

If the participant did not react to the triggers, only the time the participant accessed and exited the screen was recorded as shown in Figure 3.6

```
[sScreen17]
11:37:39=Accessed
11:39:37=Exited

[sScreen18]
11:39:38=Accessed
11:42:27=Exited
```

Figure 3. 6: No responses

Preparing the data for analysis

On conclusion of the experiment, the .ini files were downloaded from each participant's individual computer and the data was checked and verified.

Four data sets were gathered out of the .ini files, as described in Table 3.7.

The student number (V1) was the connection between data sets and a screen log kept record of the sequence in which the participant accessed the screens.

Data set	Description	Variables	Verified against
MESA	Demographic data	V1 – 10, V40 – V44	
MESB	Trigger times	V1, Lesson #, Lesson started, Screen #, Access time, Trigger #, Trigger time	First trigger should be 4 seconds after access time. Next triggers should be on intervals of 10 seconds. Number of triggers per screen was fixed.
MESC	Response times	V1, Lesson #, Screen #, Trigger #, Entry, Response time	The response times were verified after the combining MESC with MESB.
MESD	Access & Exit times on screens	V1, Lesson #, Lesson started, Entry , Screen #, Access time, Exit time	Access time should be after Lesson started time. Access and Exit times should be in sequence with the screen log.

Table 3.7: Data sets

Formula used to calculate cognitive load

The following protocol was used to determine the cognitive load per screen (CLS):

CLS = AVERAGE(($RESPONSETIME_1$ -$TRIGGERTIME_1$) + ($RESPONSETIME_2$ – $TRIGGERTIME_2$) +…($RESPONSETIME_x$ –$TRIGGERTIME_x$) where x = the number of times the trigger changed colour on the screen. If there was no response time but a trigger time, the difference was calculated as the maximum, which was10 seconds.

Using this protocol the following were calculated:

1. The cognitive load for each participant per screen (CLPS):

   CLP#$S_\#$ = AVERAGE(($RESPONSETIME_1$ -$TRIGGERTIME_1$) + ($RESPONSETIME_2$ – $TRIGGERTIME_2$) +…($RESPONSETIME_t$ –$TRIGGERTIME_t$) where t = the number of times the participant entered S#.

3. The cognitive load for each screen per Lesson (CLL#S#):

   CLL#S# = AVERAGE(CLP_1S# + CLP_2S# + ... CLP_mS#) where m = the number of participants entered L#

4. The cognitive load for each per screen per Version (CLV#S#):

   CLV1S# = AVERAGE(CLL1S# + CLL2S#)

   CLV2S# = AVERAGE(CLL3S# + CLL4S#)

5. Cognitive load per participant ($TCLP_\#$):

   TCLP# = Average(CLP#S_1 + CLP#S_2+ …. CLP#S_n) where n = the number of screens where cognitive load was measured.

6. The cognitive load per Version (TCLV#):

   TCLV# = AVERAGE ($TCLP_1$ + $TCLP_2$ + …. $TCLP_m$) where m = the number of participants who did the version.

2.4.3 The Instrument

A multimedia program, designed and developed by Strehler (2007), was used for this study. The program presents content on the Autonomic Nervous System. The complete program is illustrated in Annexure B. The sequence of the program was introduced in a main menu. The participant was required to log on using his/her student number.

The next sequential screen was a practice screen, illustrated in Figure 3.7.

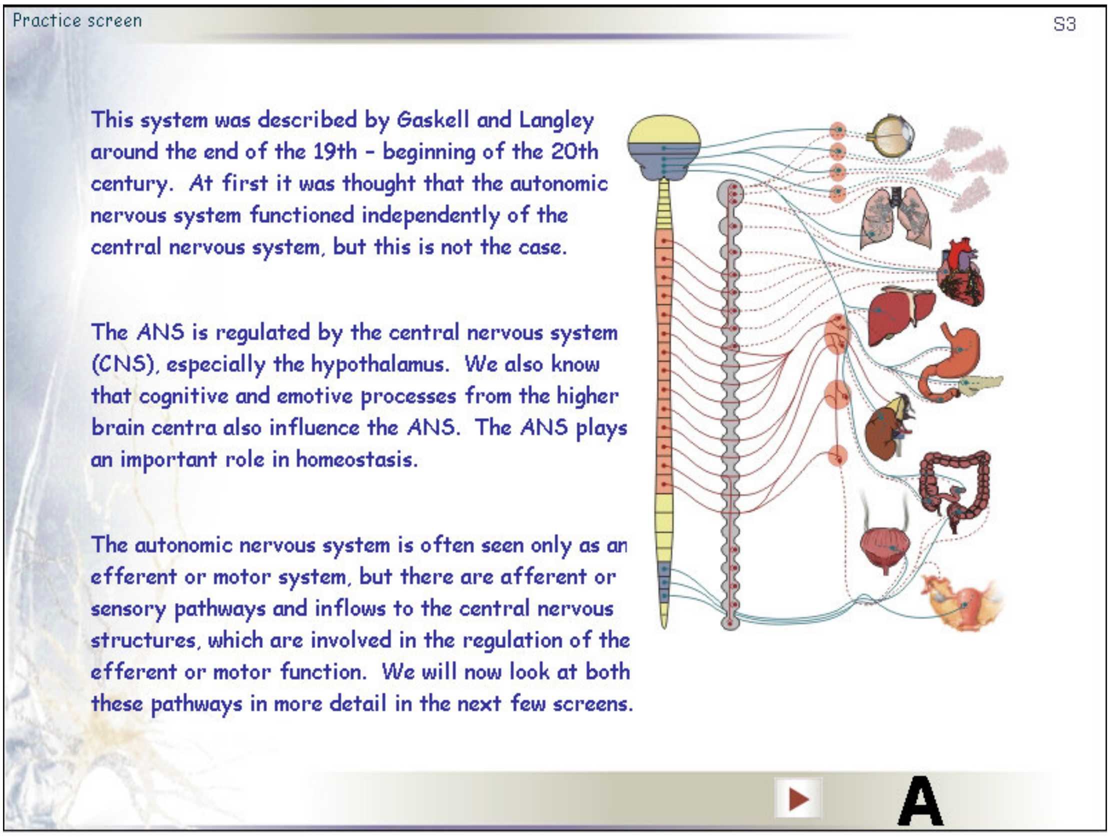


Figure 3. 7: Practice screen

The participants were given an opportunity to practice responding to the trigger in an effort to control the extraneous load that this secondary task might cause. It was made clear that the most important task was to study the content and not to respond to the trigger. The trigger times and response times for this screen were recorded but not used in the study.

The primary task

The program presents content on the Automatic Nervous System. The participant was required to study the content prior to taking a post test.

The secondary task

The capital letter A was used as the trigger and placed on different positions on some of the screens. If the symbol changed colour and the participant noticed this change they were required to press the ENTER key. In the first pilot study the symbol changed colour from black to purple, illustrated in Figure 3.8. During the pilot study I observed that the difference between the colours that were used was too small.

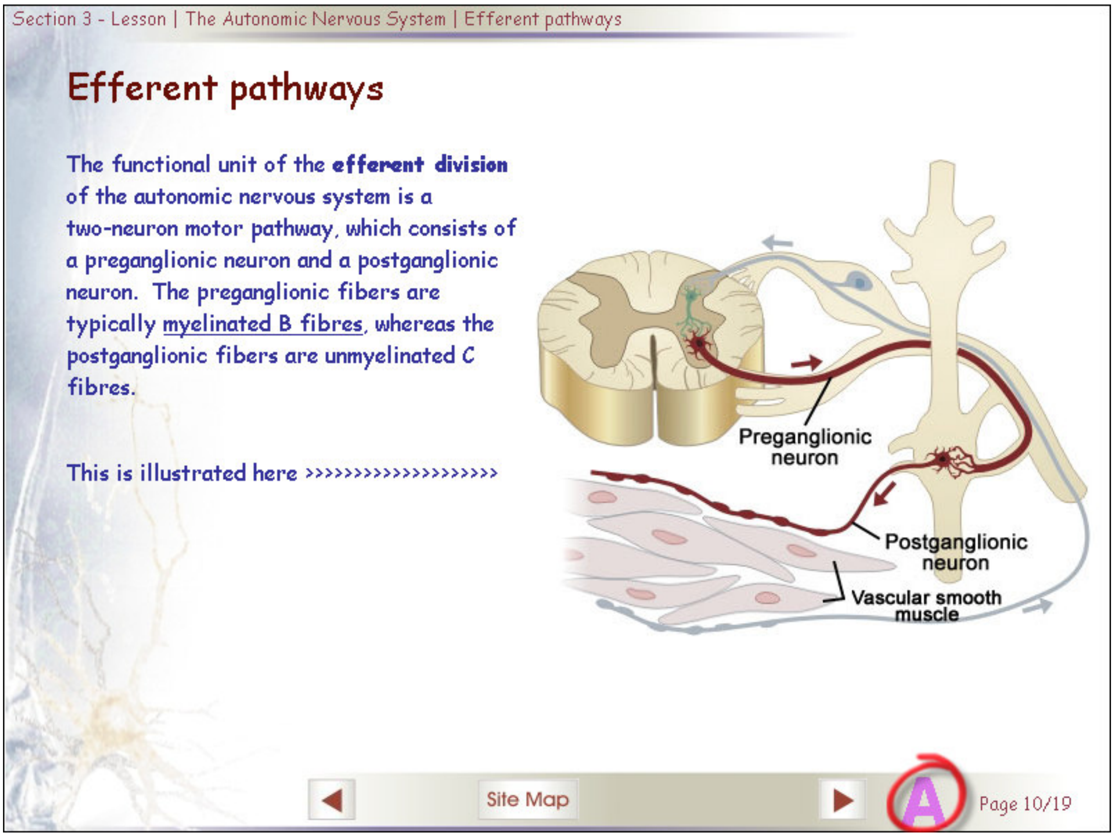


Figure 3. 8: First secondary task symbol

The program was then adjusted so that the symbol changed from black to green as illustrated in Figure 3.9.

Figure 3. 9: New colour of the secondary task symbol

Initially the symbol was placed in the bottom right corner as shown in Figure 3.9. From Pilot Study 1 it seemed that the position of the symbol might also have an influence on the response to the Trigger. The two versions of the program were duplicated so that for each version there were two formats to furthermore test the influence on cognitive load if the position is in the top right corner as shown in Figure 3.10.

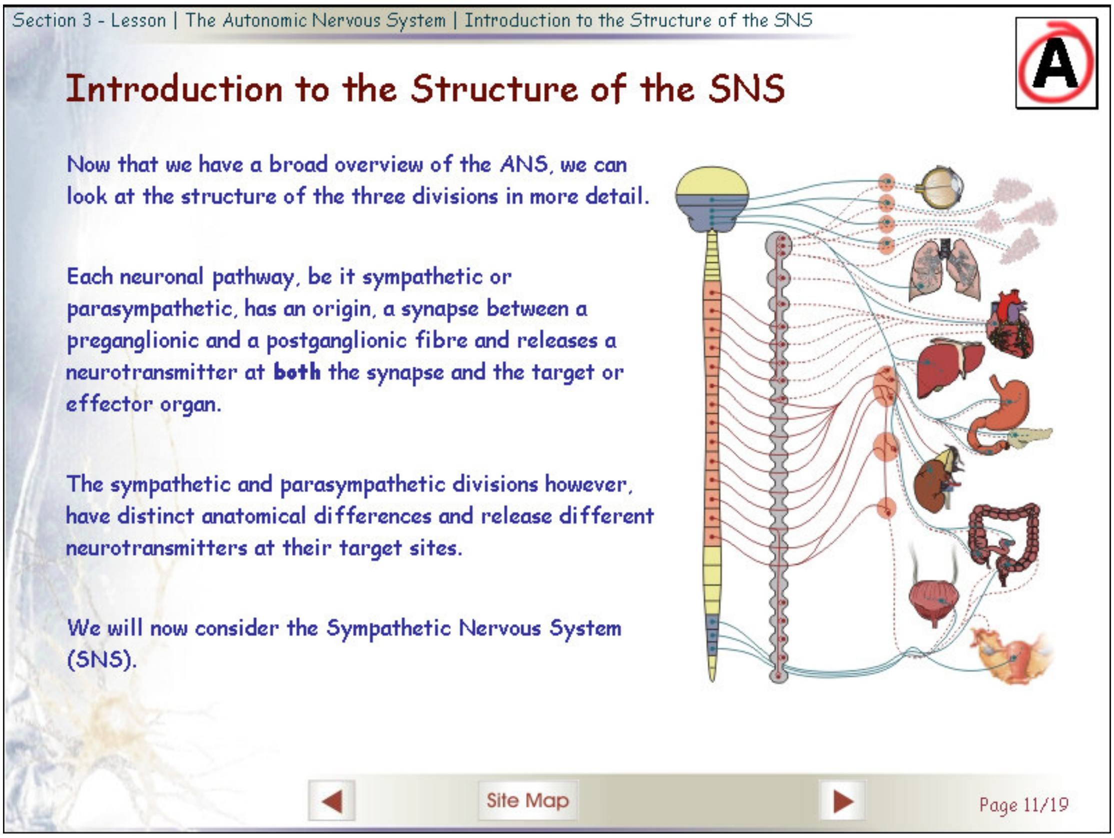


Figure 3. 10: Example of a screen with the symbol in the right top corner

The formats differed only in respect of the position of the secondary task on the screen as can be seen in Table 3.8.

Program	Presentation format	Position of the secondary task
1.1	Graphics, text, audio and animations	Bottom right
1.2	Graphics, text, audio and animations	Top right
2.1	Text and static images only	Bottom right
2.2	Text and static images only	Top right

Table 3.8: Position of the trigger

3. Conducting the study

3.1 Ethical considerations

There were three main ethical issues that were considered in the wider study.

- How were the participants' anonymity ensured?
- Should participants be informed about their particular cognitive style?
- Should all the formats be made available to the participants (particularly to those who did not learn using the better design with a lower cognitive load)?

The ethical issues in my study were the anonymity of the participants as well as the availability of the different formats to all participants.

- The measures to be put in place to obtain data anonymity might be cumbersome, but they are unavoidable if the rights of participants are to be protected. Each student received a letter of invitation to participate in the study in which the purpose of the study was explained. They were informed that their student numbers would be used in the measurement of cognitive style, self-report measures and the pre- and post-test measures, their individual student numbers would not be used in any research report. I furthermore requested each student to indicate their willingness to participate in the study by signing a consent form attached to the letter of invitation.
- To ensure that all participants had access to the format with a lower cognitive load, the learning material was made available in all the formats to all participants after the study.

3.2 Permission

Permission to conduct the study in the proposed manner was obtained from the head of the Department of Physiology and from the students (Annexure C).

3.3 Procedure

Figure 3.11 displays the sequence and time line in which the studies were conducted.

