Cognitive load theory and music instruction

Author: Owens, Paul
Publication Date: 2005
DOI: https://doi.org/10.26190/unsworks/22943
License: https://creativecommons.org/licenses/by-nc-nd/3.0/au/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/22994 in https://unsworks.unsw.edu.au on 2023-08-04
COGNITIVE LOAD THEORY
AND MUSIC INSTRUCTION

Paul Owens

A thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

School of Education
University of NSW

November 2005
Declaration

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

P.O.
Acknowledgements

After many years I tentatively returned to university in February 1998, the first lecture bringing me into contact with my eventual supervisor John Sweller. I was from this moment captivated, intrigued and inspired to keep learning; nothing has changed in the intervening years. John has been endlessly encouraging, always supportive and an extraordinary mentor, inspiring a commitment to both the rewards of research and the values of academic rigor. I will always be indebted to him for the warmth, remarkable knowledge and the guidance he has so readily given; the qualities of scholarship and aspiration I have gained from working with John are legacies to be shared with my family and students.

I am very grateful for the assistance and generous support given by Gary McPherson. Gary’s advice regarding music research was invaluable in the formulation of experimental materials. I always appreciated his relaxed manner and the complete professionalism with which he encouraged the investigation of matters musical.

I would also like to acknowledge the support of all staff from the School of Education, University of NSW. They have established a productive research community that exerts a powerful influence over those who are fortunate enough to work in this stimulating environment.

My deepest love and respect is extended to my wife Cathy; without her endless support, forbearance and invaluable advice this project would never have been brought to completion. Her capacity to anticipate the pressures and to negotiate the difficulties is simply indicative of her equable, compassionate and resilient approach to life. As with all my personal and family endeavours, this work has been a road travelled together.

To my children Michael and Daniel, who have been so patient with a father who always seems attached to the computer. Thank you for continually reminding me of what really matters and of why any research in education should be undertaken. My ardent wish is that in some way, no matter how intangible, the end result of all these many hours will also benefit you!
Without the support of Lynne Irvine the empirical work undertaken in this thesis would have been difficult to complete. As an educational leader, her liberal and receptive approach to learning continually provides an environment where the investigation and evaluation of instructional practices can take place. Your support has not gone unnoticed; thank you.
Table of Contents

Declaration ........................................................................................................................................... ii
Acknowledgements ........................................................................................................................... iii
List of Figures ........................................................................................................................................ x
List of Tables ......................................................................................................................................... xiii
List of Appendices ............................................................................................................................. xiv
Abstract ................................................................................................................................................ xv

Chapter 1: Human Cognition .............................................................................................................. 1
  1.1 Introduction: Memory and Learning ......................................................................................... 1
  1.2 Human Memory: Models and Conceptions .............................................................................. 5
  1.3 Sensory Memory ...................................................................................................................... 11
    1.3.1 Processes and meaning ..................................................................................................... 11
    1.3.2 Capacities and functions .................................................................................................. 13
  1.4 Working Memory ..................................................................................................................... 17
    1.4.1 Capacities and mechanisms ............................................................................................. 17
    1.4.2 The nature of items (elements) ....................................................................................... 22
    1.4.3 A component model of working memory ........................................................................ 25
    1.4.4 Modality-based processing ............................................................................................. 28
  1.5 Long Term Memory (LTM) ...................................................................................................... 32
    1.5.1 Structures and models ...................................................................................................... 32
    1.5.2 Encoding and retention .................................................................................................... 38
    1.5.3 Schemas .......................................................................................................................... 48
  1.6 Expertise and Problem Solving: Implications for Cognitive Mechanisms ......................... 55
Chapter 4: Introduction to Experiments ................................................................. 157

4.1 Background to Empirical Investigations ....................................................... 157

4.2 Cognitive Architecture, Cognitive Load Theory and Music Instruction:
   An Overview .................................................................................................. 159

4.3 Overview of Experiments ............................................................................. 166

Chapter 5 Experiment 1 ....................................................................................... 169

5.1 Introduction .................................................................................................. 169

5.2 Method ......................................................................................................... 171
   5.2.1 Participants ......................................................................................... 171
   5.2.2 Materials and procedure ................................................................. 172

5.3 Results ....................................................................................................... 180

5.4 Discussion ................................................................................................. 183

Chapter 6 Experiment 2 ....................................................................................... 186

6.1 Introduction ................................................................................................ 186

6.2 Method ....................................................................................................... 188
   6.2.1 Participants ....................................................................................... 188
   6.2.2 Materials and procedure ............................................................... 188

6.3 Results ...................................................................................................... 194

6.4 Discussion ................................................................................................. 197

Chapter 7 Experiment 3 ....................................................................................... 202

7.1 Introduction ............................................................................................... 202
7.1.2 The Musical Aptitude Profile - Measuring Audiation Skills

(Experiments 3-6)........................................................................................................... 205
7.2 Method ...................................................................................................................... 207
7.2.1 Participants......................................................................................................... 207
7.2.2 Materials and procedure .................................................................................. 207
7.3 Results ................................................................................................................... 214
7.4 Discussion ............................................................................................................. 220

Chapter 8 Experiment 4 .............................................................................................. 224
8.1 Introduction ........................................................................................................... 224
8.2 Method .................................................................................................................. 228
8.2.1 Participants....................................................................................................... 228
8.2.2 Materials and procedure ................................................................................ 228
8.3 Results .................................................................................................................. 234
8.4 Discussion ............................................................................................................. 240

