Literature Review and Theoretical Framework 2

2.1 Introduction

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

2.2 Overview of Chapter 2

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

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

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

Table 2.1: Detailed outline of the parts of Chapter 2

2.3 Literature sources

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

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

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

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

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

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

Part 1 - Theoretical Frameworks

2.4 Introduction to the theoretical frameworks of this study

Decisions about the development of instructional material are not only content and context dependent, but depend increasingly on a skill set that is not part of the average lecturer or facilitator’s experience (Inglis, 1999). Contemporary multimedia for face-to-face, blended and/or distance learning environments are currently developed by teams, and it is a feature of these teams that each member of a team will bring different skills and perspectives to the design and development table. It is primarily the task of the instructional designer in any such team to decide on the variety and blend of instructional strategies and methods that will be used in any particular case. When making such decisions, an instructional designer needs to take the whole gamut of current theories about learning, available research-based evidence and empirically tested theory into account.

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

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

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

Dalgarno describes three broad principles of constructivist theory:

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

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

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

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

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

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

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

2.5 The cognitive theory of multimedia learning

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

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

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

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

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

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

In multimedia learning, that is to say, learning from words, pictures and sounds presented in any number of combinations, the learner needs to process information that arrives from more than one sensory channel. The central tenet of the cognitive theory of multimedia learning is that the design of multimedia instruction affects the degree to which learners engages in the cognitive processes within these two systems.

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

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

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

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

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

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

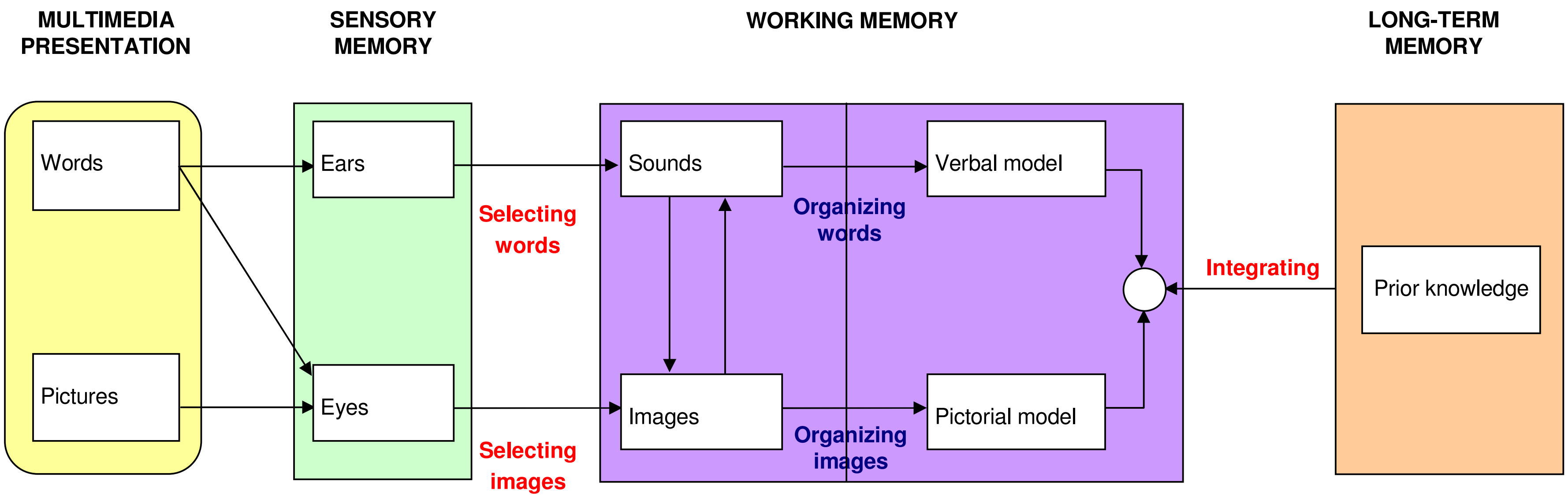

Figure 2.1: A framework for a cognitive theory of multimedia learning (Mayer, 2003:129)

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

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

2.6 Cognitive load theory

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

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

> *Research using cognitive load theory is a vibrant, active enterprise around the globe. The papers of this issue provide both theory and data attesting to the strength of this continuing work.*
>
> *Sweller, 2006b, pg 367.*

2.6.1 The human cognitive architecture

The foundation for understanding and using the CLT is an understanding of the human cognitive architecture. The assumptions about the human cognitive architecture that underpin the development of cognitive load theory are based on the well known and widely accepted limitations of working memory. Pioneering work in this field was being undertaken by Miller as early as 1956 (Miller, 1956). The architecture referred to is a cognitive system that consists of three memory structures: a very powerful long-term memory, a limited working memory and a limited sensory memory. Working memory can only handle a few new interacting elements at one time. Long-term memory expands processing ability due to the creation of schemas. Schemas are cognitive constructs that incorporate multiple elements of information into the ambit of one coherent cognitive construct (Paas, Renkl & Sweller, 2003). We are not consciously aware at any given moment of the information that exists in our long-term memory. We are only aware of whatever information appears in our working memory. Information is transferred into working memory from both the long-term memory and the sensory memory. Paradoxically, even though the working memory has a limited capacity, its limitations seem only to apply to information arriving from the sensory memory. A vast amount of information can be transferred from the long-term memory to the working memory without overloading the working memory in any way (Sweller, 2004).

2.6.2 The construct 'cognitive load'

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

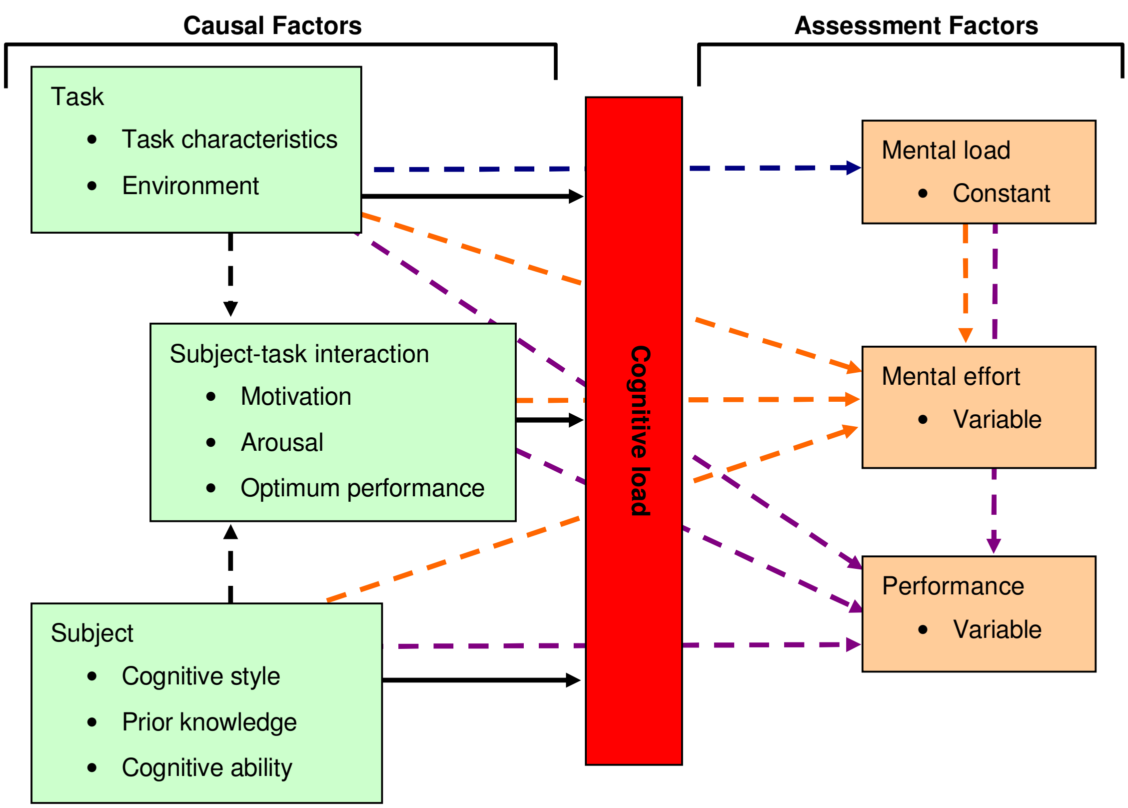

Figure 2.2: The construct cognitive load

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

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

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

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

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

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

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

- Motivation
- Arousal
- Optimum performance

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

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

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

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

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

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

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

These different forms of cognitive load are all additive.