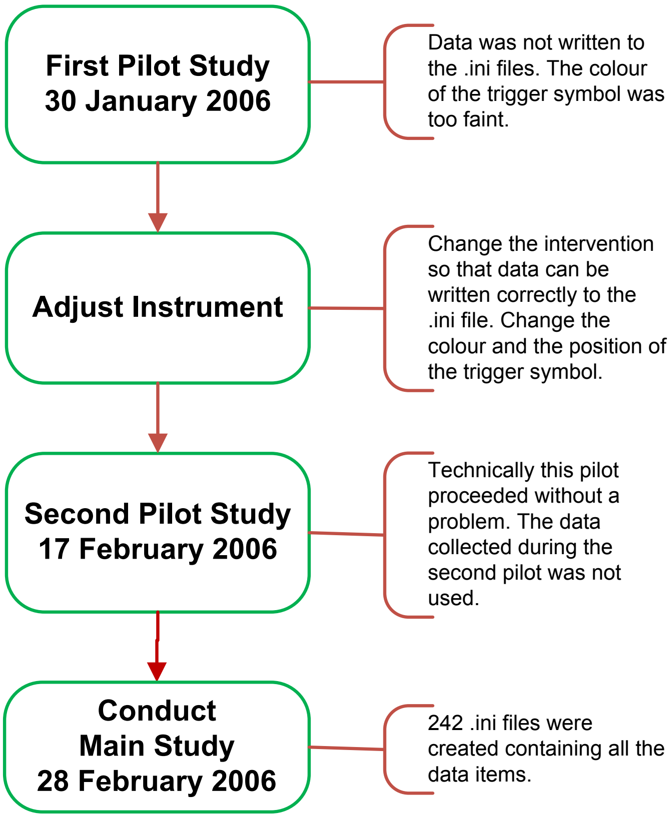


Figure 3. 11: Conducting the study

The same procedure was followed in all the studies. This procedure included the following steps:

1. Participants were randomly grouped into 4 subgroups.
2. Each subgroup was presented with a different lesson.
3. Lessons were uploaded onto the computers in the computer lab.
4. Participants were guided to their seats.
5. Participants were informed about what to expect using an informative presentation. The participants were briefed on the following:
   - If they noticed the symbol changing colour they had to press the ENTER key once on the keyboard.
   - No results would be observed on the screen after pressing ENTER.
   - Looking out for the colour change was not the purpose of the program – their main task was to learn the content.
6. Participants had the opportunity to practise using the practice screen.
7. The participants completed the program in the following sequence:
   - completed a styles test,

- signed on to the intervention using their student numbers,
- read through the instructions,
- practiced on the practice screen,
- completed the demographic data,
- completed the pre-test,
- worked through the content,
- completed the post test,
- completed a final survey and
- completed the paper-and-pencil section of the post-test.

8. At first the students had to follow a linear path through the program. After that they could navigate between the screens as they wished. Every time they accessed or exited a screen the time was recorded.
9. The log files were copied from the individual computers.

3.3.1 Pilot studies

Initially only one pilot study was intended but the electronic data was not written to the .ini files as planned and another pilot study was necessary. The same procedure was followed for both the pilot studies.

3.3.2 Main study

The main study was conducted at two different computer laboratories equipped with seventy computers each. Two sessions were held at each laboratory.

The computers were colour coded to indicate which version was loaded on the computer.

All the data was collected in four hours. The .ini files were retrieved from the computers via the LAN. All the .ini files were recovered and the cleaning of the data could begin.

4. Summary

This chapter provided a description of the quantitative research approach and experimental research design. The data collection plan included a description of the electronic intervention and data collecting instrument. Analyses of the log files containing the data followed.

The results and findings of the study are presented in chapter four.

Chapter 4: Analysis and findings

1. Introduction

The previous chapter described the research approach and design. The planning section included an explanation of the unit of analysis, the data, the instrument and a brief description of how the study was conducted.

This chapter presents the analysis of the data and discusses these findings. The findings determine the extent to which the research questions are answered. The existing literature was used to determine how the findings of this study complement or contradict findings in other studies.

The research was directed and guided by the following research questions:

- What is the cognitive load of the different presentation formats of a multimedia program, using the direct measurement technique?
- How does the position of a secondary task influence the measurement of cognitive load?

The hypothesis is that the cognitive load is influenced by the presentation format of content, but not by the position of the secondary task.

2. Clarifying the concepts

Concepts relating to statistical analysis should be unambiguous to enhance the understanding of the interpretation of the data. Table 4.1 describes the statistical terminology used in this study according to information from the electronic textbook, Basic Statistics (1994) available at http://www.statsoft.com/textbook/stbasic.html, unless otherwise stated.

Terms	Explanations
Statistics	Statistics involves collecting, summarising and analysing data that are subject to random variation as a means of finding order and meaning in apparent chaos (unknown author).
Descriptive statistics	Descriptive statistics describe patterns and general trends in a data set and are used to scrutinize one variable at a time.
Inferential statistics	Inferential statistics test hypotheses about differences or relationships in populations on the basis of measurements made on samples (Lane, 2003).
Frequency distributions	This is a way of displaying numbers in an organised manner to be able to easily answer direct questions concerning quantities. A frequency distribution is a table that displays how many times in a data set each response occurs. Relative frequency is the frequency converted to a percentage.
Mean procedure	The MEANS procedure provides data summarisation tools to compute descriptive statistics for variables across all observations and within groups of observations.
Standard deviation	The standard deviation measures the spread in individual data points to reflect the uncertainty of a single measurement.
Mean	The mean of a set of observed data is the sum of data divided by n, where n is the number of responses. The mean takes consideration of all the values and can be corrupted by extreme values.
Median	The median value is the value of the middle item of a distribution list and tells nothing about the other data in the list. If the number of values is n, the median value is the value in the n/2 position. 50% of observed data lie above the median value, and 50 % below.
Range	The range is the difference between the largest and smallest values in a data set.
SL Mean	Significance level of the mean.
Ordinal measurement	Subjects are ranked in order from greatest to least or best to worst. Again there is no precisely measurable difference between the ranks.
General linear model (GLM)	The GLM procedure uses the method of least squares to fit general linear models. The statistical methods available in PROC GLM (in SAS®) include regression analysis, analysis of variance, analysis of covariance, multivariate analysis of variance, and partial correlation (Research – AMNH, 2005).

Terms	Explanations
Probability error / p value	The p-value is the probability of obtaining a result at least as extreme as a given data point, assuming the data point was the result of chance alone. The fact that p-values are based on this assumption is crucial to their correct interpretation. P-values are used in hypothesis testing.
Hypothesis	The hypothesis is a prediction of the findings in any research study. A tentative assumption is made in order to draw out and test its logical or empirical consequences. The null hypothesis takes the position that there is no change or difference as a result of the independent variable. It is a statistical hypothesis to be tested and accepted or rejected in favour of an alternative hypothesis The alternative or research hypothesis states that there is a change or difference.
R-square value	In statistics, the coefficient of determination R^2 is the proportion of variability in a data set that is accounted for by a statistical model. R^2 is a statistic that will give some information about the goodness of fit of a model. In regression, the R^2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R^2 of 1.0 indicates that the regression line perfectly fits the data.

Table 4.1: Statistical concepts

3. Analysis of the data

Descriptive and inferential statistics were used to analyse the data in this study. The SAS®[1] system release 8.2 was used for all analysis.

The data was reduced to descriptive summaries, which include frequencies, means, standard deviations or correlation. The following statistical procedures were used to summarise and interpret the results of this study:

- frequency distributions,
- mean procedures and
- the General Linear Model.

[1] SAS version 8.2 running on VN/CSM of the University of Pretoria's mainframe

The level of significance for all reported analyses was set to $\alpha = 0.05$. The results obtained in the study are presented in an appropriate format and discussed at that point.

3.1 Descriptive statistics

In this study I used descriptive statistics to summarise the demographic data collected in order to describe the profile of the sample population. Analysis included determining frequencies and calculating means and standard deviations. The PROC MEANS procedure was used to calculate descriptive statistics, estimate quantiles, which included the median, and calculate confidence limits for the mean.

The variables included in this analysis are age, gender, culture and language of the respondents. The study year of the respondent and prior knowledge of the content is also described in this section.

3.2 Inferential statistics

Inferential statistics are used to reach conclusions that extend beyond the immediate data alone. They are especially useful in experimental research. Inferential statistics serve as an indication of whether differences or relationships are real or just a chance fluctuation (Lane, 2003).

A simple example of the use of inferential statistics is to compare the average performance of two groups on a single measure to see if there is a difference. In this study the cognitive load of different formats of the multimedia programs were compared.

The means procedure and general linear model procedure were the statistical procedures used for this analysis.

The general linear model (GLM) was used as an extension of linear multiple regression for a single dependent variable and calculates the analysis of variance in this study.

3.3 Null Hypothesis

The analyses described in this section were conducted for the purpose of establishing whether or not it was possible to reject the following two null hypotheses:

- The position of the secondary task has no influence on the cognitive load.
- The presentation format has no influence on the cognitive load of the program.

Based on the literature as summarised in Chapter 2 the alternative hypotheses would be that presentation format does influence cognitive load and that the position of the secondary task on the screen influences the measured cognitive load.

The probability error (p-value) of the statistical analysis is used to decide whether or not the null hypothesis should be accepted or rejected. Since every score has some level of error researchers must decide how much error they are willing to accept prior to performing their research. This acceptable error is then compared with the probability of error and if it is less, the study is said to be significant. The value statisticians use to be able to say that they can reject the null hypothesis with only a 5% error as an acceptance error in the rejection was applied in this study ($p\text{-value} < 0.05$).

It is expected in this study that the cognitive load imposed by animations is significantly higher than the cognitive load imposed by graphics and text.

4. Demographic data

The demographic data describes the population.

There were 262 participants enrolled for the course (Homeostasis, Block 3) but only 245 arrived at the computer laboratories for the study. The data for three of the participants were excluded during the session as they had problems with the program. It appeared as if the data was not being written out. They were allowed to continue and complete the lesson as this was part of their normal class. Their log files were not retrieved. We finally

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

retrieved log files for 242 participants, but data for four participants were excluded from the analysis because the data was incomplete. They did not complete the demographic data, and/or the pre-and post-test and had very few responses to the secondary task. They were excluded from the sample as it seems that they did not want to participate or did not understand the secondary task. The 238 participant sample included 193 medical students and 45 dental students.

Two formats of the program were presented, namely version 1, where content was presented using predominantly animations, and version 2, where content was presented using predominantly static graphics and text. The distribution between the versions was almost equal: 120 participants did version 1 and 118 did version 2.

Table 4.2 displays the frequency distribution of the age and gender of the participants. The mean age of the participants was 18.3 years.

Variable	Description	Frequency	Percentage
Age (V3)	17<=V3<=18	24	10.08
	18<V3<=20	159	66.81
	V3 > 20	55	23.11
Gender (V4)	Male	77	32.35
	Female	161	67.65

Table 4.2: Age and gender distribution

The sample was randomly selected from the unit of analysis without taking gender in to consideration. The frequency shows that there was a skewed distribution between genders. More females (163) than males (79) participated in this study. The reason for this might be that the medical field was previously male dominant but with the liberation of women currently in South Africa the field is predominantly occupied by female students.

The ethnic origin of the sample population was divided into 4 groups as displayed in Figure 4.1.

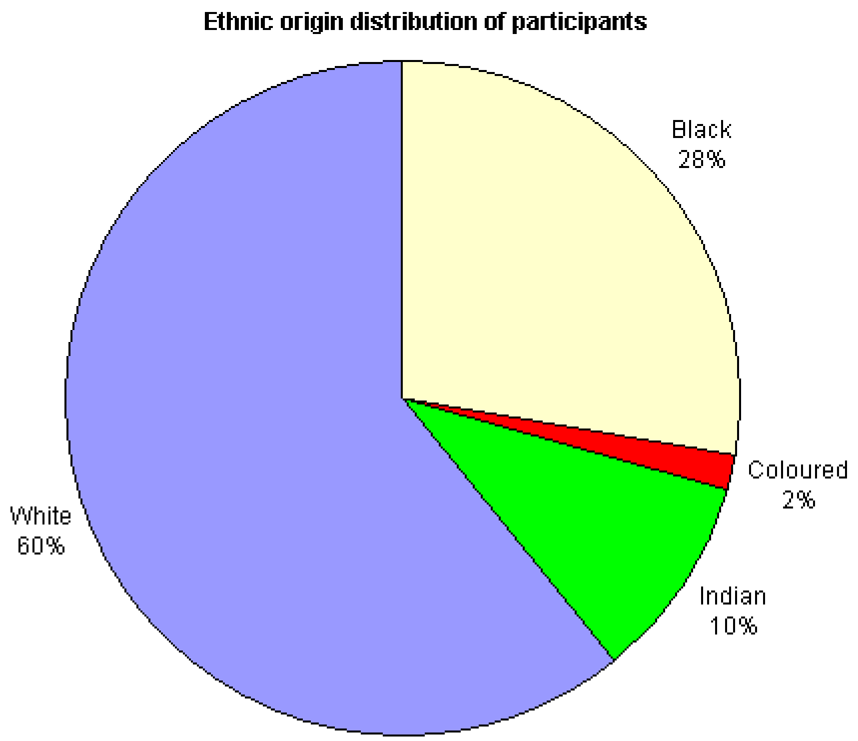


Figure 4.1. Ethnic origin distribution of participants

There are significantly more white participants than any other ethnic group. The reason for this skewed distribution is because the university was predominantly a white university and the other ethnic groups were previously politically excluded from the university. Even though the different ethnic groups now have equal opportunity, UP is a dual medium university and therefore attracts Afrikaans speaking students, who in South Africa are predominantly white.

The participants were asked to rate their knowledge of the subject before they used the program. These results are presented in Figure 4.2

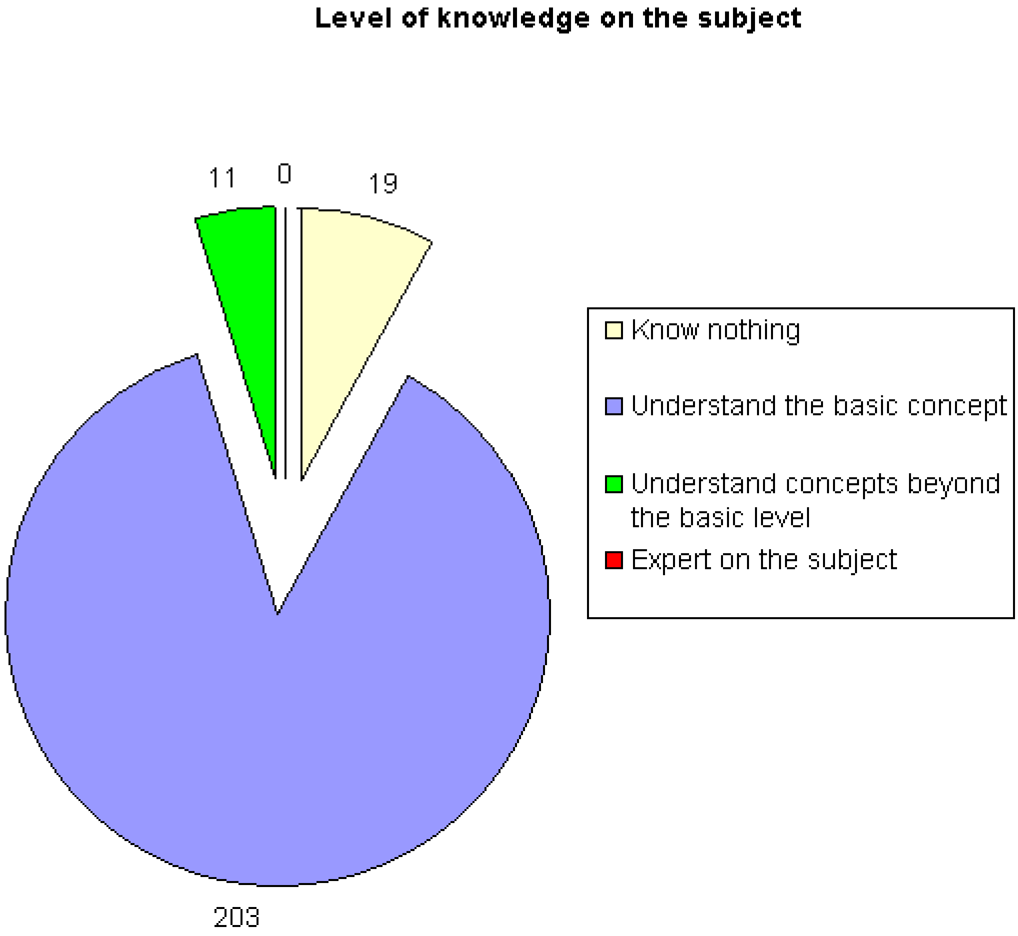


Figure 4.2. Level of prior knowledge of the subject

The highest frequency of participants understood the basic concept of the subject presented to them. Interestingly none of them rated themselves as experts. Few students (19) indicated that they know nothing. The possible reasons for this may be that they had read the study material beforehand, or had covered some of the work in anatomy or they might have had Biology as subject on secondary level at school. No students regarded themselves as experts on the subject.