Chapter 9 Experiments 5 and 6 .................................................................................. 244
9.1 Overview of Experiments ....................................................................................... 244
9.2 Introduction to Experiment 5 ................................................................................ 245
9.3 Method .................................................................................................................. 248
9.3.1 Participants....................................................................................................... 248
9.3.2 Materials and procedure ................................................................................ 248
9.4 Results .................................................................................................................. 254
9.5 Discussion ............................................................................................................. 258
9.6 Introduction to Experiment 6 ................................................................................ 260
9.7 Method .................................................................................................................................263
  9.7.1 Participants ..................................................................................................................263
  9.7.2 Materials and procedure ..........................................................................................263
9.8 Results ..................................................................................................................................267
9.9 Discussion ..........................................................................................................................271

Chapter 10: General Discussion ...............................................................................................274
  10.1 Overview of Research ...................................................................................................274
  10.2 Results ..........................................................................................................................277
  10.3 Theoretical and Future Research Implications, and Limitations of Findings ...283
  10.4 Practical Implications of Research Findings ...............................................................294
  10.5 Conclusion .....................................................................................................................298

References ...................................................................................................................................300
List of Figures

Chapter 1

Figure 1  Atkinson and Shiffrin’s (1968) modal model of memory.............................8

Figure 2  Baddeley’s (1997) model of working memory........................................25

Figure 3  Paivio’s (1986) dual-coding model.............................................................42

Figure 4  A simple example of the flexible way in which schemas help us to
interpret variations among otherwise common objects........................................51

Figure 5  Far Side Gallery (Larson, 1992); schemas and humour ............................53

Figure 6  The Tower of Hanoi problem....................................................................64

Chapter 2

Figure 7  A scale illusion (Sloboda, 1985). .................................................................75

Figure 8  Berz’s (1995) model of working memory...................................................78

Chapter 3

Figure 9  Dimensions for measuring cognitive load. ..................................................109

Figure 10 Formula for the calculation of relative condition efficiency ($E$)..........113

Figure 11 Example of a self-sufficient diagram (Chandler & Sweller, 1991). ........129

Chapter 4

Figure 12 Example of a simultaneously delivered format (Experiment 4) ............158

Figure 13 Example of split-source materials (Experiment 2).................................163

Figure 14 Example of dual-modal materials (Experiment 5).................................165
Chapter 5 - Experiment 1

Figure 15  Example of the split-attention format used in Experiment 1. ..................177

Figure 16  Example of the integrated format used in Experiment 1.........................178

Figure 17  Graph (Cartesion plane) of relative condition efficiency for
Experiment 1. ...........................................................................................................182

Chapter 6 - Experiment 2

Figure 18  Example of the dual-modality format used in Experiment 2. ..................192

Figure 19  Graph (Cartesion plane) of relative condition efficiency for
Experiment 2. ...........................................................................................................196

Chapter 7 - Experiment 3

Figure 20  Two examples (Nº 2 & Nº 4) of the explanatory statements from
Powerpoint slides (dual-modal condition) used in Experiment 3. .................209

Figure 21  Graph (Cartesion plane) of relative condition efficiency for
Experiment 3 ..............................................................................................................218

Figure 22  Graph (x, y, plane) of relative condition efficiency for Experiment 3.....219

Chapter 8 - Experiment 4

Figure 23  Example of the integrated written materials used in Experiment 4 .......229

Figure 24  Graph (Cartesion plane) of relative condition efficiency for
Experiment 4 (Written Problems) ...........................................................................238

Figure 25  Graph (Cartesion plane) of relative condition efficiency for
Experiment 4 (Total Test Problems) .....................................................................239
Chapter 9 - Experiments 5 and 6

Figure 26  Example of the integrated written materials used in Experiment 5, indicating a specific variation. .................................................................250

Figure 27  Example of the integrated written materials used in Experiment 5, indicating a general variation. .................................................................250

Figure 28  Graph (Cartesion plane) of relative condition efficiency for Experiment 5 .............................................................................................257

Figure 29  Graph (x, y, plane) of relative condition efficiency for Experiment 5, highlighting the difference between conditions with musical notation ....257

Figure 30  Example (Concept 2) of the integrated written materials used in Experiment 6. ............................................................................................265

Figure 31  Graph (x, y, plane) of significant interaction for mean cognitive load ratings found in Experiment 6.................................................................269

Figure 32  Graph (Cartesion plane) of relative condition efficiency for Experiment 6. ............................................................................................270

Figure 33  Graph (x, y, plane) of mean instructional efficiency for Experiment 6, highlighting the difference between successive conditions with and without musical notation.................................................................270
List of Tables

Chapter 5 – Experiment 1
Table 5.1 Acquisition and Test Problems for Experiment 1 ........................................180
Table 5.2 Cognitive Load and Efficiency Measures for Experiment 1 .....................181

Chapter 6 – Experiment 2
Table 6.1 Acquisition and Test Problems for Experiment 2 ........................................194
Table 6.2 Cognitive Load and Efficiency Measures for Experiment 2 ....................195

Chapter 7 – Experiment 3
Table 7.1 First Set of Similar Test Problems for Experiment 3: Features that
Remained the Same ........................................................................................................214
Table 7.2 Total Similar Problems, First and Second Transfer Problems, and
Written Problems for Experiment 3: Variations ........................................................215
Table 7.3 Cognitive Load and Efficiency Measures for Experiment 3 .....................217

Chapter 8 – Experiment 4
Table 8.1 Test Problems for Experiment 4......................................................................235
Table 8.2 Cognitive Load and Efficiency Measures for Experiment 4 .....................237

Chapter 9 – Experiments 5 and 6
Table 9.1 Test Problems for Experiment 5......................................................................254
Table 9.2 Cognitive Load and Efficiency Measures for Experiment 5 .....................255
Table 9.3 Test Problems for Experiment 6 .................................................................267
Table 9.4 Cognitive Load and Efficiency Measures for Experiment 6 .....................268
Appendices