2.6.3 Practical application of cognitive load theory

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

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

2.7 Riding's cognitive style model

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

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

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

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

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

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

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

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

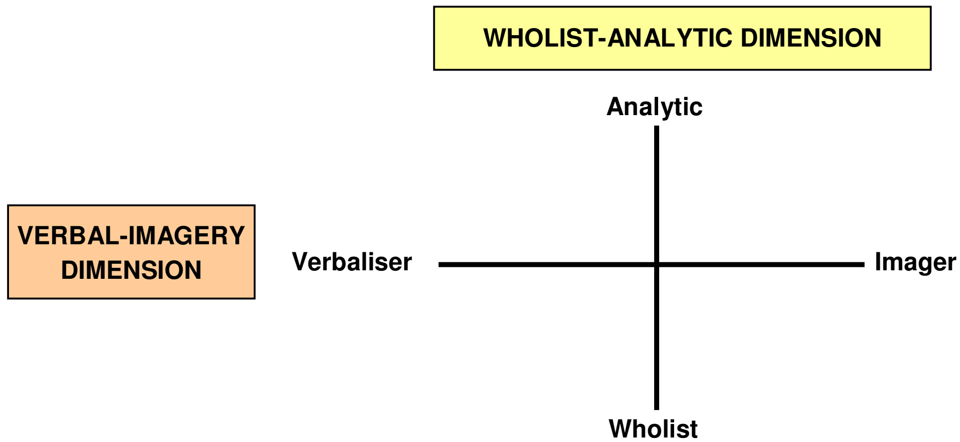

Figure 2.3: Riding's Cognitive Style Model

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

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

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

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

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

2.7.1 Wholist-Analytic Style

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

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

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

2.7.2 Verbal-Imagery Style

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

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

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

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

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

2.8 Summing up Part 1 of the Literature Review

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

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

The second theory that is used in this study is the cognitive load theory of Sweller. The foundation of this theory is a description and understanding of the function of the human cognitive architecture. Sweller takes the position that human beings possess a finite cognitive capacity which is characterised by, among other things, limited attentional resources and a limited working memory. When the limits of attention and working memory are reached, performance declines. The information that must be processed by this cognitive architecture is called cognitive load, and when cognitive load is too great, the resulting situation is called cognitive overload. Three following types of cognitive load have been identified: intrinsic load(which relates to difficulty and the interrelatedness of the content), extraneous load (which relates to design and which is a hindrance to cognitive processing), and germane load (which is concerned with motivational issues and strategies that promote deeper processing). There are several different contributors to cognitive load, and these include cognitive style. The effects of cognitive load are manifested in the performance of the learner. A learner is also able to report the amount of mental effort invested in a particular learning task, and mental effort is used as an indicator of cognitive load. This theory is applicable to all learning, irrespective of the delivery medium.

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

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

Part 2 - Literature Review

2.9 Introduction to Part 2

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

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

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

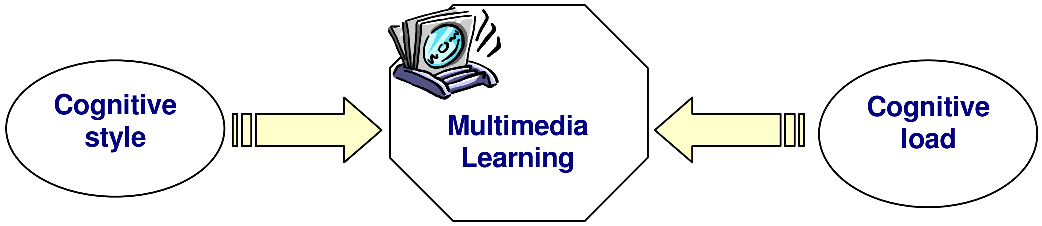

Figure 2.4: Relationships between cognitive load, cognitive style and multimedia learning

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

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

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

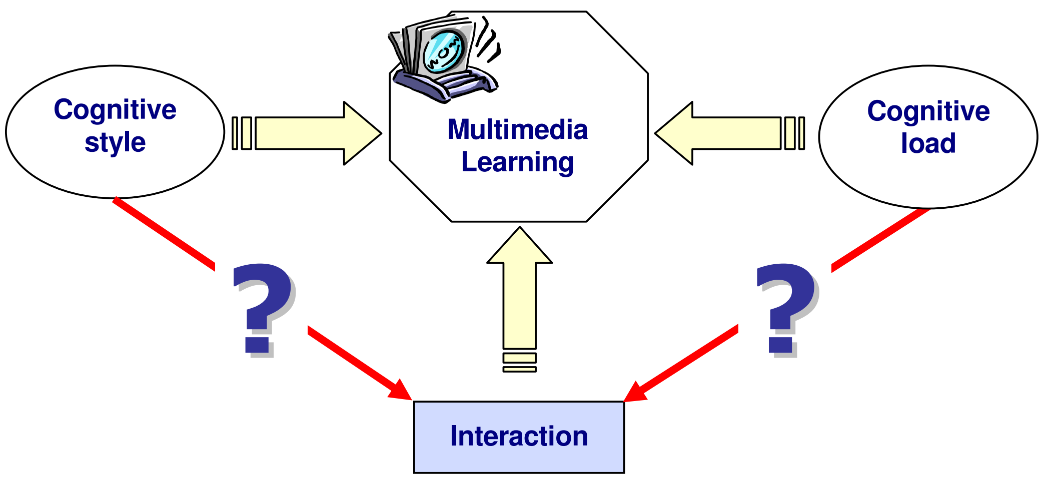

Figure 2.5: New relationships between cognitive style, cognitive load and multimedia learning

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

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

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

Table 2.3: Approach to the literature review

2.10 Cognitive styles and multimedia learning

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

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

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

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

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

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

2.10.1 Contributions to the theory

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

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

John and Boucouvalas (2002a) propose a new definition of cognitive style that takes an auditory component into account and is based on the ways in which individuals perceive different types of media. In order to test this theory, they developed a different but equivalent set of questions that can be used with the questions in Riding's CSA. In this new test the Verbal-imagery questions were spoken instead of being displayed as text, and the Wholistic-Analytic questions were transmitted by means of sound effects rather than graphics. The performance of subjects in the visual test was then compared to their performance in the audio version. The hypothesis of the researchers was that if cognitive style influences an individual's perceptions of both visual and auditory information, there should be a strong correlation between the ratios achieved in both versions of the test. The researchers found there were strong correlations between both versions of the text in the Analytic-style questions, as indeed there were with the questions in the Imager-style questions. In spite of this, the overall style dimensions did not correlate well across the two versions of the test. The differences in performance between the visual and audio experiments seems to indicate that Riding's classification of cognitive style does not explain the performance of users when information is presented using a combination of visual and audio components. Is this model of style therefore able to explain multimedia learning which – in multimedia research and the context of this study – includes both an audio and visual component? John and Boucouvalas's proposal to add a visual-audio dimension to the current Verbaliser-Imager classifications does seem to provide new direction for cognitive style theory.