5. Results of the experiment

In Chapter 2, I discussed the literature which describes the use of the secondary task method to measure cognitive load (e.g. Brünken and Leutner, 2000, Brünken et al., 2002, Brünken, Plass and Leutner, 2004). In this study the cognitive load was not measured on all the screens. Table 4.3 presents the screens where cognitive load was measured, the presentation format of the content and the number of occurrences of the secondary task.

Program	Version 1		Version 2	
	Screen	No of triggers	Screen	No of triggers
The content and format were equivalent on these screens	4	4	4	4
	5	3	5	3
	8	4	8	4
	10	3	10	3
	11	3	11	3
	14	4	21	4
	17	6	18	6
	19	8	23	8
The content was equivalent but the presentation format was different. Version 1 used one screen to present the content and Version 2 used 4 screens	12	8	13 14 15 16	4 3 3 5
The content was equivalent but the presentation format was different, only one screen was used to present the content	13	5	20	5

Table 4.3: Occurrences of the secondary task on screens where cognitive load was measured.

Individual means on the repeated measures of reaction times were calculated for each participant under each condition. These mean values were then used for further analysis.

The null hypothesis will be rejected if the p-value is smaller than or equal to the significance level where the probability error is set at $p < 0.05$.

5.1 Influence of the position of the secondary task on the cognitive load.

To answer the research question: "How does the position of the secondary task influence the measurement of cognitive load?" the secondary task was placed in the top or bottom right hand corner of the screens where cognitive load was measured. Figure 4.3 illustrates how the analysis was approached.

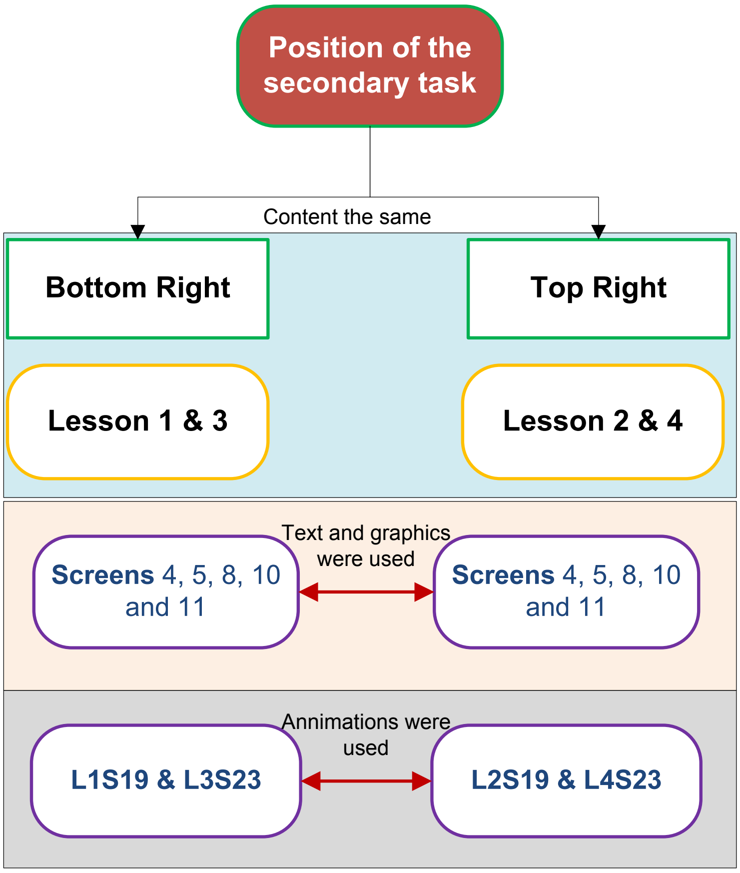


Figure 4.3. Analysis of data according to position of secondary task.

The comparison was not done at version level but at lesson level. Lessons were grouped together according the position of the secondary task. A comparison of each of the three different levels in the diagram was performed. I first looked at the position as it affects the combined lessons as a whole, and then compared screens at lesson level according to format as well.

5.1.1 Lesson-wise comparison of the mean cognitive load per position of the secondary task

The mean cognitive load for the lesson groups is displayed in Table 4.4. These results are obtained after determining the combined mean for lessons 1 and 3 and lessons 2 and 4 respectively. The lessons were grouped according to the position of the secondary task on the screen.

Lesson	Position of the secondary task	N	Mean (SD)	F value	P value	R-square
Lesson 1 & 3	Bottom right	1207	6.24609155 (3.4431348)	3.25	0.0716	0.001220
Lesson 2 & 4	Top right	1454	6.0058570 (3.4068841)			

Table 4.4: Lesson-wise comparison of the mean cognitive load per position of the secondary task

Table 4.4 indicates that the mean cognitive load is slightly higher on the screens where the secondary task was in the bottom right corner of the screen. A GLM procedure, with a two level class variable (the trigger position at bottom right or top right) was performed to test the significance of these results. The p- value is 0.0716, which is greater than 0.05 and indicates that the null hypothesis, which states that there is no relationship between cognitive load and position of the secondary task, cannot be rejected. It would therefore appear that the position of the secondary task on the screen does not influence the cognitive load, as measured using dual-task methodology.

The R square value, which provides an indication of how well the relationship which was tested accounts for the variation in the data, is only 0.00122. This means that only one tenth of a percent of the variation in the data is explained by the relationship tested. It must be assumed that other variables that were either not tested or included in the model may have contributed to the variation in the data.

5.1.2 Screen-wise comparison of the mean cognitive load per position of the secondary task

The position of the secondary task was then compared at screen level. The content and format (text and graphics) of screens, 4, 5, 8, 10 and 11 were exactly the same in all four the lessons, except for the position of the secondary task.

Table 4.5 shows the mean value of the cognitive load per group of screens grouped together according the position of the secondary task. The statistical analysis was extended and by-processing screen-wise analysis produced the results presented in Table 4.5.

Both versions	Position of the secondary task				Statistical values		
	Bottom right		Top right				
Screen	N	M (SD)	N	M (SD)	F	P	R^2
4,5,8,10 & 11	544	5.6868315 (3.4705484)	644	5.3486867 (3.3911115)	2.87	0.0905	0.002414
4	108	5.78410494 (3.5747377)	129	5.4684047 (3.276862)	0.50	0.4792	0.002132
5	109	6.03076307 (3.5280464	129	5.67892211 (3.1640590)	0.66	0.4183	0.002778
8	109	5.78640434 (3.1814855)	129	5.22316886 (3.3890256)	1.73	0.1903	0.007259
10	109	5.67776322 (3.4605558)	128	5.15164311 (3.4798034)	1.35	0.2460	0.005723
11	109	5.15601427 (3.5958250)	129	5.21976744 3.6500520)	0.02	0.8926	0.000077
19, 23	108	9.3024266 (0.9964639)	128	9.2421894 (1.1441084)	0.18	0.6696	0.000779

Table 4.5: Screen-wise comparison of the mean cognitive load per position of the secondary task.

The difference in the mean values of bottom right = 5.68 and top right = 5.34 indicates that the cognitive load is slightly higher on the screens where the secondary task is on the bottom right. The P value (0.0905) however is exceptionally high and there is only a marginally improvement in the R-square value (0.002414) and therefore we cannot reject the null hypothesis that position of the secondary task does not influence the cognitive load.

This trend can be seen in all the comparisons between screens. On screen 11 the difference in cognitive load is very small and the p value exceptionally high. The content on screen 11 was presented using a complex static graphic and static text to explain a number of concepts.

Note that the mean of the cognitive load on screen 19 & 23 is exceptional high (9.x and 9.c respectively, out of a possible 10). Both screens used animation to present the content. This finding will be elaborated on in paragraph 5.2.4.

5.1.3 Conclusion

Although there is an indication that the position of the secondary task might influence the measurement of cognitive load, the null hypothesis cannot be rejected with confidence and therefore the alternative hypothesis cannot be accepted with confidence either. The influence of the position of secondary task in this study is only a random effect.

5.2 Influence of the presentation format on the cognitive load.

To answer the research question: “What is the cognitive load of the different presentation formats of a multimedia program, using the direct measurement technique?” I compared the cognitive load of the program developed for version 1, which predominantly used animations, with the program developed for version 2, which predominantly used text and static graphics.

Figure 4.4 illustrates the analysis approach.

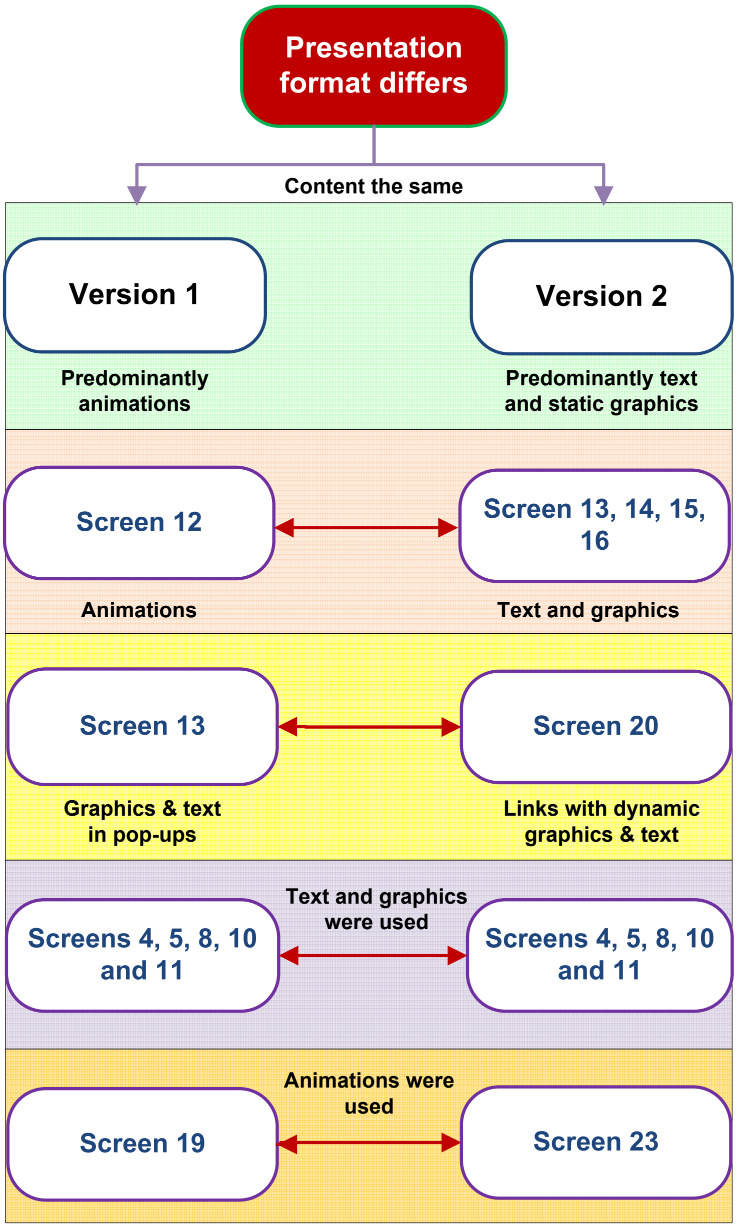


Figure 4.4. Analysis of data according to presentation format

The first comparison was between version 1 and 2, followed by further screen-wise comparisons.

5.2.1 Version 1 vs version 2

The mean cognitive load per version of the program is displayed in Table 4.6.

Version	Presentation format	N	Mean(SD)	F value	P Value	R-square
1	Predominantly animations	1198	6.64081144 (3.3347598)	52.39	<.0001	0.019322
2	Predominantly text and graphics	1463	5.68411239 (3.4385300)			

Table 4.6: Results for the comparison between versions where presentation format differs

The standard deviation gives an idea of the spread of the results and is derived from the means procedure. The GLM procedures use the least squares (LS) mean and standard error to describe the spread of the results. In this study the means and LS means were exactly the same in all instances. The mean cognitive load of 6.64 for version 1, where predominantly animations were used to present the content, was calculated from 1198 responses. The mean cognitive load of 5.68 for version 2, where predominantly static graphics and text were used to present the content, was calculated from 1463 responses.

The p value of less than 0.0001 is very enlightening and indicates a highly significant difference between the cognitive load means of the two versions. The statistical significant results indicate that in the real world the cognitive load is measurable and that presentation format does influence cognitive load.

The null hypothesis that presentation format does not influence cognitive load, can therefore be rejected and the alternative hypothesis, that presentation format does influence cognitive load, seems to be a valid prediction of the results.

The R-square is still very low at almost 2 percent. This is not problematic but motivates further analysis per screen where the presentation format was compared more rigorously.

5.2.2 Animation vs Text and Graphics.

The same content was presented using only animation on screen 12 in version 1 and using only static graphics and text in version 2.

The results of this analysis are displayed in Table 4.7.

Version	Screens	Presentation format	N	Mean(SD)	F value	P value	R square
1	12	Animation	360	7.13174619 (3.1719675)	28.32	<.0001	0.035792
2	13, 14, 15 & 16	Text and graphics	405	5.86694698 (3.3748061)			

Table 4.7: Version and screen

Again the p-value of less than 0.0001 indicates highly significant results and the improvement of the R-square value supports the conclusion that presentation format influenced the cognitive load. The mean cognitive load of 7.13 for screen 12 of version 1 where an animation of 1 min 45 seconds in duration was used to present the content is significantly higher than the mean cognitive load of 5.82 for screens 13 – 16 combined, of version 2, where static graphics and text and one very short animation were used to present the same content.

Figure 4.5 shows screen 12 of version 1.