Appendix A  Examples of perceptual recognition phenomena and Gestalt grouping laws...............................................................344

Appendix B  Examples of mnemonic techniques.................................................................345

Appendix C  The complete “washing clothes” passage from Bransford and Johnson (1973) .................................................................346

Appendix D  The Monk problem ...................................................................................347

Appendix E  Example of functional fixity ........................................................................348

Appendix F  The Water Jug problem .............................................................................349

Appendix G  Means-ends analysis model used by Sweller (1988).................................350

Appendix H  Survey of musical experience ......................................................................351

Appendix I  Music Response Sheet ...............................................................................352

Appendix J  Examples of the instructional format (Experiment 2) for the integrated condition ...........................................................................353

Appendix K  Examples of the melodic materials from Experiment 3 ..........................355

Appendix L  Examples of the melodic material used for transfer problems in Experiment 3 ...........................................................................356

Appendix M  Examples of written and aural tests from Experiment 4 ..................357

Appendix N  Examples of musical variations from Experiment 5 ..........................358

Appendix O  Estimate of interacting elements for Experiment 5 ..............................359

Appendix P  Examples of test questions from Experiment 5 ..................................360

Appendix Q  Estimate of interacting elements for Experiment 6 .............................361

Appendix R  Examples of test questions from Experiment 6 ..................................362
Abstract

Cognitive load theory assumes that effective instructional design is subject to the mechanisms that underpin our cognitive architecture and that understanding is constrained by the processing capacity of a limited working memory. This thesis reports the results of six experiments that applied the principles of cognitive load theory to the investigation of instructional design in music. Across the six experiments conditions differed by modality (uni or dual) and/or the nature of presentation (integrated or adjacent; simultaneous or successive). In addition, instructional formats were comprised of either two or three sources of information (text, auditory musical excerpts, musical notation). Participants were academically able Year 7 students with some previous musical experience. Following instructional interventions, students were tested using auditory and/or written problems; in addition, subjective ratings and efficiency measures were used as indicators of mental load. Together, Experiments 1 and 2 demonstrated the benefits of both dual-modal (dual-modality effect) and physically integrated formats over the same materials presented as adjacent and discrete information sources (split-attention effect), confirming the application of established cognitive load effects within the domain of music. Experiment 3 compared uni-modal formats, consisting of auditory rather than visual materials, with their dual-modal counterparts. Although some evidence for a modality effect was associated with simultaneous presentations, the uni-modal format was clearly superior when the same materials were delivered successively. Experiment 4 compared three cognitively efficient instructional formats in which either two or three information sources were
studied. There was evidence that simultaneously processing all three sources overwhelmed working memory, whereas an overlapping design that delayed the introduction of the third source facilitated understanding. Experiments 5 and 6 varied the element interactivity of either two- or three-source formats and demonstrated the negative effects of splitting attention between successively presented instructional materials. Theoretical implications extend cognitive load principles to both the domain of music and across a range of novel instructional formats; future research into auditory only formats and the modality effect is suggested. Recommendations for instructional design highlight the need to facilitate necessary interactions between mutually referring musical elements and to maintain intrinsic cognitive load within working memory capacity.
Chapter 1: Human Cognition

1.1 Introduction: Memory and Learning

Understanding human cognition has long been a driving force within educational psychology and over recent decades has resulted in a “substantial corpus of knowledge concerning our mental processes” (Sweller, 1999, p. 2). Historically, these mental processes have invariably been linked to learning and the inseparable qualities of human achievement. Consider briefly, our thirst for knowledge, the profound beauty of artistic endeavour, the exquisite demonstration of intricate skill, or the revelatory insight of new ideas, all of which define by action a deeper and unique characteristic of human behaviour: a capacity for complex intellectual thought.

For thousands of years this capacity has intrigued philosophers, who in the process have provided a living embodiment of its very existence. Always central to these musings was the challenge to understand human memory; its capacities most obvious in the astounding recall of individuals ranging from the Greek poet Simonides (Yates, 1966), through to the great Maestro Toscanini (Marek, 1982) and the more recent phenomenological cases documented by the likes of Chase and Ericsson (1981). Towards the end of the Nineteenth Century, psychological research embraced the scientific spirit of the age, and early steps were taken by Ebbinghaus (Bower, 2000) and others to understand both the underlying mechanisms of human memory and their

---

1 Refer to their studies involving skilled memory (see section 1.6).
2 Münsterberg, (1894) in making the President’s address before the New York meeting of the American Psychological Association in 1893 commented, “The experimental study of memory…is as yet begun. The only experiments we have are those of Ebbinghaus…” (p. 34).
3 For example, Müller and Schumann (Angell, 1894), who sought to build on the methods pioneered by Ebbinghaus.
constituent parts. From these early investigations, covering a range of human intellectual abilities, researchers such as Galton (1883), James (reprinted, 1952) and Thorndike (1991) laid the foundations for educational psychology and the promise of merging scientific precision with an understanding of human cognition and learning.

Over the next hundred or so years, educational psychology and our understanding of human memory continued to evolve, along with the way in which we conceived learning and the relationship between learner and teacher. In summarising these changes and the ways in which they have influenced educational practices, Mayer (1992) proposed three metaphors of learning: 1. The first half of the Twentieth Century was characterised by learning as response acquisition, a tradition dominated by behavioural research—especially with animals—which emphasised eliciting the correct behaviour through repetition and rote learning. 2. Through the 1950s and 1960s an underlying shift from behaviourism to mental processing heralded the rise of, and emphasis on, cognition: Learning as knowledge acquisition dominated this period, and by extension teaching was viewed as the direct conduit through which knowledge could be provided and measured. 3. From the 1970s, our understanding of human cognition developed to include the learner as an active participant and the central agent within the learning process: Learning as knowledge construction redirected educators to focus more on how we learn, rather than focussing on what was learned.