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

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

They found that:

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

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

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

2.10.2 Measurement of cognitive style

2.10.2.1 Riding's CSA and related research

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

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

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

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

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

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

> *In considering psychological assessments the most important feature of a test is its construct validity - if there is no evidence that it assesses what it purports to measure then it is of no use*
>
> *(Riding, 2005b, pg 6).*

With research into the CSA the primary emphasis has therefore been to demonstrate its validity. Riding and Rayner (1998) summarise the evidence for validity as follows:

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

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

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

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

Rezaei and Katz (2004) also investigated the reliability of the CSA in three studies that used a test-retest interval of one week for the first study and one month for the second and third studies. The range of reliability for the Wholistic-Analytic dimension across the three studies was between 0.42 and 0.55, and for the Verbaliser-Imager dimension it was between 0.30 and 0.45. This, like the findings in the other studies, is still low. Rezaei and Katz went one step further and made suggestions for improvements to the CSA by addressing the limitations of the commercial version. Their suggestions include the following:

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

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

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

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

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

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

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

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

2.10.2.2 Allinson and Hayes' CSI

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

This call was taken up by Sadler-Smith, Spicer and Tsang (2000). Their study (n=1050) sampled employees from a large range of occupational and professional contexts and included some graduate students. Sadler-Smith et al. found that the internal reliability was satisfactory across all sub-samples (0. 84 - 0. 90). They also found that it compared well with the values cited by the developers. When they performed factor analysis, they obtained results that were similar to those of Allinson and Hayes. They also examined these styles in relation to other constructs, and used Riding's CSA because it had been demonstrated to possess both construct and concurrent validity. Their findings, however, indicated that there was no statistically significant correlation between the CSI and the Verbaliser-Imager or Wholistic-Analytic dimension of Riding's CSA. They suggested that the styles measured by the CSI and the CSA are orthogonal, and their suggestions for future research include a recommendation to consider the construct validity of the CSI. Both Sadler-Smith et al. and Riding feel that it is a matter of the greatest importance to establish the validity of the instrument. These researchers feel that this is crucial to the empirical and theoretical elaboration of the cognitive style construct, otherwise the cognitive style field will suffer the same fate as the learning style field, that is, it will become so diluted and confused by the use of terms that at face value have the same label, but have different meanings to different groups of researchers and may even describe different constructs.

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

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

2.10.2.3 Determining the Verbaliser –Imager dimension of style

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

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

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

Massa and Mayer (2005) raise several concerns about the validity of the Verbal-Imager subtest of the Cognitive Styles Analysis. In their summary of the evidence, they state that this subtest of the CSA lacks face, construct and predictive validity. Their critique of face validity is that the test does not appear to measure the test-takers primary mode of processing or representing information in thought. No information is requested about how the information is processed. Massa and Mayer do, however, concede that the ability of any test to ascertain what thought modality a person is using is must be interpreted very cautiously because participants might not know whether they are using their verbal or visual information processing. The work by Mayer and Massa (2003) that was discussed in the previous paragraph, provides evidence that the Verbaliser-Imager dimension of the CSA does not have construct validity. Massa and Mayer also report on a study in which they compared learning behaviour with scores on a test of style in order to determine the predictive validity of the instrument. The process measure of the Verbaliser-Visualiser cognitive style was the number of times a learner selected a pictorial help screen rather than a verbal help screen. They used Riding's CSA and their own Verbal-Visual Learning Style Rating to determine the Verbaliser-Imager-style of the participant. Riding's CSA did not correlate significantly with any of the process measures. By contrast, the Verbal-Visual Learning Style Rating had strong, significant correlations with the process measures. The implication for my study, which will use Riding's CSA is that results must sill be interpreted cautiously due to the limitations of the instrument.

2.10.2.4 Other measures of cognitive style

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

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

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

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

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

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

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

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

> *Basic lack of agreement over nomenclature and few reliable and valid instruments suitable for assessing cognitive style in applied settings threatens the viability of this construct. But if cognitive style was a unitary construct then it would be necessary that the majority of instruments developed to measure cognitive style are inter-correlated with one another. This evidence has not been forthcoming.*
>
> *(Hodgkinson & Sadler-Smith, 2003).*

2.10.3 Cognitive style and learner characteristics

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

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

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

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

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

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

2.10.4 Cognitive style and behaviour

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

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

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

Graff (2005) researched the different web browsing strategies used by older and younger participants, on one hand, and individuals who displayed Verbaliser and Imager cognitive styles, on the other hand. He developed different web architectures: a simpler hierarchical architecture that allowed users to browse up and down the hierarchy, and a more complex architecture that allowed users to browse hierarchically and laterally across the different topics. Graff hypothesised that Imagers would visit a greater proportion of pages in a complex hierarchy as they strove to obtain a big picture view, and that they would also tend to browse high in the hierarchy rather than dig deep in their efforts to locate specific information. His findings were that Verbalisers and Bimodals visited more pages and a higher proportion of pages in the simple hierarchy, while Imagers visited more pages and a higher proportion of pages in the relational condition. This is consistent with a Verbaliser's need to obtain detail by drilling down into the hierarchy for each topic, and with the need of Imagers to get the big picture by browsing more laterally in the structure. My critique of this study is that participants were given 10 minutes to browse through 64 pages of content. Although they were told that they would be answering questions about the content at the end of the session, this was not clarified in the published article. Since ten minutes seems to be a very short time for such a task, it is valid to ask whether this would not impose a high cognitive load on the learner. Since no performance results were reported in the assessment, it is not possible to decide whether or not cognitive load might have played a role in this instance. What *is* relevant here is whether a similar pattern of navigation will be observed in my study in which one might expect Imagers to move quickly through the program to get the big picture before coming back to each screen to study the content in more detail, and in which Verbalisers might be expected to work systematically through the lesson only once and to spend the amount of time on each screen that would allow them to do this. Screen logs of this behaviour were recorded so that it would be possible to determine this information.

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

Self-perception in general has many dimensions. These are efficacy, ability, concept and esteem. While the influence of self-perception on performance will sometimes be negligible, it will at other times exert a major influence. Perception about oneself as a learner is influenced by motivation, interest in a subject and performance. Riding and Staley (1998) explored the relationship between self-perception and cognitive style with a sample of first-year Business Studies university students. The main focus of the study was on style and the differences between perception and performance. The role of self-perception for a learner is in the regulation of learning. Riding and Staley propose that the level of self-regulation that learners adopt, and hence their motivation, will be influenced both by the learning experience itself and by the influence of performance on outcomes. Learning experience includes whether or not learners find the subject matter easy, the attractiveness or otherwise of the presentation, and the extent to which the material makes sense to them. When one looks at their proposal from a cognitive load frame of reference, the extent to which learning material is easy and extent to which it makes sense to the learners speaks to the issue of intrinsic cognitive load, while the attractiveness of any particular presentation is addressed by extraneous cognitive load. If both the learning experience and performance are likely to be influenced by the cognitive style characteristics of the individual learner, then the question ‘And what role or influence does cognitive load have?’ is one that may also require investigation. The results of this study indicated that performance in all the subjects included in the study was related to the nature of the tasks within each subject area: thus, in those cases where the acquisition of the big picture was necessary for a proper grasp of the subject (as is the case in Management), Wholistic learners performed better than Analytic learners. One might therefore hypothesise that if there is a mismatch between subject and style, learners might report that they needed to invest more mental effort in order to understand the work, so indicating a higher cognitive load. What type of load is being influenced here? It is most likely to be germane load, which has to date not yet been directly measured.