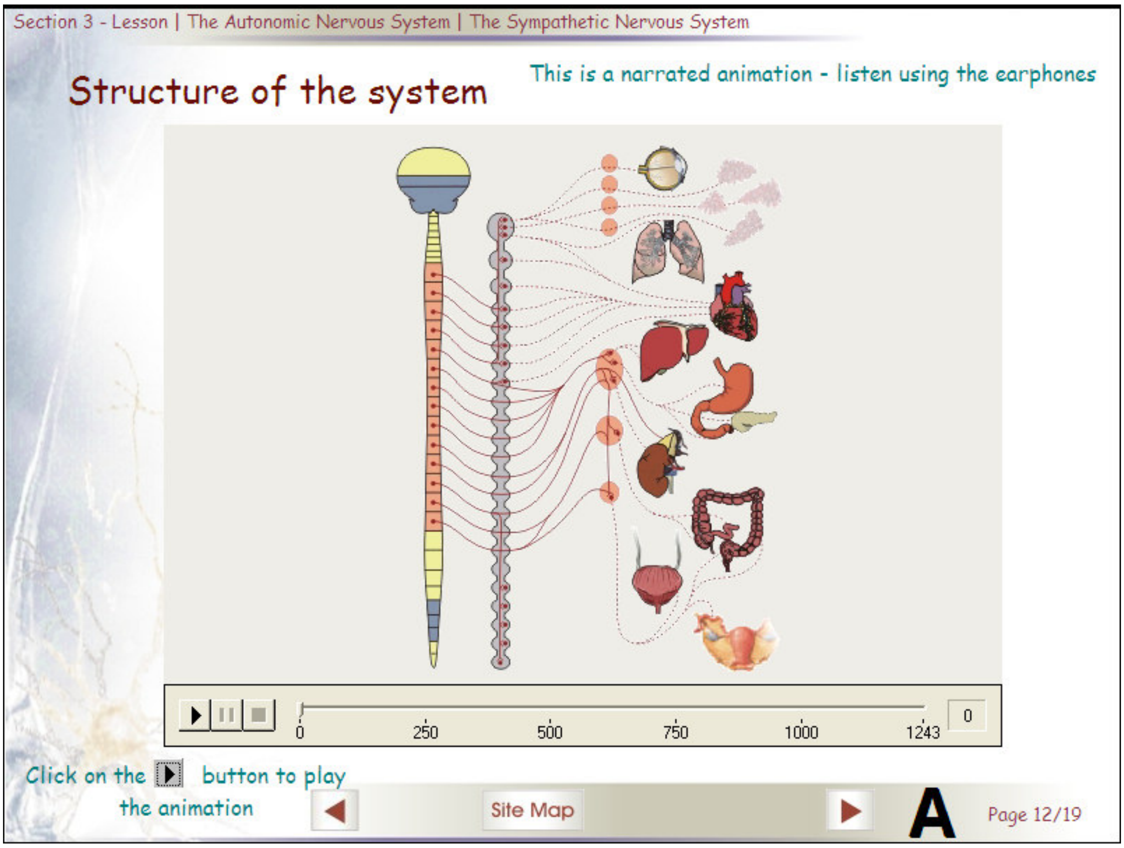


Figure 4.5. Screen 12 of version presenting content using an animation

For the primary task, the learner has to play the audiovisual animation to commence learning. The graphic on the screen changed according to the narrative explanation of the different areas. The learner has no control over the pace except to replay it. The secondary task was to hit the ENTER key if the symbol in the bottom right corner in lesson 1 or top right in lesson 2 changed colour. The auditory and visual channel in working memory is used to look and listen, leaving fewer cognitive resources available in working memory, which then in turn has a negative influence on schema building in the long-term memory.

Figure 4.6 shows screen 13 of version 2 as the first in a sequence of 7 screens where the same content was present ended using static images and pop-up text boxes.

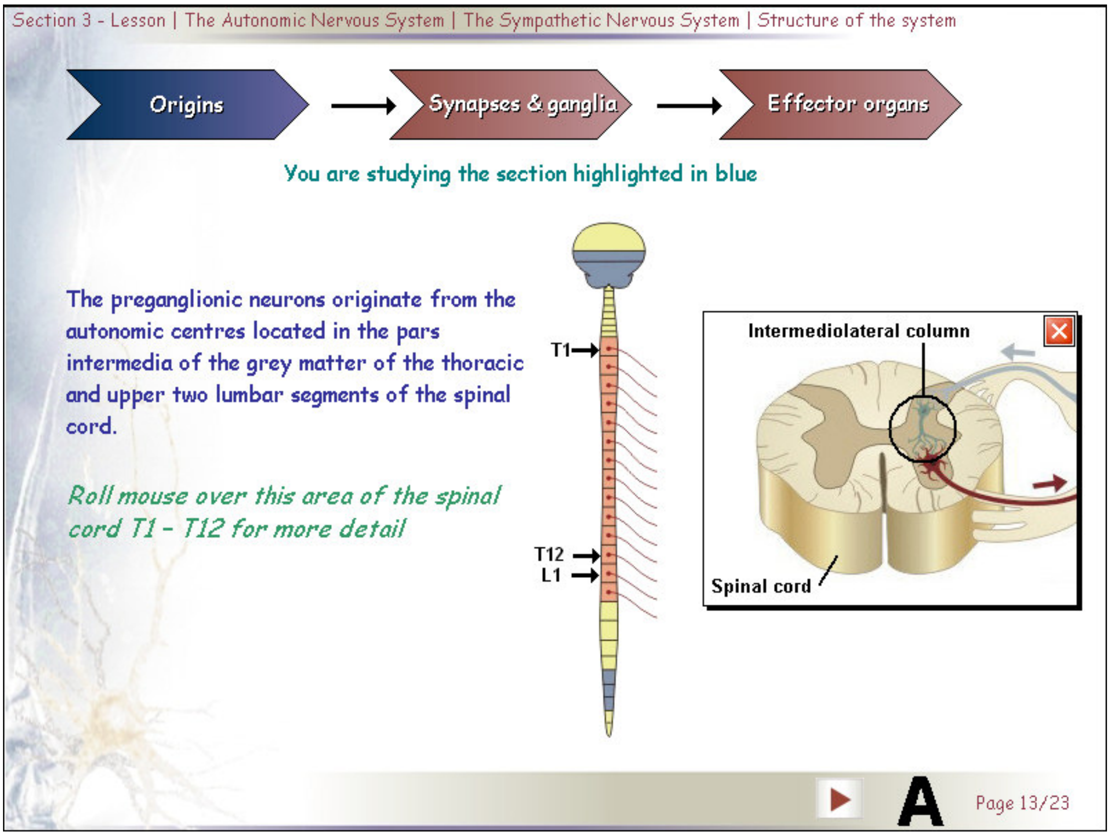


Figure 4.6. Screen 13 of version 2 where static images and text are used to present content

The learner had control over the pace of learning. Nothing happens unless the learner initiates it. The learner only used the visual channel in the working memory to eventually build schemas in long-term memory. There are now more cognitive resources available in working memory, compared to those available when the learner is using animation. The learner could subsequently pay better attention to both tasks and respond more rapidly to the secondary task.

5.2.3 Graphics & text in pop-ups vs links with dynamic text & graphics.

On screen 13 of version 1 graphics and text in pop-ups were used to present content. In version 2 on screen 20 the same content was presented using links with dynamic text and graphics.

The results of this screen-wise comparison are presented in Table 4.8.

Version	Screens	Presentation format	N	Mean (SD)	F value	P value	R-square
1	13	Graphics and text in pop-ups	120	6.51098094 (3.2445231)	6.55	0.0109	0.018225
2	20	Links with dynamic text & graphics	235	5.53164172 (3.4908239)			

Table 4.8: Comparison between screens using graphics & text in pop-ups vs screens using links with dynamic text & graphics

A p-value of 0.0109 indicates a significant difference between these two values. This is still highly acceptable even though the p value is not as considerable a p value of less than 0.0001. The R-square value remains in the vicinity of almost 2 percent. Figure 4.7 shows screen 13 of version 1 where the learner initiates learning. After clicking on either Divergence or Convergence the content is presented in a pop-up, shown in Figure 4.8.

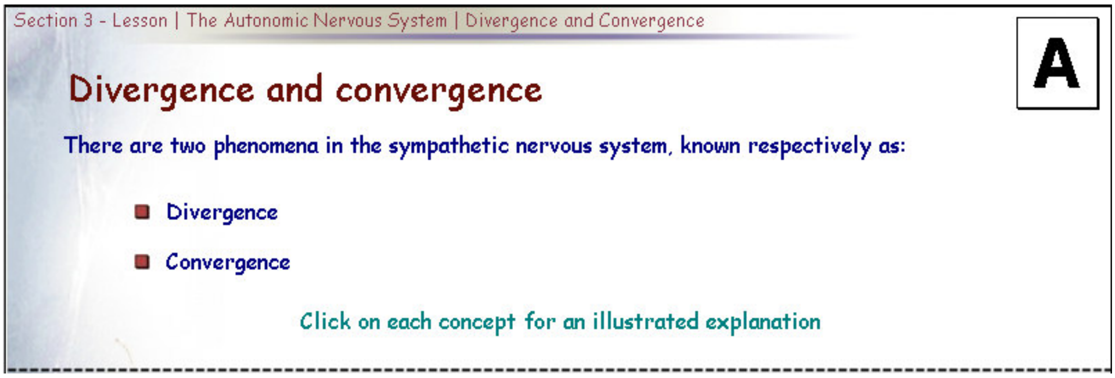


Figure 4.7. Screen 13 of version 1 content presented using pop-ups

The secondary task was either top right, lesson 2 or bottom right, lesson 1.

Figure 4.8 shows the pop-up covering the complete screen.

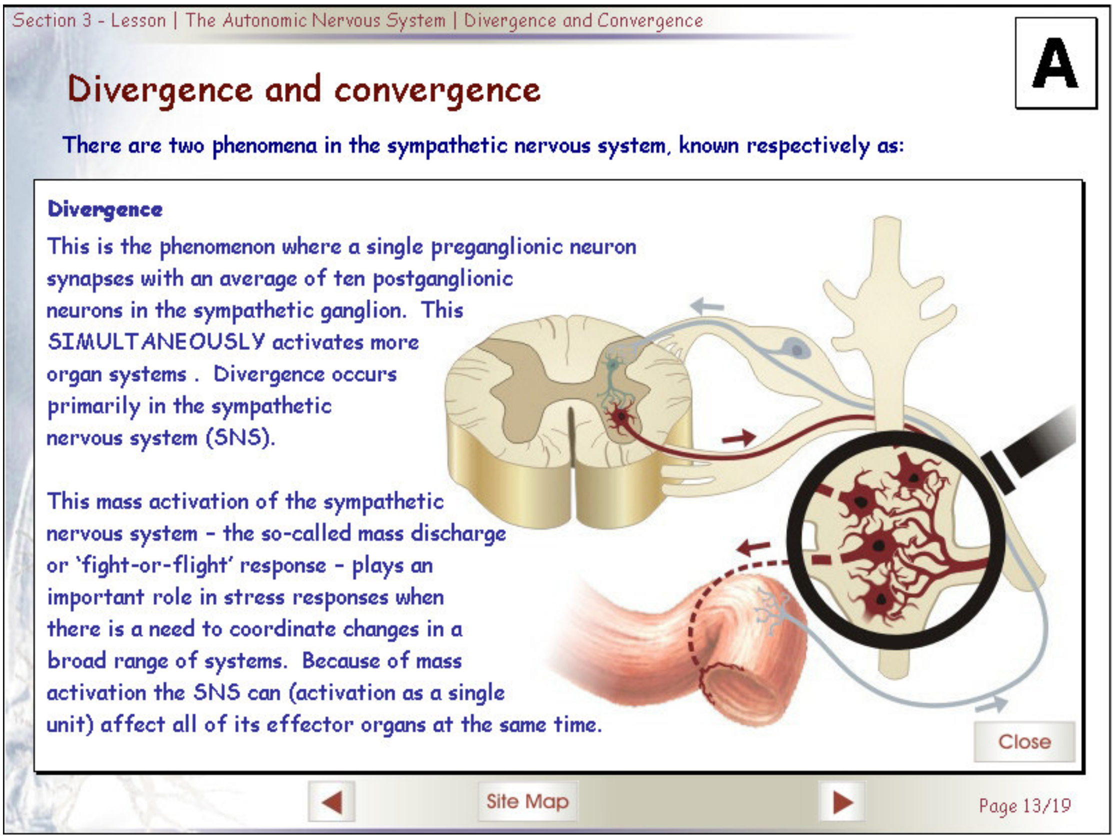


Figure 4.8. Pop-up with text and images on screen 13 of version 1

In version 2 the learner also had to initiate the learning as shown in Figure 4.9. The difference between version 1 and 2 was that now in version 2, instead of the content displaying in a pop-up, it displayed under the bulleted text. The user could toggle between the two concepts and the display updated dynamically.

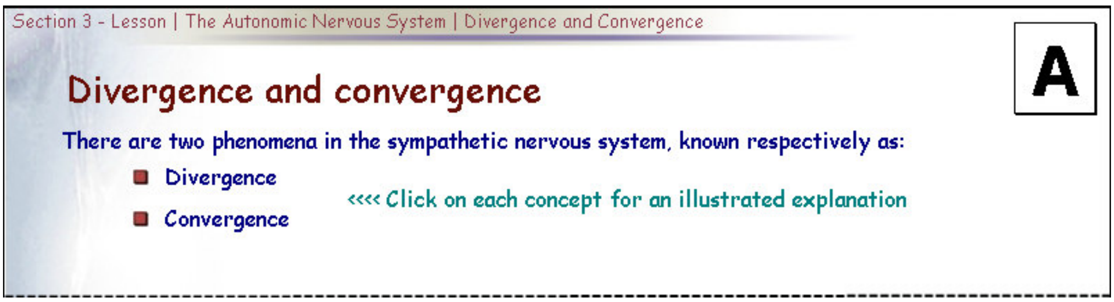


Figure 4.9. Screen 20 of version 2 using text and images

The content that was presented on the same screen under the bulleted list is displayed in Figure 4.10.

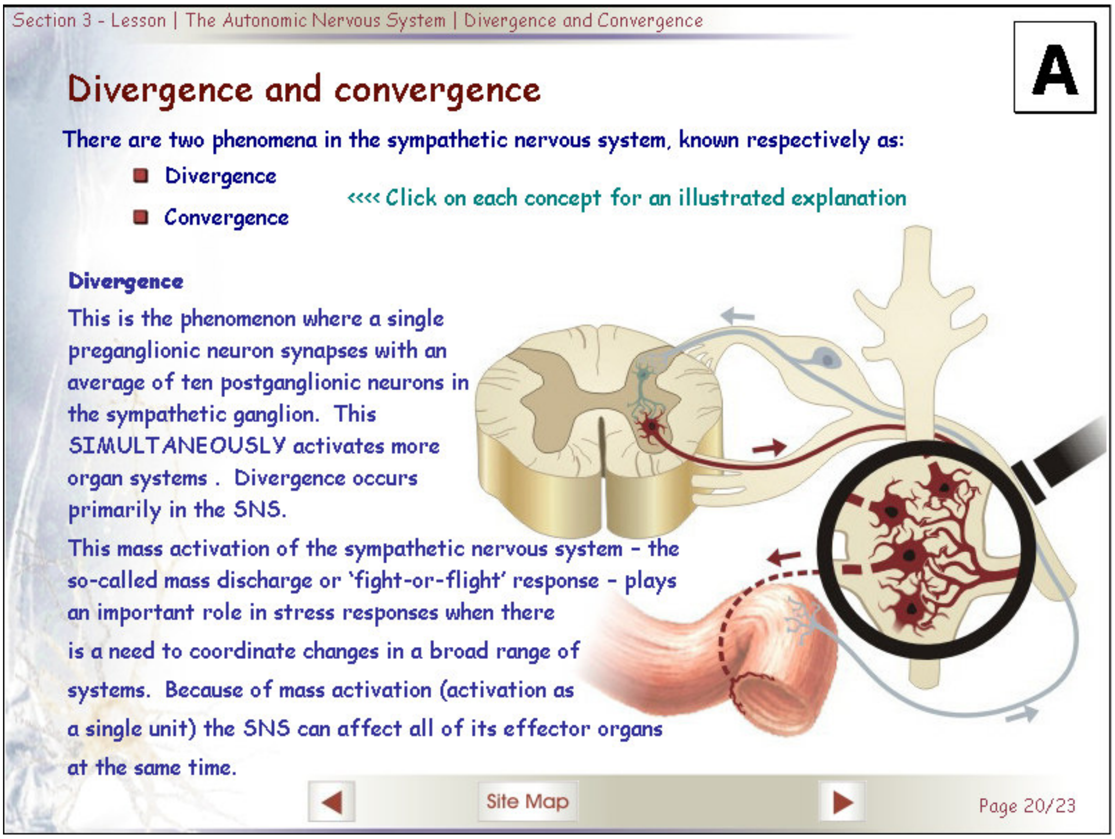


Figure 4.10. Text and images on screen 20 of version 2

The secondary task yet again was either top right, lesson 4 or bottom right lesson 3.