Cognitive psychology, with its emphasis on knowledge construction, has “become very important in helping to explain effective learning and its relationship to effective teaching” (McInerney & McInerney, 1998, p. 5). The constructivist paradigm and the many models and theories generated by the study of cognition (e.g., see

---

4 In support of this approach, R. L. Thorndike (1991) commented when writing about his father, that above all he was “an empiricist, a conductor of investigations and an analyser of data” (p. 140).
Conway, 1997) have now become part of the larger field of cognitive psychology.

Bruning, Schraw and Ronning (1999) included the theoretical perspectives and study of human perception, thought, and memory within this field, and they underlined the place of cognitive psychology over recent decades as a major influence on mainstream education. In short, this branch of psychology seeks to understand the nature of human information processing, a widely adopted term and flexible approach to describe cognitive mechanisms (Baddeley, 1997). As Newell and Simon (1972) noted, the individual terms of information and processing were long established concepts, however their coupling as a useful metaphor coinciding with the emergence of digital computers offered new ways of exploring the complex processes underlying human cognition.

Both the models of memory discussed in the following sections, and the empirical work presented in this thesis, are firmly grounded within the information-processing paradigm.

Current understanding of cognitive architecture has considerably strengthened over recent decades and the increasing consensus regarding its structures and processes have helped initiate many new and productive lines of research (Sweller, van Merriënboer, & Paas, 1998). More than any other factor, advances in representing and describing human memory have generated both theoretical and practical outcomes for learning. In part, this is due to the on-going activity of many researchers who now work in the area of instructional design (ID), and who readily seek to empirically validate the application of cognitive theories to authentic learning environments (e.g., Mayer & Simms, 1994; Moreno & Mayer, 2000; Mousavi, Low, & Sweller, 1995; van Merriënboer, Kirschner, & Kester, 2003). In one sense their work is a convergence of

---

5 Neisser’s significant contribution, *Cognitive Psychology* (1967), provides an early example of how the information processing view shaped both the nature of future research and the interpretation of its results.
laboratory-based research (e.g., Conrad & Hull, 1964; Peterson & Peterson, 1959) and what Neisser (1987) emphasised as ecologically valid research\(^6\). These traditions within cognitive psychology have continually altered and adapted to the way in which human memory is conceived and applied to intellectual activity. It is to some of these recent and present conceptions of memory that we now turn.

\(^6\) To distinguish these two paths in the study of memory, Neisser (1982) referred to the exploration of mental mechanisms through well controlled experiments as the *high road*, and investigations of the specific manifestations of memory through ordinary human experience as the *low road*. 
1.2 Human Memory: Models and Conceptions

Memory, from an information-processing perspective, is conceived as a central repository of knowledge, the conscious interface of mental functioning and the essential component in determining an individual’s capacity to learn. This definition is descriptive of the important role memory performs in human cognition, and yet tells us little of what memory is, or how it functions. The development of models that explain the way in which memory supports cognition has engaged numerous researchers since the emergence of the cognitive revolution (Allport, Antonis, & Reynolds, 1972; Atkinson & Shiffrin, 1968, 1971; Baddeley & Hitch, 1974; Craik & Lockhart, 1972; Sweller & Chandler, 1994; Waugh & Norman, 1965). The process of gathering evidence for these models has been gradual and, by design, particulate in the way each new finding contributed to a general and increasingly shared conception of memory. Even so, that general conception has been, and remains, somewhat illusory; this singular term—memory—belying the complex processes and numerous interactions that continue to emerge as characteristic of its deeper structures.

Early processing conceptions of memory (1960s) represented what was then a major debate between unitary and componential models. Unitary models were an outcome of both the sustained emphasis that had been placed on behavioural or stimulus-response psychology (Atkinson & Shiffrin, 1971) and the economical explanation they offered of experimental phenomena up until that point. The key mechanism used to justify a unitary system was based on interference theory (Ceraso, 1967; Melton, 1963)\(^7\), which provided a “mounting body of evidence (cf. Postman, 1964) for the continuity of the principles of interference governing short-term and long-

\(^7\) At this time Melton (1963) states, “the duplexity theory of memory storage must, it seems to me, yield to the evidence favouring a continuum of STM and LTM” (p. 19).
term retention” (Postman & Phillips, 1965, p. 138). Interference effects are well
documented (e.g., Murdock & Carey, 1972; Underwood & Postman, 1960) and
indicated that the memory trace of prior (proactive interference) and subsequent
(retroactive interference) learning will interfere with current recall; the more closely
associated the items to be learned, the greater the interference.

These effects did not exclude the possibility of alternative explanations and a
series of important studies gradually gave impetus to a componential interpretation of
memory. For example, Peterson and Peterson (1959) identified what they termed a
second kind of learning, and concluded that “short-term retention is an important,
though neglected, aspect of the acquisition process” (p. 198). On reflection, by
separately analysing short- and long- recall intervals, the Petersons found an indication
of differential effects that interference alone could not easily explain (Baddeley, 1997).
Similarly, Hebb (1961) established that improvements in the recall performance of
digits were attributable to information from the long-term store: Subjects were unaware
that one particular list of digits was repeated on every third trial, Hebb reasoning that in
the absence of any structural changes, each successive trial would “wipe the slate clean
and set up a new pattern of activities” (p. 41). This repetition however, gradually
improved recall accuracy.