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

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

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

> *The unitary aspect of this cognitive style also adds another requirement: this style can deal with large amounts of information.*
>
> *(Evans, 2004, pg 512).*

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

2.10.5 Cognitive style and instructional design practice

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

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

Graff (2003b) made use of a web-based lesson to investigate the influence of segmentation of information and the extent to which an overview facilitated learning. His findings suggested that cognitive style and segmentation have an effect on performance, and that the Wholist-Analytic dimension of cognitive style is the style that determines the degree of success that individuals have when they attempt to learn from web-based systems. Segmentation may exacerbate the Analytics tendency to see information in parts because it encourages them to focus more strongly on the parts at the expense of seeing the whole picture, and this – in the long run – may be detrimental to learning. Analytics should therefore be able to learn more efficiently from material that is not heavily segmented. For Wholists, segmentation would make very little difference because of their inherent ability to see the whole. But the nature of the content is also important. Which is more important: the big picture or are the parts? Because of its inherent ability to link information that is conceptually related, the web has the potential to present instructional information more effectively than traditional linear methods. But one has to take into account the disadvantages of fragmentation and the lack of discourse cues (indicators of how information is presented). The findings also suggested that the provision of an overview had little effect on learning performance. This leads me to return to the question: 'Is the big picture more important or are the parts more important?' If the learning of the parts is more important, the provision of an overview would tend to benefit Wholists because they strive to see the big picture. It would make little difference to Analytics. This angle, which considers the nature of the content, was not part of Graff's investigation. In addition to this, users of web-based systems have to multitask (navigate, read and understand). All these activities increase the cognitive load. Graff's study contains a very brief reference to the fact that there might be a relationship between cognitive style and cognitive load, where the load is affected by the amount of extraneous load, and which includes tasks such as navigation, rather than information processing at a cognitive level. Graff did indicate that it is possible for an instructional system with different subject information and also (possibly) a system incorporating a multimedia component, to yield a different finding. Multimedia brings with it a greater potential for cognitive overload. This suggests another reason why it might be helpful to explore the relationship between cognitive load and style in the context of web-based and other multimedia learning.

2.10.6 Cognitive style and achievement

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

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

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

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

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

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

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

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

The implication of this is that instructional strategies need to be put in place to assist those learners whose style does not naturally 'fit' the nature of the content. Riding and Agrell's study leads us to believe in this regard that while the more intelligent and adaptable learners will probably be able to solve this problem for themselves, the less and adaptable intelligent ones will need and should be trained in the art of developing strategies. In those cases where, by contrast, the styles are naturally appropriate, there will be little incentive to develop strategies because learners will usually be able to cope reasonably well.

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

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

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

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

2.10.7 In summary

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

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

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

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

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

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

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

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

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

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

2.11 Cognitive load and multimedia learning

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

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

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

Not all multimedia research literature addresses the question of cognitive load (that research will be discussed in Section 2.11.4 and 2.11.5 of this chapter). Multimedia learning environments were also not a topic of investigation in early cognitive load research(Chandler & Sweller, 1996; Sweller & Chandler, 1994). While research into instructional strategies that use paper-based material has been ongoing (Robinson, 2002; Verdi & Kulhavy, 2002), the proliferation of technology in education has increased multimedia learning environments to such an extent that research on this topic is now included in the cognitive load research stream.

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

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

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

2.11.1 Instructional design issues

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

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

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

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

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

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

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

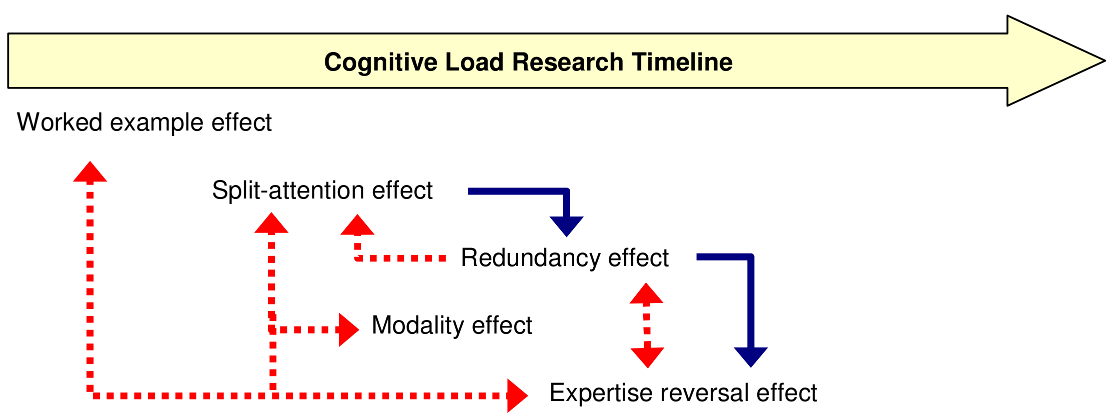

Figure 2.6: Progression for the 'effects' research within the cognitive load research stream

My review of the literature reveals that once these various effects had been described and tested in practice, it became increasingly difficult for subsequent researchers to consider each of the effects in isolation. The split-attention effect for one group and context, for example, later became a redundant effect for another group and context (Yeung, Jin & Sweller, 1997). Another example is that researchers investigated the modality effect as a possible solution for solving the problems generated by the split-attention effect (Kalyuga, Chandler & Sweller, 1999). For the purpose of this review, however, I will discuss each effect separately and make appropriate references to the other effects where necessary.

2.11.1.1 Split attention effects

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

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

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

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

The research of Kalyuga, Chandler and Sweller (1999) overlapped to some extent with the research of Mayer and Moreno (1998). These researchers investigated alternatives to split-attention instructional designs from a cognitive load perspective. These alternatives included narration rather than text, as well as colour to limit the load of searching for information in diagrams. The authors pointed out that dual-mode presentations did not reduce cognitive load per se, but rather increased the effective working memory capacity. In the first experiment there was evidence that the dual-mode format (audio-text) was significantly more effective than the visual only (diagram-text) and the audio-diagram-text formats. They concluded that one could use the modality effect to reduce the negative effects of split-attention, and they produced evidence to show that the redundancy effect (diagram plus text plus audio) negated the positive findings of a modality effect. The use of a diagram **and** text **together with** narration induced a redundancy effect which overloaded working memory and influenced learning performance negatively. In contrast to what happened in the early research of 1991 (Chandler & Sweller, 1991), these researchers also measured the cognitive load. They found that the diagram-audio format returned the lowest cognitive load rating. In their second experiment ,Kalyuga et al. (1999) were able to demonstrate that colour coding, if used with care, could also be used as a technique for reducing the cognitive load imposed in those situations where learners are compelled to integrate content from a diagram and from text. In cases where the same colour was used for text and related sections in a diagram, learners did not have to waste cognitive resources in a search for information so that they could integrate it. While a split-attention effect was still operative, its impact was reduced through the use of appropriate colour coding. In this experiment, the colour-coded version was also rated as having a lower cognitive load by the participants.