Once more the learner controlled the pace of learning and only the visual channel in the working memory is used to build schemas in long-term memory. However, the pop-ups covering the whole screen was a source of split attention and the learner needed to simultaneously attend to two sources of information (Cooper, 1998). More of the cognitive resources are used to keep the learner focussed on the content in context, whilst the text displayed on the same screen stayed in context and there are now more cognitive resources available in the working memory, compared to those available when the learner is using pop-ups. Again the learner can pay better attention to both tasks and respond more rapidly to the secondary task.

5.2.4 The same presentation format were used in both versions

A screen-wise comparison of content that was presented using exactly the same format, static graphics and text on screen 4, 5, 8, 10 & 11 for both version 1 & 2 and animations on screen 19 of version 1 and screen 23 for version 2 seemed unnecessary as no differences were expected. These results are displayed in Table 4.9, and contrary to what was expected, show a marginally significant difference for the screens using static text and graphics.

Version	Screens	Presentation format	N	Mean(SD)	F value	P Value	R-square
1	4, 5, 8, 10 &11	Static text and graphics	599	5.87384247 (3.4453768)	14.24	0.0002	0.01186
2	4, 5, 8, 10 &11	Static text and graphics	589	5.12692521 (3.3765773)			
1	19	Animation	119	9.3825071 (0.789328)	2.65	0.1050	0.01119
2	23	Animation	117	9.1550766 (1.300291)			

Table 4.9: Screens with the same presentation format in both versions

The p value of 0.0002 is significant at $\alpha = 0.05$ and this is an unexpected finding as the content and presentation format was exactly the same. Though the difference between 5.873 and 5.126 are not even one unit of cognitive load in the real world it is still unexpected. If screens 4, 5, 8, 10 & 11 were presented after screen 12 of version 1 and screens 13-17 of version 2 where we expected a difference in cognitive load, then we could draw the conclusion that the participants might have followed the trend. But as these screens were presented first the cognitive load should have been the same. If the screens were at the end of the program the reason for the results might have been fatigue. Various factors might have influenced the results like

- the time spent on the screens,
- age, gender and culture,

- prior knowledge of the subject and
- learning performance in the pre- and post assessment of the content.

The p-value of 0.1050 for the comparison between screen 19 and 23 is expected, and indicates that there is no significant difference in the cognitive load for the two screens.

The high cognitive load on the screens where animation was used to present content, screen 19 (9.38), screen 23 (9.155) and screen 12 (7.13) supports the previous findings of studies done by Brünken, et al (2002, 2003) that animation increases cognitive load, which could negatively influence learning.

The animation was 1.5 minutes long. It covered the content about neurotransmitters, which was new work for this group and therefore it can be expected that the intrinsic cognitive load is high as the number of elements that are to be integrated into a to-be-learned schema, is high (Gerjets, Scheiter & Catrambone, 2004). Participants had to listen and look using dual channels to build schemas in the long term memory. This instructional techniques that require learners to engage in activities that are not directed at schema acquisition, increased the extraneous cognitive load (Sweller, 1994) leaving less working memory available for schema acquisition. The animation could be stopped by the participant, but user interaction was not tracked at this level. We also did not track how many times the user actually looked at the animation.

5.2.5 Summary

In theory, with a sample of this size, one can say with 95 percent certainty that the results have a statistical precision of plus or minus 95 percent of what they would be if the entire participant population had been targeted. Unfortunately, there are several possible sources of error in all experiments that are probably more serious than theoretical calculations of sampling error. They include refusals to respond to the question or task, or weighting by demographic control data. It is difficult or impossible to quantify the errors that may result from these factors.

6. Conclusion

In this study the significance of the results as indicated by the p-values allow me to draw the following conclusions:

- The position of the secondary task when using the dual-task method to measure cognitive load has **no influence** on the cognitive load.
- The presentation format of the content **does influence** the cognitive load. The cognitive load on screens where animation was used, were a great deal higher than those screens that used static images and text.

Recommendations for further studies in this field as well as a summary of the study are discussed in chapter 5.

Chapter 5: Closure

1. Introduction

A literature review of empirical research in the field of instructional design and multimedia highlighted the fact that an instructional designer needs to pay attention to cognitive load. Cognitive load theory describes the causes and influences of design practice on cognitive load (Kirschner, 2002). The research in this field provides guidelines for developing content that minimises this cognitive load (Mayer & Moreno, 2003).

Measurement of cognitive load becomes an important issue because if it is not measured, the designer can only make assumptions about its influence. These assumptions might not necessarily be correct.

The opportunity to study the measurement of cognitive load came with the invitation to research the measurement of cognitive load as part of a wider study, which also compared the correlation between two different techniques for measuring cognitive load. The wider study measured the cognitive load of two programs where the same content was presented using different presentation formats.

This chapter presents a summary of the study, a reflection of the methodology and a discussion of what can be learned from the findings described in chapter 4. Finally, I recommend topics for further research and practice.

2. Summary of the study

This study investigated some factors that might influence the direct measurement of cognitive load. The research questions for this study were:

- What is the cognitive load of the different presentation formats of a multimedia program, using the direct measurement technique?

- How does the position of a secondary task influence the measurement of cognitive load?

The literature review in Chapter 2 discussed the concepts multimedia, cognitive load and measurement of cognitive load.

Over time, as technology is used more and more in the field of education the definition of multimedia changed from a narrow view of words and pictures to a broader view that includes audio and visuals (pictures) in both static and dynamic format.

Mayer conceptualised the theory of multimedia learning, which addresses processing of multimedia in memory. This theory is illustrated in Figure 5.1

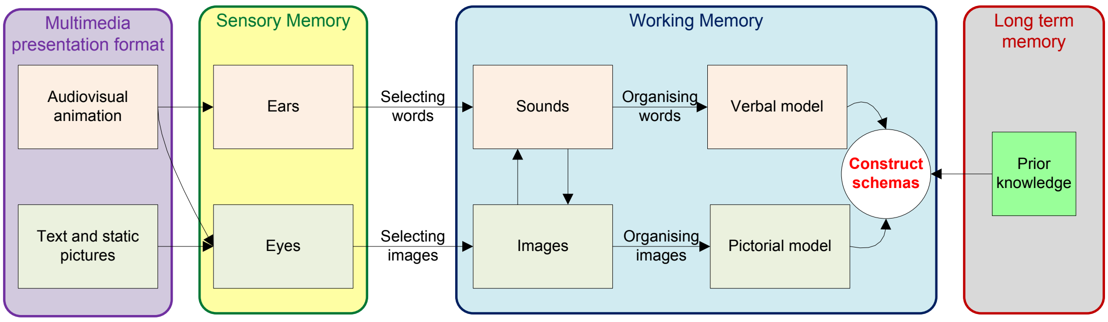


Figure 5.1. The theory of multimedia learning

The cognitive load theory of Sweller (1998) implies that good instructional design of multimedia programs focuses on two issues:

- decreasing extraneous cognitive load, which increase the amount of working memory that can be used in schema construction;
- increasing germane cognitive load.

Brünken, Plass & Leutner (2003) categorise cognitive load measurement on two dimensions namely objectivity and casual relations. Objectivity refers to whether the data collecting method is objective or subjective. Direct or indirect measurement refers to the casual relation between the method and the participants. The dual task method is an objective direct measurement method and was used in this study. Subjective measurement, using rating scales to indicate mental effort, was used in the wider study.

Brünken, et al (2003) showed how cognitive load can be quantitatively measured in an "authentic" learning environment. They define free cognitive resources as the difference between total cognitive load and zero cognitive load. They used the dual task approach which uses reaction time to a secondary task. Their work provided the direction for the protocol set up to measure cognitive load in this study.

A quantitative research approach, using an experimental research design was used as described in Chapter 3. Initially only one pilot study was planned in order to test the design of the program and the data collection process, but because the data files were not created, another pilot study was necessary. The program was adjusted to investigate the influence of the position of the secondary task on the measurement of cognitive load. Two hundred and forty two second year medical and dental students participated in the main study. The data for 4 participants had to be dropped from the study as the data files were empty.

Four different multimedia programs were used to measure the cognitive load and test the influence of the position of the secondary task on the measurement of cognitive load and the presentation format. The differences in the programs used in this experiment are illustrated in Figure 5.2.

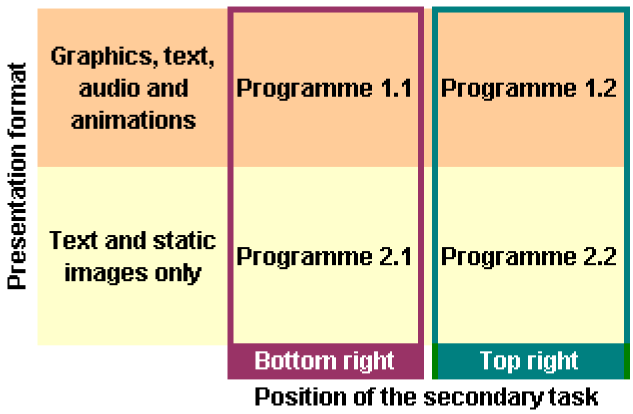


Figure 5.2. Differences in the programs

The following data was collected electronically and recorded in an external log file:

- Time the lesson was accessed.
- Time the lesson was exited.
- Time a screen was entered.
- Time the learner left a screen.
- The trigger time of the secondary task.
- The response time of the learner to the secondary task

The data was written to an .ini file, together with demographic data, style analysis results, pre- and post-test results and the subjective measurement of the cognitive load according to a 9-point rating scale. All of the latter were used in the wider study. The cleaning up of the data was an enormous task and an investigation should be done to reveal an easier way to record and process the data.

The pilot studies and main study were conducted in exactly the same way. The logistical problems to accommodate such a huge sample were handled well and the eventual loss of only 4 data files can be seen as a major achievement for this study.

3. Methodological reflection

One of the main characteristics of quantitative experimental design is random sampling. Although this was an empirical study in an authentic environment it was still possible to randomly allocate participants to the 4 different groups of the program.

Another important characteristic of experimental designs is the controlled environment of the experiment. In an authentic environment it was however not possible to control the complete presentation format of the 2 versions of the program, as both contained animation and static images. A comparison between the different formats was therefore only possible on screen level and not per version. The research program was part of the students' study programme and was conducted during their education time and therefore each student had to have access to the complete content.

Another question that arises is whether random timing of triggers would not have been better rather than triggers at set intervals. Brünken and Plass (2003), used random intervals in their studies. Researching the influence of the timing of triggers was not part of this study, but may provide valuable insight if tested.

The study produced some logistical problems like working with large groups and within a specific time frame. It is difficult to control 60 students in a computer laboratory and there were two sessions in two different labs within the same time frame. These challenges had to be overcome to maintain the authentic environment for the study.

3.1 Substantive reflection

There are several publications that discuss the dual-task methodology (Brünken et al, 2003, Chandler & Sweller, 1996, Marcus, Cooper & Sweller, 1996), but little empirical work has been done using this technique. No study was found that took the position of the secondary task into consideration.

After the first pilot study the colour of the symbol that was used as secondary task was changed from light purple to brilliant green as it seemed that the change in colour from black to purple might not be patently visible. No other colours were used. The decision to test the influence of the position of the secondary task on cognitive load was taken after the first pilot study. The secondary task was only positioned at the bottom right and top right. Other positions of the secondary task were not tested, for example bottom and top left, bottom and top centred or left and right centred.

In spite of the methodological considerations a comparison of ‘only apples with apples’ was possible. We were able to select and do screen wise comparisons that isolated the presentation formats. As the purpose for this study within the context of the wider study was to determine and investigate the cognitive load for the different presentation formats, there were enough screens where the presentation format differed enough to provide clear results. Therefore the results obtained for this study is what was expected and was not in contradiction with what was found in worked done by Brünken, et al (2003), where the dual task method was also used to measure cognitive load in an objective direct way.

Like Brünken and Plass (2003), I found that it is possible to measure cognitive load using the dual-task method, but it is not an easy technique – it can be viewed as an intrusive element in a multi media program. However I did not explore this and will recommend it for further research.

Strehler (2007) measured cognitive load using the same materials and sample via the subjective technique and then investigated whether there was a correlation between these two measurements. She found no correlation between the measures obtained using two different techniques. This is interesting and should open a debate on what method of measuring cognitive load is the most reliable. In this field Paas, van Merrienboer and Adam (1994) launched a subjective indirect measure of cognitive load that used a questionnaire where learners are asked to report on the mental effort used to complete a task.

3.2 Scientific reflection

It is necessary to consider the extent to which the research methodology influenced the results. The dual task method - using a secondary task symbol that changed colour to measure cognitive load, rather than playing a sound or animation as secondary task was used. It is important to use only a visual or audio trigger that the learner has to respond to in order to keep as much working memory free as possible. Since only a visual trigger was used I am confident that the method did not influence the results. The learning style of the learner – auditory of visual - might have influenced the results, and the findings of Strehler (2007) will provide an indication of whether cognitive style influences cognitive load.

I could not find any literature in which research on cognitive load in tertiary health science or medical education was described. This field uses lots of visuals and multimedia programs should therefore pay more attention to the presentation format of the content.

The program was designed for an authentic environment and not specifically for a controlled experimental environment. Therefore both versions of the program used animations, static images and text to present the content. This however had no influence on the results as a screen-wise comparison was possible where the same content was presented using animation compared to using text and static images. Furthermore the duplication of the programs to accommodate bottom right or top right for the position of the secondary task made comparison between all formats possible.

Though the findings on the influence of the position of the secondary task on the measurement of the cognitive load were not statistically significant it revealed some interesting facts. In all instances the cognitive load on the screens where the position was bottom right the cognitive load was higher than on the screens where the position was top right. I found that odd, as most navigation buttons are positioned bottom right.

4. Recommendations

Further studies in the field where presentation format of multimedia programs are investigated, should include a study where the screens are manipulated to have a more precise and defined difference in the presentation format in a controlled research environment. Further research should include investigations into:

- the time spent on a lesson
- learning performance in terms of pre- and post test results.

This research provided an enormous field for new thoughts for educational researchers and instructional designers.