In a later study Glazner and Cunitz (1966) collected data from two experiments
that hypothesised differential short- and long- term storage effects for the serial position
curve in free recall. These predicted effects were identified, the results providing
“further support to the hypothesis of two distinct storage mechanisms” (p. 358). As a
further example, Baddeley (1966a; 1966b) demonstrated encoding differences between

---

8 As pointed out by Baddeley (1997), Postman and Phillip’s (1965) results can also be interpreted as
supporting a two-store model.
short- and long-term memory stores that reflected acoustic and semantic coding respectively.

Additional support for a componential model of memory arose from at least two other areas: The first of these, as mentioned above, was inspired by the computer analogy. Its characteristics of a large storage capacity and means of conscious control were descriptive of the ways in which memory was observed to operate. Specifically, there is a resemblance between the way computers and memory both need to process information before a response can be made (Stern, 1985). The second area of influence arose from neuropsychological research (Banich, 2004): Shallice and Warrington (1970) studied a now famous patient known as K. F. who suffered a profound repetition defect resulting in severe short-term, but not long-term, memory deficits\(^9\). Other neuropsychological findings of brain impairment (e.g., Baddeley & Warrington, 1970) similarly indicated partial deficits rather than the more systematic disruption to memory a unitary model would otherwise suggest.

The research outlined above is indicative of the increasing evidence that eventually coalesced into functionally integrated componential models, the Atkinson and Shiffrin (1968; 1971) modal model the foremost example\(^{10}\). Their model was generally consistent with the major processing models at that time (Eysenck, 1977), most of which functionally divided memory into three primary processing structures: 1. a sensory store in which information rapidly decays; 2. a short-term store in which only limited information can be held for a short period; and 3. a virtually unlimited

---


\(^{10}\) Atkinson and Shiffrin (1971) suggested the notion of a two-component store was acknowledged as far back as the 19th Century by the likes of John Stuart Mill and William James.
long-term store (Figure 1). Characteristic of these structures were the properties of encoding, storing and retrieving information (Murdock, 1972).

![Modal Model of Memory](image.png)

*Figure 1.* Atkinson and Shiffrin’s (1968) Modal Model of Memory (p. 93)

Atkinson and Shiffrin (1971) highlighted the discretionary nature of the control processes and the pivotal role these played in the flow of information into and out of short-term storage. In addition, they stated that the flow of information through this store was limited and transitory by nature, its retention dependent on rehearsal by the individual. Although this new conception of memory helped unify empirical findings, its authors understood that it did not imply a literal mapping from schematic to neurological structures. They did however equate the short-term store with consciousness and, because of its controlling processes, considered it to be a working memory. This notion of an operational- or working memory- centred model was deliberate and represented a retrieval rather than interference model of memory. In support of these structures, Atkinson and Shiffrin demonstrated that probe cues (most likely temporal) in working memory governed the expediency and accuracy of search in long-term storage. In so doing they proposed that long-term storage was permanent and
they rejected interference explanations of memory trace erosion.

With somewhat of a prophetic note, Atkinson and Shiffrin (1971) concluded that their model did not purport to be a complete theory and therefore would undergo the same progressive changes that led to its inception. Indeed, some of the earlier cited neuropsychological studies not only supported a componential model, but also challenged some of its early principles (Baddeley, 1992). For example, the modal model indicated that impairment to working memory should, as the central interchange of information, lead to similar decrements in the long-term store; this was clearly not the case. In a further re-evaluation of the modal model, Baddeley (1997) suggested that it was difficult to reconcile the results of some earlier studies (e.g., Tulving, 1966) with the view that the transfer of information to long-term storage was simply a function of rehearsal in working memory\(^{11}\). That is to say, the evidence was not always consistent with the proposed cause and effect relations indicated by the model.

Although the modal model proved extremely useful as a theoretical stepping stone, it was equally clear that memory would need to be reconceived if it were to accommodate the extent of empirical findings continuing to arise at this time. Influenced by information-processing conceptions, some researchers advocated that memory should be considered in terms of processing levels, rather than as compartmentalised or multistore structures. For instance, Craik and Lockhart (1972) thought multistore models were taken too literally and that a processing approach “provides a conceptual framework—a set of orienting attitudes—within which memory

\(^{11}\) Tulving (1966) found that recall was enhanced only when material was organised into higher-order memory units (see section 1.4.1).
research might proceed” (p. 681)\textsuperscript{12}. This approach proved a useful way of examining how the same stimulus might be processed in different ways and also the effect processing has on recall from, and retention in, long-term memory (Eysenck, 1986).

In a major contribution to the conception of memory, Baddeley and Hitch (1974) pursued the model-based approach and posed the question, “what is short-term memory for” (p. 86)–rather than focussing on what it is–advancing considerable evidence for the functions and processing capacities of a multistore working memory. In this sense, their approach shared some common ground with the likes of Craik and Lockhart (1972). Not only did the Baddeley and Hitch model attract broad agreement, it also offered a productive basis upon which these processes could be applied to learning. As part of a larger conceptual scaffold, the explication of this model (see section 1.4.3) offered, on one level, a way of viewing the interrelated processes within working memory, and on a higher level, understanding the processes between the three major components of sensory, working and long-term memories. Each of these major components is discussed across the following sections in conjunction with current understandings of these processing structures.

\textsuperscript{12} As with Hebb (1961) and Atkinson and Shiffrin (1971), they also made the point that their approach did not constitute a theory of memory. Although all of these researchers made substantial contributions to our understanding of memory, it is clear that there existed a healthy respect for the sheer complexity of generating memorial conceptions!
1.3 Sensory Memory

1.3.1 Processes and meaning.

Considerable research into memory has been dedicated to sensory processes, although a significant proportion of this to what Neisser (1967) described as the iconic (visual) and echoic (auditory) stores. Of these, the visual store has received the greatest attention, although it is assumed that all sensory modalities involve associated mechanisms by which information is briefly captured and then transferred for further processing (Stern, 1985). In relation to these sensory stores, research attempts to address the three areas of how we perceive, recognize, and attend to incoming stimuli (Bruning et al., 1999), areas which also represent a microcosm of memory at large (Baddeley, 1997).