2.11.1.2 Redundancy effects

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

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

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

The results of this experiment indicated that verbal redundancy was the better format because the addition of an animation before the verbal learning still produced superior learning performance for the redundant verbal format. In contrast to the findings of Kalyuga et al. (1999), the learners did not ever have to split their attention between visual and verbal material. In a second experiment, Moreno and Mayer added visual material that was displayed at the same time as the narration or narration-text format. They also presented animation before the narration and narration-text formats. In two of these formats the learner had to process a corresponding animation simultaneously, thereby creating the split-attention effect. They found that when the material was presented sequentially, the verbally redundant condition produced better learning, but that when the presentation was concurrent (a split-attention effect), the redundant conditions produced less efficient learning. In a final experiment they included non-verbal auditory materials (background sounds), and looked at how this would affect learning performance. They found that the addition of sound did not help students to understand the learning material. Moreno and Mayer (2002) concluded that the most effective technique for enabling learning with multimedia explanations was to use the auditory and visual modalities simultaneously for verbal information. In order to avoid cognitive overload this strategy would not be effective if other visual material such as diagrams and animation) was presented simultaneously. A learner cannot watch verbal and visual material and listen to verbal material simultaneously without experiencing high cognitive load.

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

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

2.11.1.3 Expertise-reversal effect

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

Kalyuga, Chandler and Sweller (1998) investigated the role of learner's prior knowledge in the context of training for trade apprentices using elementary electrical engineering instructional materials. A series of three experiments was conducted. Because the learners in Experiment 1 were inexperienced, the results provided evidence for the split-attention effect. These inexperienced learners performed best when the material (diagrams and text) were integrated. Although the learners in Experiment 2 were initially inexperienced, they were given training at different points in time and tested after each stage. The researchers then compared the relative improvement of the groups for the different interventions. The findings indicated that as learner expertise and knowledge increased, the need for using integrated material decreased. Because these results yielded no clear evidence for a full redundancy effect, the researchers concluded that the level of expertise was in all probability not high enough. They therefore conducted a final experiment with more experienced learners and these findings demonstrated a redundancy effect – experienced learners performed better when there was no redundancy. Split-attention effects in the experiment did not appear to influence performance. Kalyuga et al. (1998) made no reference to the work of Yeung et al. (1997), who obtained similar results in another context and subject domain.

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

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

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

2.11.1.4 Modality effects

The modality effect states that learning will be enhanced if textual information is presented in an auditory rather than (the conventional) visual format, and if such an auditory format is accompanied by visually based information such as graphs, diagrams or animations (Ginns, 2005). Modality effects have been researched in both the cognitive load and multimedia learning research streams. In this section, I will look at the research undertaken from a cognitive load perspective, and in Section 2.11.4, I will examine the research undertaken from the perspective of multimedia learning.

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

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

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

The effect size used in the analysis was *d*, which is the difference between the means of the different conditions divided by the pooled standard deviation. Between-subject and within-subject designs were analysed separately. When the between-subject designs were compared, the overall weighted mean effect size was large $d = 0.72$ (95% confidence interval 0.52–0.92). Further analysis indicated that the mean effect size was greater for high element interactivity materials than for low element interactivity material. The pacing of the presentation was also a significant moderator, with system-paced instruction yielding a substantially higher effect size than self-paced material. The results also suggested that the modality effect might be particularly strong for students who were learning by means of virtual reality media. Ginns's analysis supported the three major hypotheses, even though some of the studies did publish contradictory findings. Ginns (2005) concluded that there was a need for considerably more research. The reasons for requiring more research were as follows:

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

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

2.11.1.5 In conclusion

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

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

2.11.2 Theory development and directions of research

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

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

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

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

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

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

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

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

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

Dixon says:

> *In theory development, the experiments provide the basis for distinguishing among different possible accounts of the cognitive processing that occurs in the situation of interest.*
>
> *(Dixon, 1991, pg 345).*

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

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

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

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

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

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

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

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

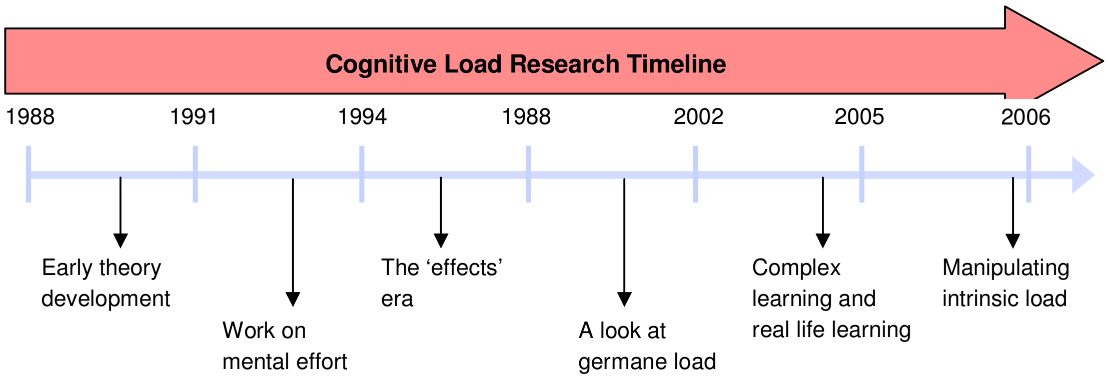

Figure 2.7: Cognitive load research timeline

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

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

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

Kirschner (2002) proposed that designing for a competency-based paradigm, for which there are growing calls in education, calls for new approaches to instruction. While he was of the opinion that cognition, meta-cognition and transfer were the most important variables in this new paradigm, he still regarded CLT as a good tool for helping designers to understand and to take into account the limitations of the human mind. He felt that designers needed to make use of this knowledge in designing for transfer of learning. In his introduction to the special issue on cognitive load theory, Kirschner (2002) described the trend in the following way: it is a movement from research that had, as its sole focus, sought ways to decrease extraneous load toward research that had started to consider germane load in instructional design. While he left the mapping of future directions to Valcke (2002), he highlighted several consequences for CLT that arise out of the research published in the special edition.

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

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

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

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

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

In yet another special edition from 2003, Paas, Renkl and Sweller (2003) offered an overview of research that had investigated the more dynamic approaches to instructional design. These approaches to design were based on new understandings of cognitive load theory, which proposed that changes occur in cognitive load as the learner moves from being a novice to being a expert within a particular domain of knowledge. The cognitive load research stream had started to look at manipulating more than just extraneous load in order to reduce the total cognitive load, and was focusing more on the various design strategies that accommodated a novice or expert learner's ability to cope with cognitive load (Renkl & Atkinson, 2003). Cognitive load research was also beginning to consider the role of the goal of instruction from the point of view of the teacher and the learner (Gerjets & Scheiter, 2003).