The study did not address the fact that the format of the symbol might have an influence on the results either. A capital letter A was used as a symbol. There is a possibility that if another symbol like a shape or drawing was used the results could have been different.

In measuring cognitive load using the dual task method and recording reaction time to the secondary task, further studies should include a comparison where

- the secondary task is at random intervals opposed to set intervals.
- the position of the secondary task include top and bottom, left centred and right.
- the symbol is a shape opposed to a character or picture
- the screens where cognitive load is measured is not known upfront, implicating that the symbol is not visible on entrance to the screen but appears at random intervals on set screens.
- the trigger differs, for example a sound that is played as opposed to a symbol changing colour.

From this study we draw the conclusion that presentation format does influence the cognitive load but the position of the secondary task does not influence the measurement of cognitive load significantly. These findings present lessons learnt when using the dual task approach to measure cognitive load and recommendations for further studies to underpin the reliability of measuring cognitive load directly.

Reference list

Antonenko, P. (2005). Considering the Learner in Educational Hypermedia Design: Cognitive Load or Cognitive Flexibility? *Iowa State University*. Retrieved May 25, 2007, from http://www.hci.iastate.edu/TRS/REPORTS/ISU-HCI-2005-02.pdf.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A Miyake & P Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge, England: Cambridge University Press.

Baddeley, A.D. (1992) Working memory. *Science*, 255, 556-559.

Brünken, R., & Leutner, D. (2001). Direct Measurement of cognitive load in Multimedia Learning. *Educational Psychologist,* 16(2), 45-52.

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, 49, 109–119.

Brünken, R., Plass, J.L. & Leutner, D. (2003). Direct Measurement of cognitive load in Multimedia Learning. *Educational Psychologist,* 38(1), 53-61.

Brünken, R., Plass, J.L. & Leutner, D. (2004). Assessment of cognitive load in Multimedia Learning with Dual task Methodology: Auditory Load and Modality Effects: *Instructional Science* 32, 115–132.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction,* 8, 293-332.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Pshycology,* 10,151-170.

Chi, M., Glaser, R. and Rees, E. (1982). Expertise in problem solving. In R. Sternberg, (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, N.J.: Erlbaum, (pp. 7-75).

Chipperfield, B. (2004). Cognitive load Theory and Instructional Design, *University of Saskatchewan*. Retrieved May 25, 2007, from http://www.usask.ca/education/coursework/802papers/chipperfield/index.htm .

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 53(2), 445- 459.

Cooper, G. (1998). Research into cognitive load Theory and Instructional Design at UNSW . *School of Education Studies, The University of New South Wales, Sydney, NSW 2052, Australia.* Retrieved July 18, 2007, from http://education.arts.unsw.ed.au/CLT.NET.html.

Cooper, G. (1990). Cognitive load Theory as an aid for instructional design. *Australian Journal of Educational Technology*, 6(2), 108-113. Retrieved July 18, 2007, from http://www.ascilite.org.au/ajet/ajet6/cooper.html.

Gerjets, P., Scheiter, K. & Catracombone, R. (2004). Designing Instructional Examples to Reduce Intrinsic Cognitive Load: Molar versus Modular Presentation of Solution Procedures. *Instructional Science*, 32, 33-58.

Gick, M. L. and Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.

Kirschner, P. (2002). Cognitive load Theory: implications of cognitive load Theory on the design of learning, *Learning and Instruction*, 12, 1-10.

Lane, D. (2003). Introduction to Sampling Distributions, *Connexions*. Retrieved July 13, 2007, from http://cnx.org/content/m11130/latest/.

Lansman, M., & Hunt, E. (1982). Individual differences in secondary task performance. *Memory & Cognition,* 10, 10–24.

Larkin, H., McDermott, J.,Simon, D., & Simon, H. (1980). Models of competence in solving physics problems. *Cognitive Science*, 11, 65-99.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, 88, 49-63.

Mayer, R.E. & Moreno, R. (2003). Nine Ways to Reduce cognitive load in Multimedia Learning. *Educational Psychologist*, 38(1), 43 – 52.

Mayer, R.E. (2001). Multimedia Learning, *Cambridge University Press,* New York, NY, USA.

McMahon, R. (1994) Trial and error: an experiment in practice. In J. Buckeldee, R. McMahon, (eds) *The Research Experience in Nursing.* London*:* Chapman & Hall.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review,* 63, 81–97.

Moreno, R. & Mayer, R. E. (2000). Meaningful design for meaningful learning: Applying cognitive theory to multimedia explanations. *ED-MEDIA 2000 Proceedings* (pp. 747-752) Charlottesville, VA: AACE Press.

Moreno, R. & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, 94, 156–163.

Neumann, A., Badke-Schauba, P. & Lauche, K. 2004. Measuring the development of shared mental models in product design teams. *Faculty of Instructional Design, Delft University of Technology,* Delft, The Netherlands.

Oliver, A, (1998). What are the main elements of a Multimedia program? Retrieved August 9, 2007, from http://www.herts.ac.uk/ltdu/technology/what_is_multimedia.html.

Paas, F., Tuovinen, J.E., Tabbers, H. & Van Gerven, P. W. M.,(2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.

Paas, F., Van Merriënboer J.J.G., & Adam, J.J. (1994). Measurement of cognitive-load in instructional research. *Perceptual and Motor Skills*, 79, 419-430.

Paas, F., Van Merrienboer, J. (1994) Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351-371.

Paivio, A. (1986). Mental representations: A dual coding approach. *New York: Oxford University Press.*

Schultheis, H & Jameson, A. (2004). Assessing cognitive load in Adaptive Hypermedia Systems: Physiological and Behavioral Methods. *Proceedings of Third International Conference, AH 2004, Eindhoven, The Netherlands*: 225-234. Retrieved June 17, 2007, from http://springerlink.metapress.com/content/r21696v81c2a/.

Strehler, A. (2008). The relationship between cognitive load and cognitive style in multimedia learning. *Unpublished Doctoral dissertation.* University of Pretoria, South Africa.

Sweller, J. (1999). Instructional design in technical areas. *Camberwell,* Australia: ACER Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning, *Cognitive Science*, 12, 257-285.

Sweller, J. (1994) Cognitive load Theory, learning difficulty, and instructional design, *Learning and Instruction* 4(4): 295 – 312.

Sweller, J. & Cooper, G. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2(1), 59-89.

Sweller, J.& van Merriënboer, J.J.G. (2005). Cognitive load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, 17(2), 148 – 177.

Sweller, J., Chandler, P., Tierney, P. & Cooper, M. (1990). Cognitive load as a factor in the structure of technical material. *Journal of Experimental Psychology: General*, 119, 176-192.

Sweller, J., 2003, 'Evolution of human cognitive architecture', *The Psychology of Learning and Motivation: Advances in Research and Theory*, 43, 12-30.

Thompson, C. (1999). If you could just provide me with a sample: examining sampling in qualitative and quantitative research papers. *Evidence-Based Nursing*, 2(3), 68–70.

Trochim, W.M.K, (2005). Design. Retrieved December 12, 2006, from http://www.socialresearchmethods.net/kb/design.htm).

Van Gerven, P.W.M, Paas, F, Van Merrienboer, J. J. G, Hendriks, M & Schmidt, H. G. (2003). The efficiency of multimedia learning into old age. *British Journal of Educational Psychology,* 73, 489–505.

Van Gerven, P.W.M, Paas, F, Van Merrienboer, J. J. G, & Schmidt, H. G. (2000). Cognitive load theory and the acquisition of complex cognitive skills in the elderly: Towards an integrative framework. *Educational Gerontology*, 26, 503-521.

Van Merriënboer, J.J.G., Kirschner, P.A. & Kester, L. (2003) Taking the load of a learner's mind: Instructional design for complex learning, *Educational Psychologist*, 38(1), 5–13.

Verwey, W. B.,&Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. J*ournal of Experimental Psychology: Applied*, 2, 270–285.

Wickens, C. D. (1984). Processing resources in attention. *In Parasuraman, R. &. Davies, D. R (Eds.), Varieties of attention* (pp. 63–102). London: Academic.

Appendix A: Letter of information

A study of the cognitive load of multimedia learning material

This study is being conducted by Riekie Smith under the supervision of Professor Johannes Cronje of the Faculty of Education, University of Pretoria, and Mrs Anne Strehler, PhD candidate.

The research is done in fulfilment of the requirements for RGO850 as credit towards the Masters degree in Education [CAE].

Dear Participant

You are invited to participate in a research study. We hope to be able to measure the cognitive load of electronic learning material. This research forms part of a wider study being conducted to investigate the relationship between multimedia learning, cognitive load and cognitive styles.

The subject for this study is Physiology. As a participant in this study, you will be asked to work through a multimedia learning program on the Autonomic Nervous System. We will use two different multimedia programs developed for the wider study . The two programs will present the same content but in different formats. Program 1 will make use of animations to explain a concept while program 2 will explain the same concept using text and graphics only. The measurement instrument is built into the program. You will be randomly assigned to one of the programs.

As a participant in this study, you will be asked to take part in at least one session of 60 – 90 minutes in length. Depending on the results of the study, another session of the same length might be necessary. You may decide at that time whether you wish to participate in that part of the study.

All information collected from participants in this study will be aggregated. Your name or student number will not appear in any report, publication or presentation resulting from this study. This will ensure your anonymity.

If you have any questions about participation in this study, please feel free to ask the researchers. If you have additional questions later, please contact Riekie Smith at (012) 368-8464 or by email at riekie.smith@epiuse.com.

In appreciation for the time you have given to this study, you will receive a copy of both formats of the learning material by the end of February 2006.

Please complete the attached letter of consent before you commence.

Regards

Riekie Smith

Researcher:

Appendix B: Letter of consent

A study of the cognitive load of multimedia learning material

I certify that I understand the nature of the study. The program will take approximately 60 - 90 minutes to complete. I willingly consent to participate and I will answer the questions to the best of my knowledge and react appropriately to interventions. The information obtained will be kept in confidence and will be used for statistical purposes only.

Signed at Pretoria on this 30th day of January 2006:

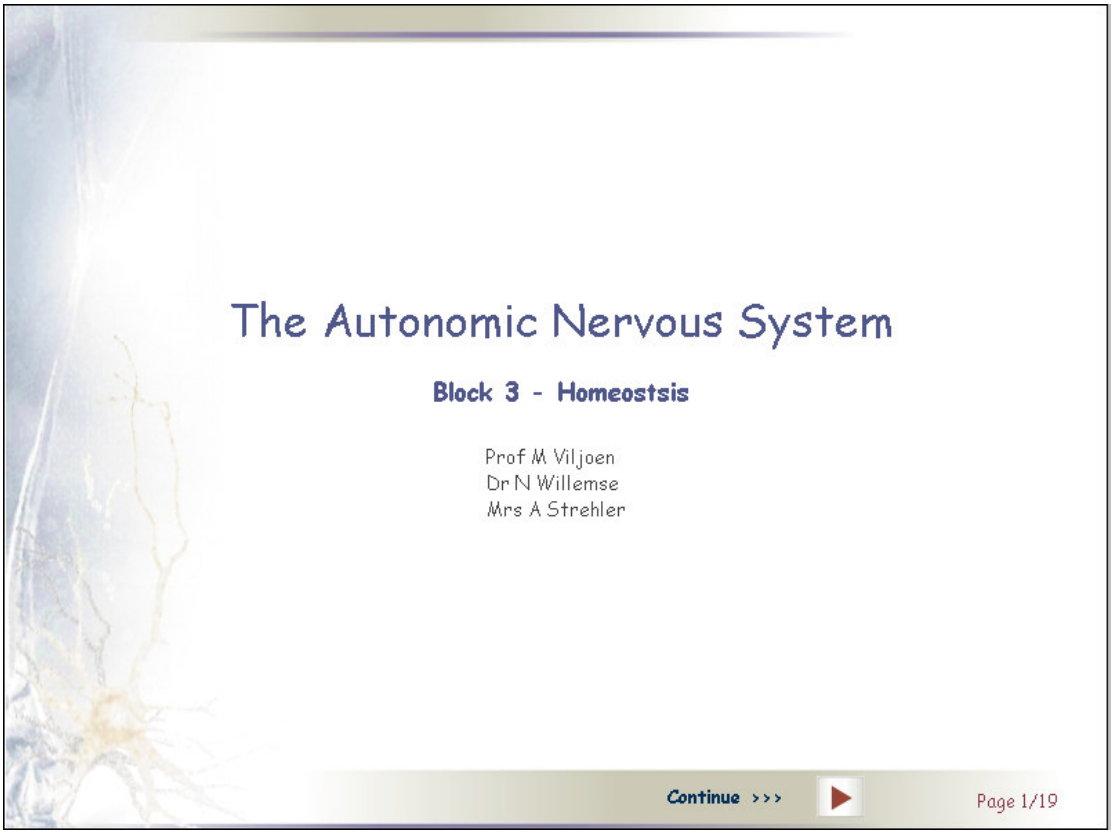
Student number of participant: ______________________________

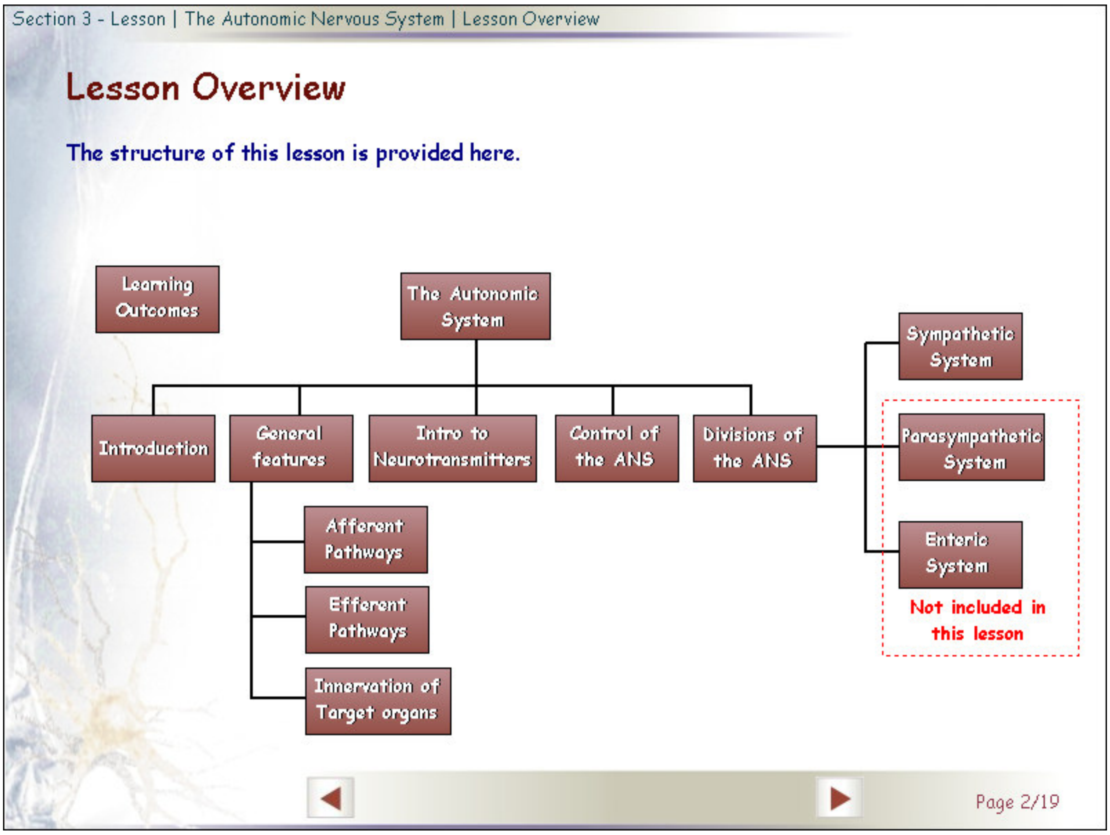
Signature of participant: ______________________________

If you would like a copy of the abstract, please provide your e-mail address:

____

Appendix C: Version 1





Section 3 - Lesson | The Autonomic Nervous System | Learning Outcomes

Learning outcomes

The learning outcomes for this lesson are to:

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

Click on the ▶ button to continue

◀ Site Map ▶ Page 3/19

Section 3 - Lesson | The Autonomic Nervous System | Introduction

Introduction

The Autonomic Nervous System (ANS) is involved in maintaining homeostasis of the internal environment of the body through the regulation and modification of the following activities.