Sensory perception, along with other memory stores, does not conform to a literal translation from external object/s to memorial sensation; other factors can either distort or influence how the outside world is perceived. Any deficit in sensory perception, whether due to either the nature (e.g., foreign vocabulary) or fidelity of the stimulus, may interfere with the way in which it is processed. From this perspective, it becomes readily apparent that perceptual recognition is also reliant on a person’s current knowledge and their capacity to assign appropriate meaning to a percept. Rumelhart (1977) spoke to this issue when he stated, in relation to pattern recognition, that it is a “process of extracting abstract features (such as contours, lines, angles, etc) from the retinal image and comparing these abstracted qualities with those sets of qualities that are known to characterize the items that the observer expects” (p. 53). Recognition is therefore based on a correspondence between these features—or contextual cues—and items already in memory (Massaro, 1970).

Contextual cues highlight both the interconnected nature of our memory systems
and the mediational factors that influence sensory perception (Appendix A.) \(^\text{13}\). To this end, Coltheart, Lea and Thompson (1974) argued that “selective transfer from iconic memory can be executed on the basis of simple physical attributes of the stimuli, but not on the basis of more complex semantic attributes” (pp. 633-634). In addition, Turvey (1973) proposed that the iconic store was characterised by both peripheral and central visual processes, stating that the latter “represents an interface between decisions based on context-independent features and decisions based on context-dependent features” (p. 45).

In total, the evidence indicates that sensory memory should be viewed as neither a rudimentary store, nor for that matter, little more than a simple conduit to working and long-term memories. As a multidimensional construct, sensory memory mirrors the outside world, but only in ways that reflect that world from within.

\(^\text{13}\) Not only is meaning strongly influenced by these cues, once established, it may persist long after a contrary interpretation has been attributed to the stimulus.
1.3.2 Capacities and functions.

An important issue of early research was to establish how much visual information could be captured at the moment a stimulus is perceived. Sperling (1960), who carried out much of the pioneering work in this area, ran a number of experiments to answer this question. He initially flashed a series of letters and numbers on a screen, which were arranged in three lines of four items, and found that subjects, irrespective of exposure time (15-500 ms), recalled on average just over four items.

Rather than accept this initial storage estimate, Sperling (1960) reasoned that his earlier experiment had not excluded the possibility that subjects, between seeing and reporting the image, may have forgotten some of the items. To correct for this, he devised within the same experimental design a partial reporting procedure in which subjects were now cued to recall any one of the three lines of four items presented. Sperling reported that subjects could see on average about nine items, verified by the fact that recall of 3-4 items in each line remained constant regardless of which row was cued. Furthermore, by delaying the time between the image and its reporting, he found that the visual trace had disappeared after approximately 0.5 s. In short, Sperling demonstrated that only limited information is held in the visual sensory store and that it decays rapidly.

Of some importance to this thesis are the mechanisms of auditory processing\(^{14}\), including the auditory equivalent of visual sensory memory. To investigate this issue, Massaro (1970) used a retroactive masking technique in which two rapidly consecutive sounds, the first a target and the second a mask, were produced with varying silent

\(^{14}\) Music materials—including auditory excerpts—were used throughout the empirical work reported in Chapters 5-9.
interitone intervals of 0-500 ms. Subjects had to identify the target sound as either high or low, an otherwise straightforward task without the accompanying masking tone. He established that the first sound was indistinguishable from the second when separated by short intervals of up to about 250 ms. That is, within this interval, the masking tone interfered with perceptual processing. Massaro (1970) suggested that this interval represented the time course or storage limit of auditory sensory memory, and that “this image can be processed, while it lasts, in order to identify the stimulus that produced the image” (p. 417). Moreover, in a related study, Massaro (1974) found similar limits for auditory speech perception (200-250 ms) and concluded, “perceptual units appear to be of roughly syllabic length” (p. 207).

In a replication of Sperling’s earlier work, Darwin, Turvey, & Crowder (1972) applied the same experimental technique to the investigation of the auditory store. Subjects were presented with numbers and letters from three orientations: left ear, right ear and from behind (left, right and combined headphone channels respectively). As previously described, cued recall of partially reported lists was incrementally delayed, in this case from 0-4 s. In line with the earlier visual results, responses became increasingly inaccurate as the delay between stimulus and recall was lengthened. After approximately 4 s, recall was no better whether given with or without cues. With due consideration of all their data, they concluded that the useful span of the auditory sensory store is around 2 s.

In further support of these auditory storage estimates, Glucksberg and Cowen (1970) designed a clever experiment in which subjects were required to shadow prose heard in one ear while hearing unattended prose in the other ear (different headphone channels). Digits were also intermittently embedded in the unattended prose and when subjects saw a green light they were asked to indicate which digit had been heard within
the last 0–20 s. The longer the time between the digit and the green light the greater the decline in performance, which levelled off after approximately 5 s\textsuperscript{15}. In a more recent study, Cowan, Lichty, and Grove (1988) demonstrated that if unattended material is heard while silently reading, the “unattended syllable of speech was found to decay gradually across a 10-sec period” (p. 331). Moreover, memory for vowels was superior to that of consonants, an indicator that an auditory rather than a phonetic form of sensory memory was involved.