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

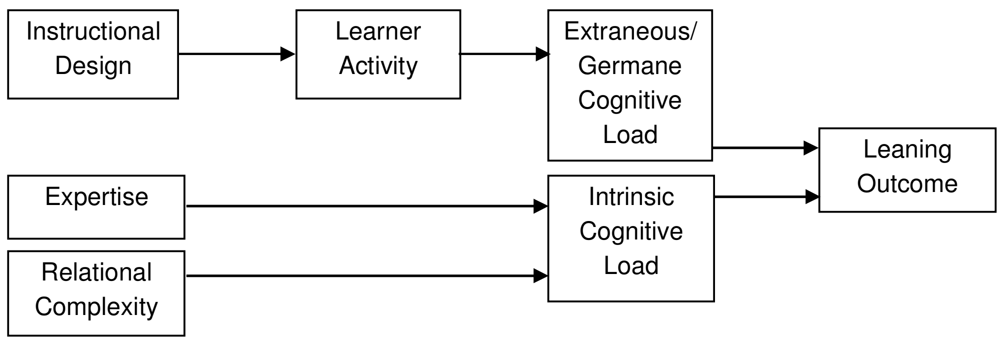

Figure 2.8: Assumptions of cognitive load (Gerjets & Scheiter, 2003, page 35)

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

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

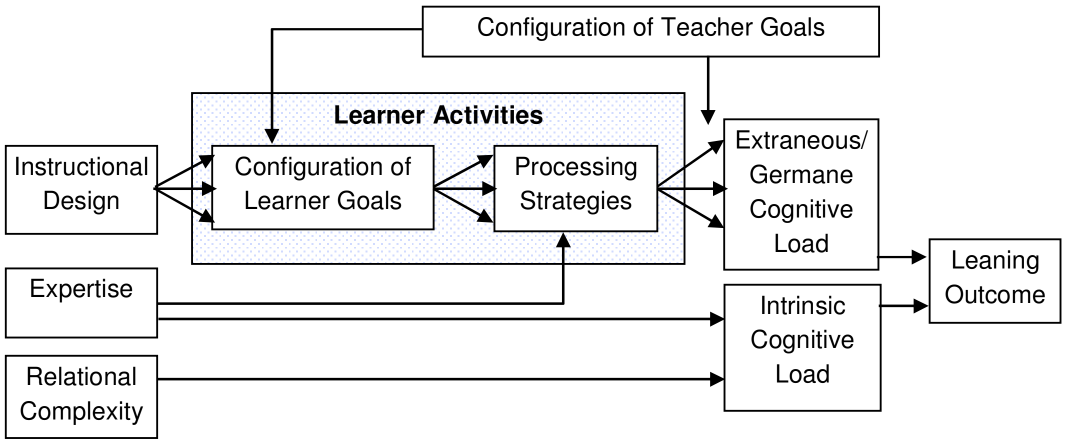

Figure 2.9: Goals and strategies as moderators between instructional design and cognitive load (Gerjets & Scheiter, 2003, page 36)

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

The emergent empirical research of 2004 could be divided into two streams: the research that extended the investigations carried out by previous work, and new directions in research. The new theoretical perspectives of that year included drawing an analogy between evolution by natural selection and human cognitive architecture (Paas, Renkl & Sweller, 2004; Sweller, 2004). The research field, while continuing to increase in diversity, was still engaged in investigations into the worked-example effect, the role of learner expertise, and the synergies between these two areas.

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

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

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

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

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

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

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

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

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

2.11.3 Measurement of cognitive load

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

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

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

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

Table 2.4: Assumptions underlying different techniques for measuring cognitive load

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

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

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

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

After the first study (Paas, 1992), Paas et al. (1994b) took a more rigorous approach to investigating the measurement of cognitive load using mental effort techniques. Mental effort may either be measured subjectively by using rating scales, or objectively by using physiological parameters. By using the data from two previous studies (Paas, 1992; Paas & van Merriënboer, 1994c), the researchers investigated the reliability and sensitivity of the subjective rating scale. In the 1992 study, the coefficient of reliability (Cronbach's alpha) was 0.90 with the use of 28 measurements, and in the 1994 study, it was 0.82 with the use of 6 measurements. 'Sensitivity' was defined in these studies as the capability of a measurement technique to reflect differences between training and transfer conditions. When defined in this way, the subjective rating scale was found to be a 'sensitive' measure. The physiological method that used heart-rate variability (Paas et al., 1994b) proved to be a less reliable method. Although all the measures, except one, correlated significantly, the correlations were so low that Paas and van Merriënboer concluded that the technique was not reliable. This method could also not be defined as a sensitive measure in terms of the definition described above. The subjective rating scale is very easy to use and implement in a classroom situation. It is not an intrusive or invasive test. When it was used together with performance measures, it proved to be a promising technique for providing additional information about cognitive load. This subjective rating technique remains widely used today and, when reported, demonstrates good reliability since it returns a Cronbach alpha of more than 0.80 (Kester, Kirschner & van Merriënboer, 2004; Stark, Mandl, Gruber & Renkl, 2002; Tabbers, Martens & van Merriënboer, 2004).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Fink and Neubauer (2001) investigated the potential of using the subjective estimation of time as an index of cognitive load. What they found was that as the cognitive load increased, so the subjective estimation of the time duration on task decreased. They also explored this subjective time estimation in participants who registered different levels of intelligence. They found that although the subjective time estimation decreased for both groups as the cognitive load increased, the group with higher intelligence produced more realistic time estimations. They also found that speed of processing was related to intelligence, and concluded that since the more intelligent participants processed more quickly when performing cognitive tasks, they had more attentional capacity left to estimate time compared to the less intelligent participants who processed more slowly and therefore had no spare capacity left to estimate time accurately. These findings raise the question of cognitive load and intelligence. Fink and Neubauer (2005) extended their work to consider explanations other than mental processing speed to explain the relationship between subjective time estimation and intelligence. They looked at whether the number of tasks completed played any role, but could find no evidence for this explanation. They then conducted two different studies in which they gave the participants tasks that required task processing in working memory and tasks that looked at the speed of processing. They confirmed yet again that subjective time estimation can be used an index of working memory load even though there were also still differences in subjective time estimation that were dependent on intelligence. Although these researchers conclude that there is not sufficient evidence to suggest that the subjective estimation of time is a valid index of task difficulty or cognitive load, I would be cautious in using this technique to measure cognitive load for at least the following three reasons:

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

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

2.11.4 Learning from multiple representations

If multimedia are to be used effectively within the learning environment, then designers and instructors need answers about the potential of multimedia to influence learning outcomes. Mayer (1997a) in the article, “Multimedia learning: Are we asking the Right Questions?”, refers, on the one hand, to the media debate that was conducted between Clark and Kozma in the early 1990s, but, on the other hand, discusses how the consensus among educational/cognitive psychologists has moved in the direction of calls for reframing the research questions that are applied to media. Multimedia learning research (Mayer, 1997a) had focused on establishing whether one medium is better than the other (the so-called media effect)(Chang, 2002; Chong, Balen & Jewesson, 2003; Frith, Jaftha, & Prince, 2004; McDonough, & Marks, 2002; Rutherford & Lloyd, 2001; Sinclair, Renshaw, & Taylor, 2004; Stern et al., 2001). While Mayer discussed three main directions for future research, the essence of these directions can be summarised as a call for research that looks at how the instructional design/treatments within various media influence the cognitive processes and consequences.

This call was for research that will consider

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

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

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

The purpose of the study by Moreno and Mayer (1999) was to clarify and test two cognitive principles: the contiguity principle and the modality principle. The contiguity principle states that learning is enhanced when printed text and pictures are physically integrated or close to each other (spatial contiguity). It also states that learning is enhanced when visual and spoken materials are temporally synchronised rather then presented successively (temporal contiguity). This principle is also described in the cognitive load literature under the heading of split-attention effect. According to the modality principle, words should be presented as auditory narration rather than as visual on-screen text. The question that they set themselves was: 'In multimedia learning with animations, what are the relative contributions of spatial contiguity and modality to multimedia learning, and what are the relative contributions of temporal contiguity and modality?' They hypothesised that if the advantage of narration over on-screen text resides in a modality principle, then the advantage for auditory-visual presentations should not disappear when the presentations are sequential rather than contiguous. Their first experiment provided support for both the modality and spatial contiguity principles. A second experiment, in which they used sequential presentations, allowed the researchers to distinguish between the modality and the contiguity effects. Their findings provided more evidence for the modality effect than for the temporal-spatial contiguity effect.