- All involuntary muscle tissue.
- The secretory activity of all exocrine glands.
- Some endocrine glands.
- Some adipose tissue.

The **TWO** main functions of the system are to:

- Regulate the activity of the visceral organs and glandular structure responsible for the basic (vegetative) bodily processes - during rest and digest.
- Respond to a stressor - also known as the typical 'fight or flight' response.

Click on the underlined text

◀ Site Map ▶ A Page 4/19

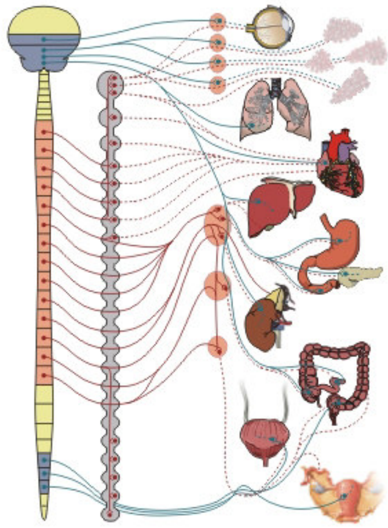
The ANS has three parts, the

- The sympathetic nervous system.
- The parasympathetic nervous system.
- The enteric system.

Click on the underlined text and look at the image

Many of the systems regulated by the ANS receive both excitatory and inhibitory signals from sympathetic and parasympathetic divisions. This provides fine control of the regulated system.

We will look at each of these in turn.



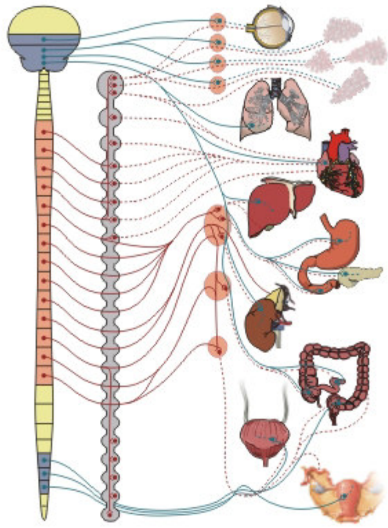
The Autonomic Nervous System

Section 3 - Lesson | The Autonomic Nervous System | Introduction

This system was described by Gaskell and Langley around the end of the 19th - beginning of the 20th century. At first it was thought that the autonomic nervous system functioned independently of the central nervous system, but this is not the case.

The ANS is regulated by the central nervous system (CNS), especially the hypothalamus. We also know that cognitive and emotive processes from the higher brain centra also influence the ANS. The ANS plays an important role in homeostasis.

The autonomic nervous system is often seen only as an efferent or motor system, but there are afferent or sensory pathways and inflows to the central nervous structures, which are involved in the regulation of the efferent or motor function. We will now look at both these pathways in more detail in the next few screens.



Afferent pathways

These autonomic fibres are unmyelinated C fibres . Stimulation is both internal and external.

Internal stimulation

The fibres originate in receptors that provide sensory information about various bodily functions in the visceral organs. These receptors can be categorised into two main groups. Each group has further categorisation. This is summarised in the table below.

Pressure changes	Baroreceptors
	Stretch receptors
	Volume receptors
Chemical changes	Chemoreceptors
	Osmoreceptors

Section 3 - Lesson | The Autonomic Nervous System | Afferent pathways

Activity: Afferent pathways

Match the example with the type of receptor. Drag the example to the correct place in the table. Click on the Check Answer button to mark your effort

Pressure changes	Baroreceptors	
	Stretch receptors	
	Volume receptors	
Chemical changes	Chemo - receptors	
	Osmo - receptors	

Stimulated when the bowel is full

Decrease in O_2 saturation of blood

Changes in electrolyte balance

An increase in blood pressure

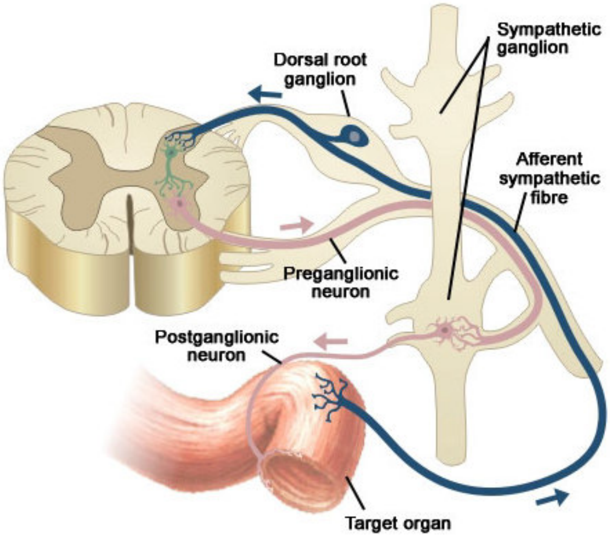
A full bladder

Check answer

External stimulation

- Changes in the external environment can also stimulate the afferent pathways - temperature changes, extreme climates and environments e.g. underwater pressure during diving.
- Psychological stressors.

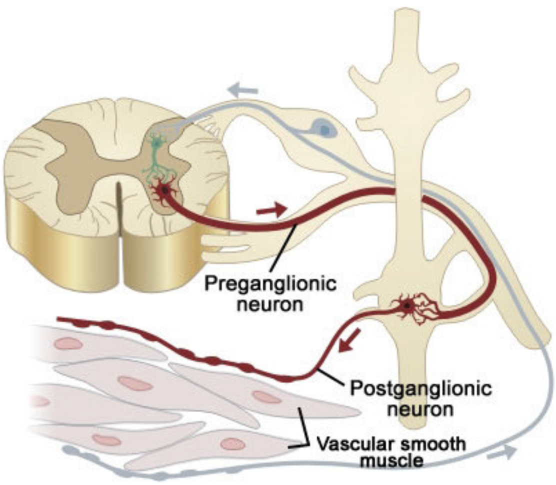
The afferent pathway is illustrated here (using blue).



Efferent pathways

The functional unit of the **efferent division** of the autonomic nervous system is a two-neuron motor pathway, which consists of a preganglionic neuron and a postganglionic neuron. The preganglionic fibers are typically myelinated B fibres, whereas the postganglionic fibers are unmyelinated C fibres.

This is illustrated here >>>>>>>>>>>>>>>>>>>>



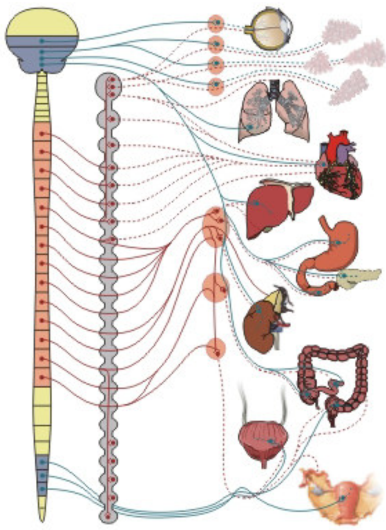
Introduction to the Structure of the SNS

Now that we have a broad overview of the ANS, we can look at the structure of the three divisions in more detail.

Each neuronal pathway, be it sympathetic or parasympathetic, has an origin, a synapse between a preganglionic and a postganglionic fibre and releases a neurotransmitter at **both** the synapse and the target or effector organ.

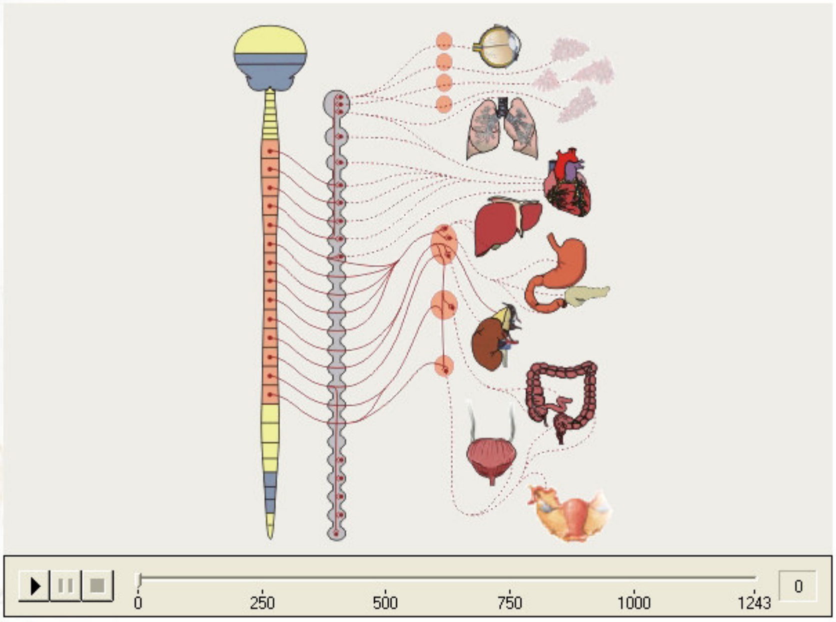
The sympathetic and parasympathetic divisions however, have distinct anatomical differences and release different neurotransmitters at their target sites.

We will now consider the Sympathetic Nervous System (SNS).



Structure of the system

This is a narrated animation - listen using the earphones



Click on the button to play the animation

Divergence and convergence

There are two phenomena in the sympathetic nervous system, known respectively as:

- Divergence
- Convergence

Click on each concept for an illustrated explanation

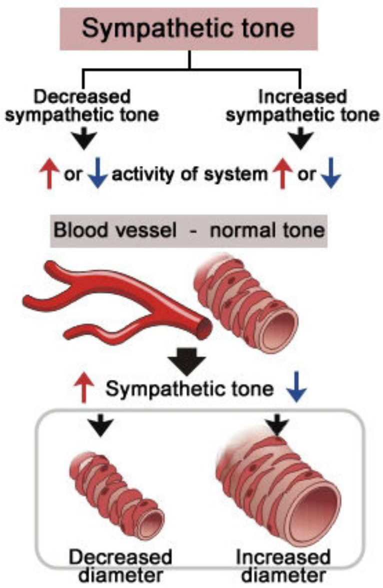
The sympathetic nervous system is continually active. This basal continuous rate of activity is known as the sympathetic tone. This feature enables the sympathetic system to either increase or decrease the sympathetic-regulated activity of the system - by either increasing or decreasing the stimulation.

A good example, illustrated here, is the control of the vascular diameter and thus blood pressure through...

- Increasing alpha-1 stimulation which results in a decrease in vascular diameter and thus an increase in blood pressure

OR

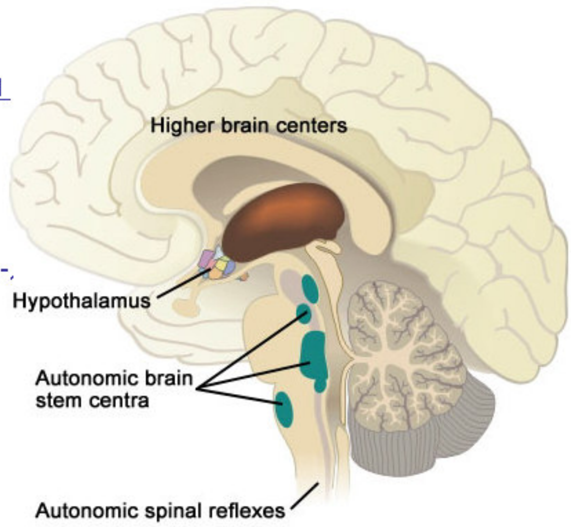
- Decreasing alpha-1 stimulation which results in a increase in vascular diameter and thus a decrease in blood pressure



Control of the ANS

The activity of the sympathetic nervous system is controlled by:

- autonomic spinal reflexes - first level of control,
- autonomic centra in the brain stem - second level of control,
- the autonomic nuclei of the hypothalamus - third level of control -, and
- by higher brain centra.



Autonomic Control Areas

Innervation of the target organs

The postganglionic neurons of both the sympathetic and parasympathetic systems end at the neuroeffector junction of the target organ. Here the release of the neurotransmitters takes place. The neurotransmitters have various effects on the target organ. More detail about this is another lesson.

Some target organs have dual nerve supply, i.e. they are affected by both the sympathetic and parasympathetic system, while others are mainly innervated by one or the other divisions of the ANS.

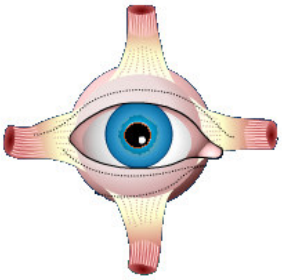
This is explained and illustrated in more detail on the next screen.