From the preceding series of results, it becomes clear that two estimates emerge for the length of auditory sensory memory. In a comprehensive review, Cowan (1984) put a strong case for the existence of both stores. He argued that considerable evidence gathered from sensory persistence, temporal integration and detection masking studies strongly support a short auditory sensory store that ranges from approximately 200–300 ms. Equally, techniques such as modality and suffix effects, dichotic listening, and partial reports provided compelling evidence for a longer auditory sensory memory with a durational capacity of several seconds. Furthermore, Cowan noted that the two memories have distinctive properties, the short store holding only unanalysed information, which cannot exist as a temporal sequence. On the other hand, the recognition stage of auditory perception “outputs a synthesized percept that becomes available to the next stage of auditory information processing” (Massaro, 1972, p. 143).

In total, these studies primarily indicated that sensory memory provides ways of perceiving and holding external stimuli for very short periods of time, and that these periods are longer for auditory storage. As a key development on earlier conceptions, sensory memory moved beyond a simple capture and storage model, and as a result is

\textsuperscript{15} Glucksberg and Cowan (1970) found that memory for “unnoticed verbal material persists for at least 3 sec but no more than 5 sec” (p. 153).
now viewed as the first step in determining the meaning, and therefore significance, of incoming stimuli. In this way, sensory stores act as an initial interface between the environment and deeper cognitive processing in working memory, the mechanisms of which are discussed in the following section.
1.4 Working Memory

1.4.1 Capacities and mechanisms.

The gradual re-designation of the short-term store (STS) to working memory reflected a changing conceptualisation of this underlying cognitive construct. For some, this distinction defined functional differences between the two, although initially these differences were represented by the terms STS and short-term memory (STM). Stern (1985) suggested that STS denoted storage capacity and STM signified the strategies that underscored the interaction between short- and long-term memories. Similarly, Craik and Lockhart (1972) reserved STS for representing a storage system, but used the term STM for the experimental situations in which shorter or transient processing functions were examined. As awareness grew of the crucial role that the STS assumed in controlling cognition, it was increasingly and more consistently viewed as a working memory.

More recent conceptions considered working memory synonymous with consciousness (Sweller, 1999) and mental activity (Chandler & Sweller, 1996); a “place where information is processed for meaning” (Bruning et al., 1999, p. 36) and which includes both storage and active processing capacities (Eysenck, 1986). In many respects, distinguishing between these capacities represents the evolution of experimental tasks that has led to our current understanding of this important mental construct: Early research, using assorted recall protocols, quickly established both various storage limitations and the means by which information is held (see section 1.4.2). In contrast, the focus on active processing (e.g., reasoning and comprehension)

---

16 Refer to the Atkinson and Shiffrin (1971) modal model (section 1.2) in which the STS is also sub-headed Temporary Working Memory. The term working memory will generally be adopted throughout this section rather than continually using different designations.
emerged a little later, although linking these mechanisms within an operational working memory was initially more difficult to establish (Baddeley & Hitch, 1974).

Miller (1956), in a now celebrated study—*The Magical Number Seven, Plus or Minus Two*—attempted to establish some of the limits on our capacity to hold and recall basic information. He referred to this feature of immediate memory as *the span of absolute judgment* and stated that it is “usually somewhere in the neighbourhood of seven” (p. 90). Miller showed that accuracy of recall decreased as further items were added, and argued that this was also consistent with multi-dimensional data, which likewise indicated accuracy declined if more than one item is attended to simultaneously. He further demonstrated that the amount of information held for immediate recall is a function of the amount of information represented by each item\(^\text{17}\), a process that is known as chunking (see section 1.4.2).

Not only were early investigations directed towards the quantitative capacity of working memory (number of items), but also the amount of information that could be held as a function of time. To explore this issue, early free recall tasks presented subjects with a series of items to be memorised, immediately followed by a *distracter task*\(^\text{18}\); recall accuracy was then measured after predetermined lengths of time. Typically, results indicated that recall failed somewhere between 6-18 s. From such tasks, Peterson and Peterson (1959) characterised forgetting as a gradually decaying trace, a prevailing view at this time (Greene & Hicks, 1984).

Murdoch (1961) produced comparable results to those of the Petersons, but in a

\(^{17}\) The nature of items is an issue taken up in the next section and is of fundamental importance to cognitive load theory (see section 3.2.1).

\(^{18}\) These tasks were usually interpolated items, e.g. counting backwards by threes or fours. They denied the use of rehearsal, a strategy for maintaining information in working memory.
further experiment altered the rate of the interpolated activity (distractors) in order to study its effect on retention. He concluded, “Since in Exp. III it was the number of interpolated items and not their rate of presentation that was the critical variable, these results are more compatible with an interference theory than with a decay theory” (p. 624)\(^{19}\). When all the data from these early experiments are considered, they reveal that both time and quantity impose a heavy limitation on working memory.

Taking an alternative approach to measuring capacity, Sternberg (1966) asked subjects to confirm whether a target item belonged to a larger set of similar items (e.g., digits). By plotting set size against reaction time he was able to calculate that for each additional item it took 38 ms to respond. In these same experiments, Sternberg also explored whether scanning for the target item was either serial or parallel and, once the target item was confirmed, whether search was self-terminating or exhaustive. Response times increased as a function of set size irrespective of either the target item’s position in the series or whether it was a negative or positive response. He concluded that processing was serial at an average rate of 25-30 symbols per second (also see Sternberg, 1969), and that all options were searched (exhausted) before the subjects responded.

Although alternative explanations have been offered for Sternberg’s results, they have proved broadly replicable (Baddeley, 1997) and they have provided further insights into the mechanisms constraining working memory. In support of these findings, Newell and Simon (1972) also claimed that information processing is unequivocally serial, however they added that this does not exclude either an awareness of multiple items, or the existence of a single process containing various amounts of

\(^{19}\) The same number of interpolated items was used over a time range of three or four to one.
information.