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

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

More recent research has investigated the effects of adding audio (Moreno & Mayer, 2000a; Moreno & Mayer, 2000b; Mayer, Sobko & Mautone, 2003) to multimedia instruction. Moreno and Mayer (2000a) tested the hypothesis that personalised explanations promote deeper understanding because the learner becomes actively engaged in the elaboration of the materials. They proposed that learners used less cognitive effort to process verbal information when it was presented in a familiar style. In a series of five experiments, they found that deeper learning took place when personalised messages were used. A design guideline that flowed from this research was that the learner should be addressed as a participant rather than an observer. Apart from motivating the inclusion of audio from a dual-processing/cognitive load perspective, additional reasons for considering adding audio are that a speaker's voice in multimedia lessons carries important social cues that can influence the process and outcome of learning. This promotes the idea that the inclusion of voice adds a social dimension to learning in a computer-based environment, the lack of which in the use of computers for education and training has been occasionally criticised.

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

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

Moreno and Mayer (2000b) investigated the use of other sounds in multimedia such as background music and sounds. There are two theoretical positions in this regard: the one position claims that adding audio makes the learning task more interesting and increases the learner's enthusiasm. The other position holds that the addition of audio can overload the auditory processing channel. The corollary of the second position is that sound that is not required to make the lesson intelligible or that is not integrated with the rest of the material will reduce the effective working memory capacity and interfere with learning. Moreno and Mayer devised two experiments whose results demonstrated that when they added unnecessary background music that was irrelevant to the content, the performance of participants on recall and problem-solving transfer tests dropped significantly. One of their experiments showed that relevant environmental sounds that have been integrated into the learning material facilitate processing and exert no negative influence on learning. They were not, however, able to replicate this finding in the second experiment. These findings are consistent in general with accumulated knowledge about the limitations of working memory and the cognitive load and dual-processing theories.

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

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

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

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

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

As an instructional designer I have often been involved in debate about the use of multiple representations within the same program. Such practice is often defended with claims along the lines of 'It accommodates different learning styles', or 'It is a good thing for learners to see the topic from different perspectives' or 'Repetition reinforces learning'. I have already discussed the issue of redundancy (the cognitive load perspective) in this chapter. A study undertaken by Kozma (2003) looks at the issue from a multimedia learning perspective. Kozma looked at the material features of external, multiple representations and the cognitive and social affordances they provide in support of understanding chemistry. He also looked at the different ways in which scientists and students used these multiple representations. In a laboratory, experimental setting Kozma (2003) found that experts were able to cluster apparently dissimilar problems or situations into large meaningful groups on the basis of underlying principles. He found that expert chemists were able to use conceptual terms to label their clusters and that they also used a greater variety of representations in their clusters. In contrast to this, he found the novices (the students) labelled their subject groupings on the basis of surface features. Another task that participants were required to perform was to transform each representation into another form. They would have to, for example, transform animation into a corresponding graph and a video into an equation. The experts were significantly better at this than novices, who tended to use the surface features of the displays in attempts to build an understanding of the chemical phenomena they represented. Kozma then replicated this research in a naturalistic setting. Observation of the environment in which the chemists worked revealed that representations of chemistry were everywhere in the laboratory: there were, for example, diagrams on whiteboards and posters on the walls. One of his observations was how frequently the chemists used these diagrams to explain a point or to seek clarification while solving problems. These experts were able to make effective use of different representations to understand the chemical phenomenon that they were examining. They used these representations to support their arguments and to explain and justify their reasoning. The students, on the other hand, did not function in this way at all. Kozma was of the opinion that the study conducted in the naturalistic setting confirmed the findings of the laboratory experiment.

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

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

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

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

Seufert's study was designed to investigate how these types of help influence coherence formation in learners who have different levels of prior knowledge. The first research question looked at the relative value of these two kinds of help. Directive help was found to be significantly superior to non-directive help for recall tasks, while there was no significant difference with regard to the different types of help when it came to comprehension tasks. When she looked at the influence of prior knowledge on performance, Seufert found that no help and directive help were more useful for low prior-knowledge learners, and that non-directive help exerted a detrimental effect on performance. She also found that learners with a medium level of prior knowledge benefited the most from directive help, which in fact seemed to be the best type of help to provide all round for recall tasks. Seufert (2003) concluded that the type and amount of help to give depends on the learning goal. Directive help seemed to support recall performance because of its summarising and repeating function. Directive help was also more effective for comprehension, contrary to the expected finding that non-directive help would be more effective. Prior knowledge also mediated the effects of help. Low prior knowledge learners, for example, cannot use help effectively at any point: their need is for additional instruction on the content. A notable observation made during this study was that the high knowledge group, who did not seem to need help, did not reach the maximum level of performance. Seufert describes this as the 'illusion of knowing'. It means that learners with high prior knowledge overestimate their abilities and do not seem to realise that help can enable them to perform even better. She recommended that instructors might need to introduce help to this group in another way – and to point out to them that help could very well enable them to improve their performance.

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

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

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

> *…once psychologists knew almost all about the psychology of the command-drive interface, there were no command-driver user interfaces in use anymore*
>
> *(Reimann, 2003, pg 251).*

Van der Meij and de Jong (2006) investigated the conditions that provided the best support for learners who use multiple presentations in a dynamic simulation-based learning environment. This study aimed at extending the work of previous research which had investigated the integration of images and text (Chandler & Sweller, 1991) or images, text and audio (Mayer & Moreno, 1998; Tabbers, et al. 2004). Since the learning environment had also been designed to be more complex, simulation was being used as an instructional strategy rather than an interactive tutorial. These researchers compared the strategies of integration and dynamic linking. The simulation interface had anywhere between three and five representations, depending on the format. Integration combined the five different representations. In the dynamic linking condition, actions performed on one representation were shown simultaneously in the other representations.

There were three different representation formats:

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

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

2.11.5 Animations, including animated pedagogical agents

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

Chandler summed it up well:

> *.….despite this seemingly endless potential and unbridled enthusiasm for technology based instruction, there is little empirical evidence to indicate that the widespread use of dynamic visualisations has resulted in any substantial benefit to learners.*
>
> *(Chandler, 2004, page 353).*

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

Animation has several uses. It can support 3-D perception by showing an object from different perspectives; it directs attention to important aspects of a display; it conveys procedural knowledge; it demonstrates the dynamics of the subject matter; it enables exploratory learning through the manipulation of a displayed object. Animation can help a learner to perform a cognitive process that they could not otherwise have performed without this kind of external support (Schnotz & Rasch, 2005). These researchers are of the opinion that the conditions under which animated pictures really enhance comprehension and learning is still an open question.