Innervation of the target organs

Click on the organ, then click on either the SNS or the PNS button below

Sympathetic Innervation | Parasympathetic Innervation | Reset

Eyes



Still under construction - the link for the Eyes is active

Summary of the Autonomic Nerve Supply and effects on visceral organs

A. Organs with dual supply

Organ	Sympathetic effects	Parasympathetic effects
Iris	Pupil dilation	Pupil constriction
Bronchioles	Dilation	Constriction
Heart	Rate ↑	Rate ↓
Digestive system	Motility ↓	Motility ↑
Sphincters of digestive tract	Contraction	Relaxation
Urinary bladder	Relaxation of wall ⟶ filling	Contraction of wall ⟶ emptying

B. Organs with mainly sympathetic supply

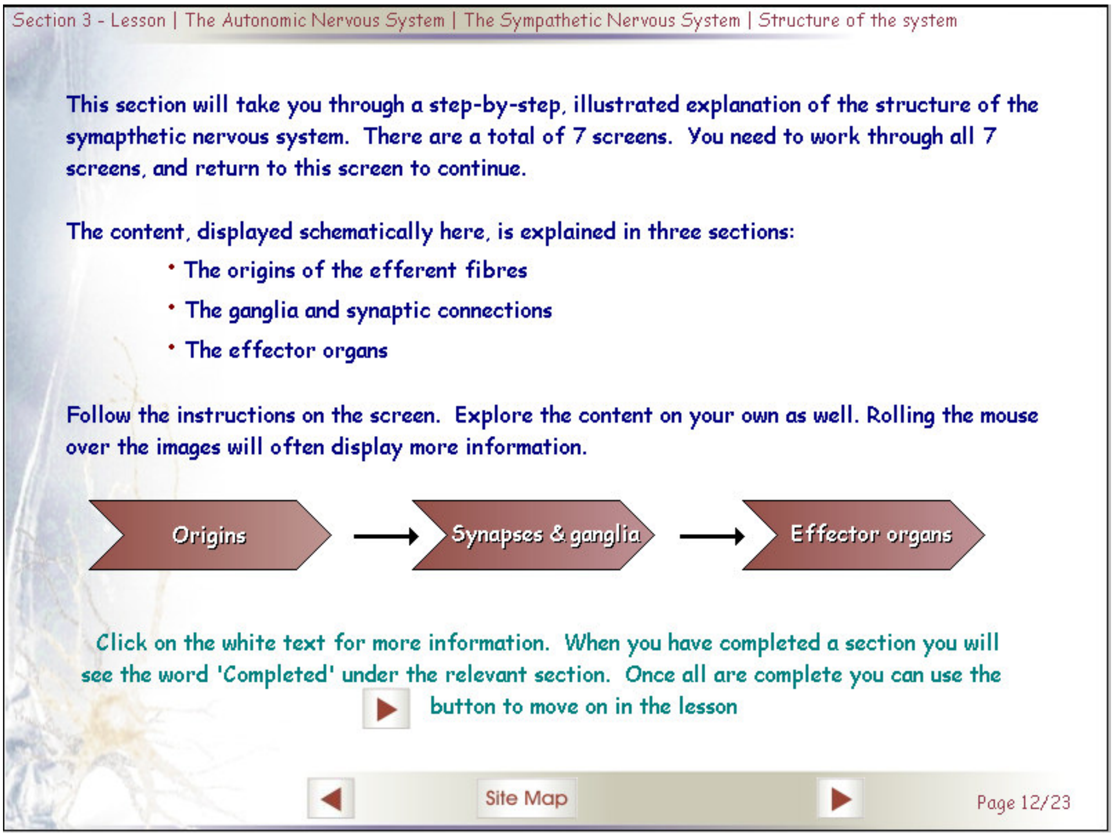
Organ	Absence of sympathetic activity	Increased sympathetic activity
Blood vessels of skin	Vasodilation	Vasoconstriction
Sweat glands	None	Secretion

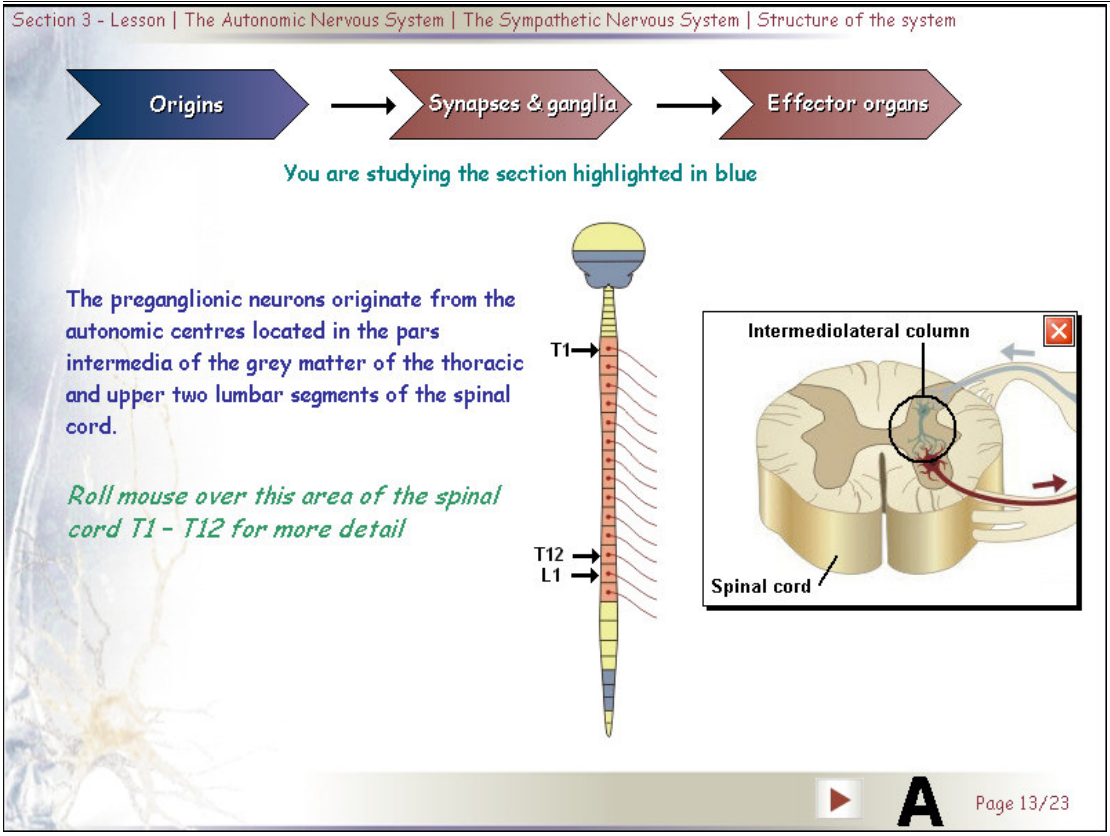
C. Organs with mainly parasympathetic supply

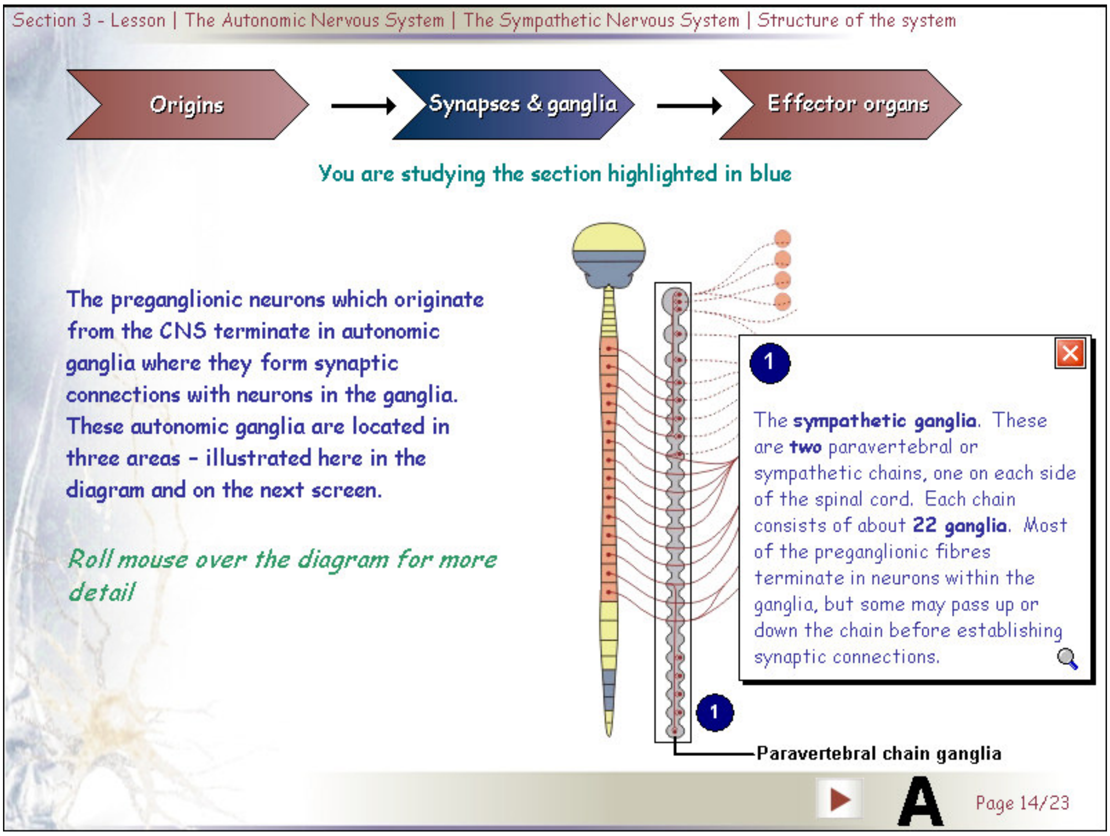
Salivary glands	Secretion of large amounts of watery enzyme-rich saliva. (Stimulation of sympathetic fibres produce small amounts of a thick, viscous secretion)
Stomach glands	Secretion of enzyme-rich gastric juice
Pancreas glands	Secretion of enzyme-rich pancreatic juices
Sex organs	Vasodilation in erectile tissue ⟶ erection

Appendix D: Version 2

Screen 1 – 11 are the same for Version 1 and Version 2



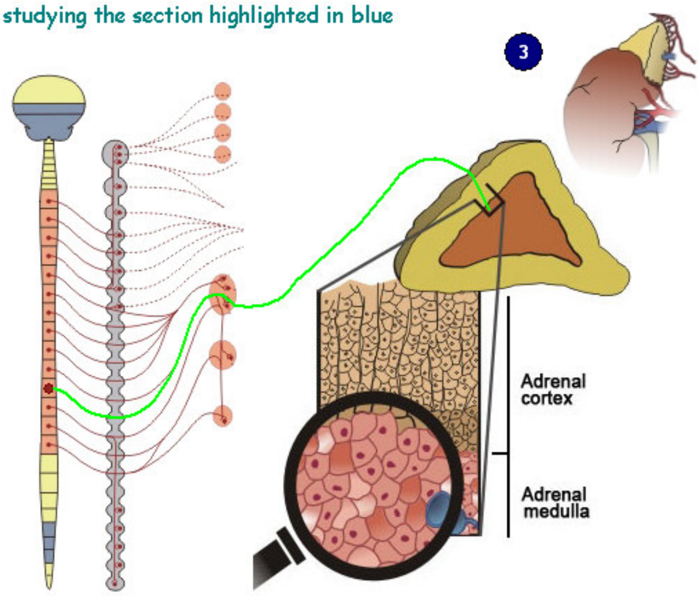

Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system
Origins
Synapses & ganglia
Effector organs
You are studying the section highlighted in blue
The preganglionic neurons originate from the autonomic centres located in the pars intermedia of the grey matter of the thoracic and upper two lumbar segments of the spinal cord.
Roll mouse over this area of the spinal cord T1 – T12 for more detail
T1
T12
L1
Intermediolateral column
Spinal cord
Page 13/23


Section 3 - Lesson | The Autonomic Nervous System | The Sympathetic Nervous System | Structure of the system
Origins
Synapses & ganglia
Effector organs
You are studying the section highlighted in blue
The preganglionic neurons which originate from the CNS terminate in autonomic ganglia where they form synaptic connections with neurons in the ganglia. These autonomic ganglia are located in three areas – illustrated here in the diagram and on the next screen.
Roll mouse over the diagram for more detail
1
The sympathetic ganglia. These are two paravertebral or sympathetic chains, one on each side of the spinal cord. Each chain consists of about 22 ganglia. Most of the preganglionic fibres terminate in neurons within the ganglia, but some may pass up or down the chain before establishing synaptic connections.
1
Paravertebral chain ganglia
Page 14/23

Origins → **Synapses & ganglia** → Effector organs

You are studying the section highlighted in blue

The preganglionic fibres (illustrated in green) leave the spinal cord and pass directly to the **medulla of the adrenal gland**, which can be likened to a modified sympathetic ganglion. This is the third area where there is a synaptic connection. The cells of the adrenal medulla (illustrated in the magnified area) are innervated by the preganglionic sympathetic fibers.



A

Origins → Synapses & ganglia → **Effector organs**

You are studying the section highlighted in blue

Innervation of the target organs

The postganglionic neurons of both the sympathetic and parasympathetic systems end at the neuroeffector junction of the target organ. Here the release of the neurotransmitters takes place. The neurotransmitters have various effects on the target organ. More detail about this is another lesson.

Some target organs have dual nerve supply, i.e. they are affected by both the sympathetic and parasympathetic system, while others are mainly innervated by one or the other divisions of the ANS.

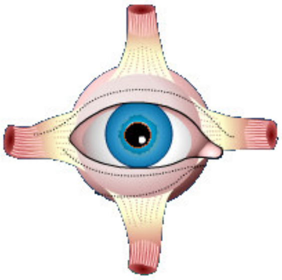
This is explained and illustrated in more detail on the next screen.

Innervation of the target organs

Click on the organ, then click on either the SNS or the PNS button below

Sympathetic Innervation | Parasympathetic Innervation | Reset

Eyes



Still under construction - the link for the Eyes is active

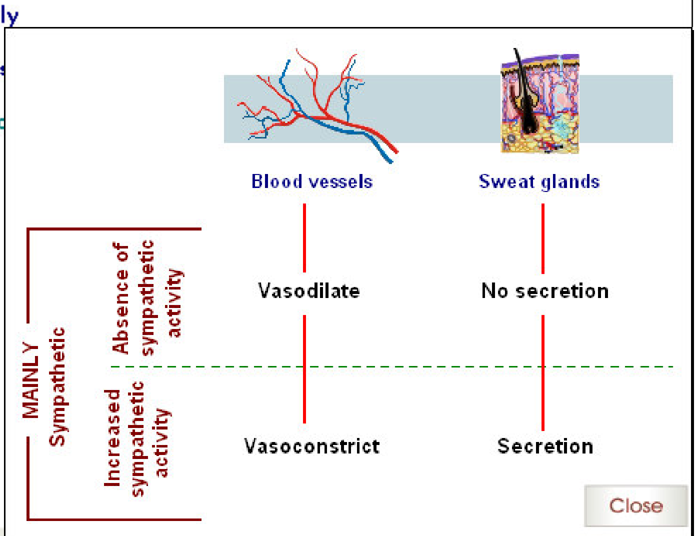
A

Summary of the ANS and effects in visceral organs

Visceral organs have:

- Both sympathetic and parasympathetic supply
- Mainly sympathetic supply
- Mainly parasympathetic s

Click on the bulleted text fo



Divergence and convergence

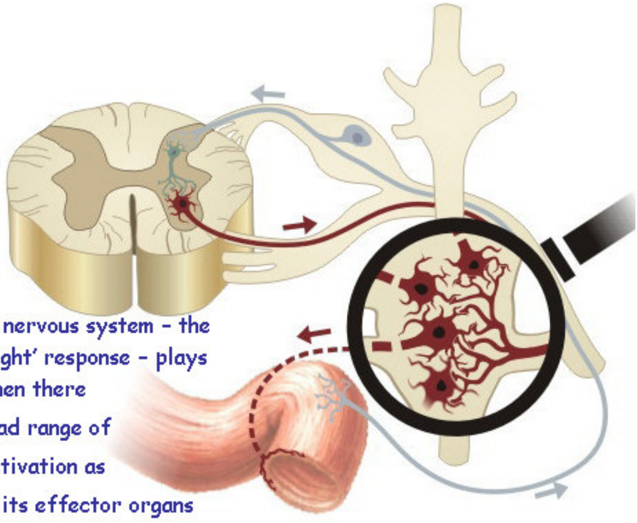
There are two phenomena in the sympathetic nervous system, known respectively as:

- Divergence
- Convergence

<<<< Click on each concept for an illustrated explanation

Divergence

This is the phenomenon where a single preganglionic neuron synapses with an average of ten postganglionic neurons in the sympathetic ganglion. This SIMULTANEOUSLY activates more organ systems . Divergence occurs primarily in the SNS.



This mass activation of the sympathetic nervous system – the so-called mass discharge or 'fight-or-flight' response – plays an important role in stress responses when there is a need to coordinate changes in a broad range of systems. Because of mass activation (activation as a single unit) the SNS can affect all of its effector organs at the same time.

Site Map

A Page 20/23