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

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

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

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

She identified the following three categories of strategy:

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

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

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

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

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

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

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

They propose that there are three types of dynamic representation, and each has distinct informational and computational properties and different uses. **Time–persistent** animations display a relation between at least one variable and time. For example, in a health care context a table comparing the incidence of HIV/AIDS and Tuberculosis (TB) over a number of years could be animated. This would provide a visual representation of the increases and/or decreases of HIV/AIDS and TB over a specific time period and how they might or might not re related to each other. This is similar to static images except that the dynamic presentation displays the data incrementally rather than presenting the whole set at once as in a static display. While no new information is added because of the animation, it might make the features of that information more salient. This type of animation displays the greatest amount of information, and this can place a high cognitive load on the learner. While **time–implicit** animations (which are often used in simulations) show a range of values, they do not show the time when these occurred unless you are looking at the animation. Using the same example, a graph could be created with HIV/AIDS on one axis and TB on the other. The information appearing in this graph is then animated, using time as the variable that determines the sequence of the visual display in the animation. When the simulation as been stopped, the learner needs to invoke internal representations to compare current values to a particular previous state in order to answers questions about the timescale. **Time–singular** animations display one or more variables at a single instant in time. For example, a pie chart could be created to show the relative proportions of HIV/AIDS and TB for each year. The series of pie charts is then animated, but when the animation is stopped the learner only sees the representation for one time period and must hold the other views in memory if any comparison is needed. This can place a high load on the cognitive resources of the learner.

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

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

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

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

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

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

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

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

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

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

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

Chan and Black (2005) proposed a format-support hypothesis of learning, which was then tested in a study that compared the performance of learners when using either static images, a system – controlled animation or a version where the user could manipulate the animation. The rationale behind this hypothesis was that comprehension and learning would be enhanced if the presentation format supported the learner's need to construct a dynamic mental model of a particular phenomenon. They argued that the benefits of using animation could only be tested and ultimately realised if there was a match between the format of the animation and the learning outcome. The learning task in their study required an understanding of the dynamic interchange between kinetic energy, potential energy and total energy. The target group was 7^{th} grade students. They used a roller coaster ride as the context for explaining the phenomenon. There were three assessment tasks: written recall, create a diagram of the phenomenon and problem-solving tasks. Participants who used the user-controlled / manipulation version of the animation performed significantly better on the recall and drawing tasks than the groups using the other versions. Participants who used the static visual version performed significantly better than the groups using the other versions on the transfer tasks. These findings might seen surprising. Why would the group who used static images outperform the groups using animation? The explanation for this finding is that when learners use a static visual they must animate the process mentally. This requires deeper cognitive processing and engagement with the content, which helps develop a solid understanding of the content. Similar findings have been obtained in several other studies (Garg et.al. 1999, Schnotz & Rasch,2005). These findings are also similar to those of Hegarty, Kriz and Cate (2003), who found no evidence of better understanding of a dynamic process after the use of animated diagrams compared to static diagrams.

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

2.11.6 In summary

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

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

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

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

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

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

The last two themes considered learning from multiple representations and animation. There has been extensive research into the use of text, images and audio, in various combinations, with solid evidence that designs using the principles of dual-coding result in superior learning performance. Multiple representations can also result in cognitive overload. The more expert the learner the better able they are in the use of multiple representations, both in the experimental laboratory and naturalistic setting. The findings regarding the use of animation as an alternative to static images & test are still mixed. The evidence for the superiority of animation over static images and text is not overwhelming. Learners are generally unable to extract the important information from an animation, especially if their prior knowledge is poor. They often tend to focus on the superficial qualities of the animation. This poses a design challenge to ensure that there is nothing in the animation that distracts the learner. Researchers still query whether the benefits of using animation justifies the cost of development and production.

2.12 Multimedia in health sciences education

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

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

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

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

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

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

2.12.1 The media effect

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

* indicates that this study also addressed perceptions and attitudes

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

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

Table 2.5: The media effect in health sciences multimedia education

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

2.12.2 The multimedia effect

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

* indicates that this study also addressed perceptions and attitudes

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

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

Table 2.6: The multimedia effect in health sciences multimedia education

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

2.12.3 The Contiguity effect

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

Discipline	Study	About the study	Findings
Anatomy **Intervention: The first study** used two strategies – key views that rotated by 90° and key views with multiple views (10° rotation). **The second study** used key views that rotated by 180°and multiple views that rotated by 10°.	Garg, Norman, Spero & Taylor, 1999	This study assessed carpal bone spatial ability by making use of a 36-item test.	Familiarity with more views made no difference to student performance. When they were briefed, the students remembered only certain key views, and then rotated the image mentally so that they could answer questions presented from the non-standard views. The authors concluded that it was possible that the multiple views were redundant and/or confusing.
	Garg, Norman & Spero, 2001	This study assessed carpal bone spatial ability by making use of A 50-item test.	In this study there was a significant difference in spatial carpal ability in favour of the key view format. Spatial ability also affected performance. Students also first considered the key view and then mentally rotated the view before answering the question. Those who used the multiple view model spent more time looking at the key views in this model than the other views.

Table 2.7: The contiguity effect in health sciences multimedia education

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

2.12.4 Interaction effect

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

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

Table 2.8: The interaction effect in health sciences multimedia education

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

2.12.5 Perceptions and attitudes

NOTE: Italic text indicates author's emphasis.

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

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

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

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

2.12.6 Meta-analyses

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

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

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

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

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

2.12.7 Multimedia in Physiology Education

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

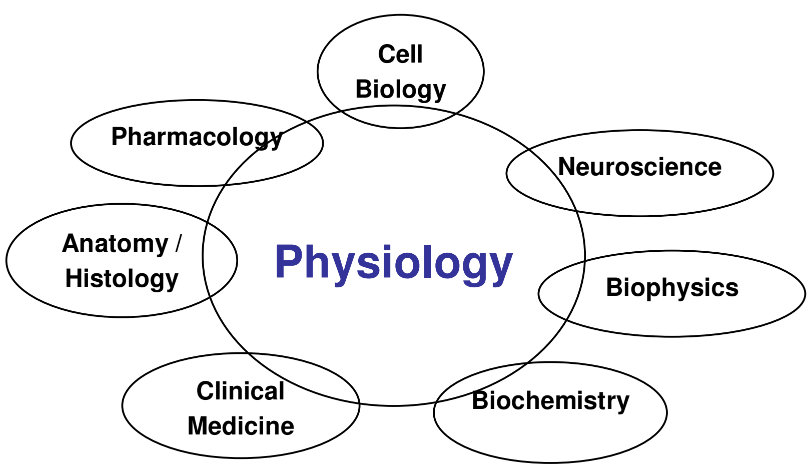

Figure 2.10: Relationships of physiology with related Medical Science (Sefton, 1998, pg. 54)

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

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

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

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

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

Both Rawson and Quinlan (2002) and Griffin (2003) address the nature of the subject and the requirement that students have to integrate multiple sources of information in order to understand some of the concepts and processes in Physiology. Another interesting factor in Michael's 2007 study (Michael 2007), but lower in the ranking, was the fact that students tended to ignore the graphs, tables and figures in their learning material. Is this indicative of a Verbaliser profile (which would focus on using textual rather than visual material) in the student group, or does it point to a lack of appropriate study skills? I did not find any reference to the cognitive styles of students in the Physiology education literature, and a discussion of study skills is beyond the scope of this review.

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

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

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

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

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

Table 2.11: Review of recent research in Physiology education

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

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

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

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

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

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

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

While calls continue to be made for research-based evidence for the need to change teaching and learning practices in health science education (Michael, 2006), the evidence is there and it has been extensively and thoroughly discussed in this literature review. Health science educators need to look beyond their own field of education and take note of this evidence. Health science education needs to move beyond research into students' attitudes toward and perceptions of multimedia programmes (see Table 2.9), and it needs to focus more on finding effective instructional strategies that meet two important challenges: firstly, that of a field where the knowledge base that is growing faster than people can keep up with, and, secondly, the life-long learning needs of the health science practitioner.

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

2.12.8 In summary

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

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

2.13 Conclusion following the literature view

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

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

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

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

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

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

Table 2.12: The literature review and the research questions

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