The early attempts to measure capacity were based upon the notion of a discrete store. Recall however, that Hebb’s (1961) experiment (see section 1.2) found gradual improvements in working memory performance that were attributed to long-term memory. Moreover, plotting serial position recall of lists usually rendered a U-shaped curve, demonstrating the tendency of subjects to remember both earlier (primacy effect) and more recent (recency effect) items in a list (e.g., Postman & Phillips, 1965); again, the former effect was assumed to be associated with long-term memory. Undoubtedly, studies such as these suggested that working memory capacity should be considered in relation to other memory mechanisms.

Evidence of interactions between memory stores prompted researchers to move beyond digit span as a measure of capacity, and focus more on tasks that combined both storage and active processing of materials. Under these conditions, Baddeley and Hitch (1974) found that working memory loads of more than three items interfered with cognitive tasks, and when loads exceeded six items, sizeable disruptions occurred. Notably, they located these limitations within a limited capacity workspace, their data indicating a trade-off between the allocation of capacity to either storage or processing. These findings were also supported by more recent research in the area of instructional design: Sweller, et al. (1998) stated that when working memory is “used to process information in the sense of organizing, contrasting, comparing, or working on that information in some manner, humans are probably only able to deal with two or three items of information simultaneously when required to process rather than merely hold information” (p. 252).

20 The digit span procedure usually involved gradually increasing the number of items to be recalled until memory consistently failed (ca. 50% criterion).
In summary, where information is held without rehearsal, limitations on working memory are both quantitative (5-9 independent items), and temporal (2-3 s). Once capacity is exceeded, the information is either displaced or forgotten. Available capacity is a trade-off between the number of discrete items in storage, and/or whether these items require active processing. Active processing—the interaction of as few as two or three items—places further limitations on working memory. Given these severe restrictions, the definition of what constitutes an item becomes pivotal to understanding human memory and therefore to understanding intellectual performance and learning. The nature of an item is discussed next.
1.4.2 The nature of items (elements).

The true significance of the aforementioned study by George Miller (1956) was establishing the nature of an item rather than item capacity per se. He tested whether the capacity for bits of information (e.g., letters) that comprised an item or chunk (e.g., words or phrase) was invariant, or whether the number of chunks, irrespective of bits, remained constant. He indeed confirmed that the “span of immediate memory seems to be almost independent of the number of bits per chunk” (p. 93). As a result, Miller made a distinction between the number of chunks (absolute judgment; ca. 7), and the variable amount of information per chunk that could be retained in working memory.

Miller (1956) categorised a chunk as something familiar, which in turn was based on considerable learning. More recent definitions viewed a chunk as a “collection of elements having strong associations with one another, but weak associations with elements within other chunks” (Gobet, Lane, Croker, Cheng, Jones, Oliver, & Pine, 2001, p. 236)\(^{21}\). Considerable work on chunking has taken place since Miller’s groundbreaking work and this strategy is now assumed to be of fundamental importance in the learning process (McInerney & McInerney, 1998). This was cleverly illustrated by Bower and Springston (1970) who presented subjects with letter strings, pausing briefly between each\(^{22}\): It was thought that these pauses determined whether meaningful recoding of familiar units took place. The first series contained groups of three random letters, and the second series familiar combinations such as IBM, CIA etc. Recall of the familiar series was superior, the difference accounted for by the requirement to learn unrelated and therefore multiple element groups in one series, and only one element per

\(^{21}\) The term element is fundamental to cognitive load theory and represents, in relation to a learner, a meaningful unit or item (see section 3.1).

\(^{22}\) These strings were ordered in digrams, trigrams or quadrigrams.
group (meaningful chunk) in the other.

In an effort to better understand the nature of a chunk, Simon (1974), using a parameter-estimating paradigm, “explored some of the interactions between research on higher mental processes over the past decade or two, and laboratory experiments on simpler cognitive processes” (p. 487). His results indicated that none of the measures were as constant as Miller’s (1956) original estimate might suggest: A more consistent number of chunks emerged when materials (word phrases) were reconceived in relation to meaningful units (chunks of varying size). Simon readjusted the capacity of working memory to five chunks, and suggested variations in this capacity for longer word phrases were due to rehearsal difficulty.

Zhang and Simon (1985) further refined the notion of chunking through a series of experiments that compared the use of Chinese and English language materials (logographic and alphabetic respectively). The different bases upon which language symbols are formed allowed them to test whether memory was limited by either a maximum number of chunks or the maximum number of syllables that could be spoken in a given time (approximately 2 s) 23. They concluded that working memory capacity “is constant neither in terms of chunks alone nor in terms of syllables alone, but can be expressed in terms of a weighted sum of chunks and syllables” (p. 200) 24. Moreover, as chunk size increased (syllables per word), memory span for chunks and syllables (bits) decreased and increased respectively, providing further support for the notion of a flexible store mediated by its processing operations.

23 Chinese logographs represent whole words or ideas, whereas alphabetic materials are formed by their constituent sounds or syllables.

24 Their findings were also interpreted in relation to “Baddeley’s hypothesis of an articulatory loop [see section 1.4.3], with a fixed duration of 2 or 3 sec.” (p. 200).
Considerable evidence suggests that chunking is a mechanism that both provides ways of circumventing limited processing capacity, and which uses what we already know to reorganise, and by so doing minimise, the number of elements that need to be held in working memory\textsuperscript{25}. In this way, a large number of otherwise unfamiliar elements can be treated as a single chunk and as a result the capacity limitations of working memory apply only to novel information\textsuperscript{26}. Chunking is but one mechanism by which information is efficiently encoded, stored and retrieved through the conscious processing of working memory. In order to better explain and understand these processes, a functional model of working memory was developed from the early modal conceptions: These developments are taken up below.

\textsuperscript{25} Section 1.6 examines chunking in relation to human expertise, a connection that has been comprehensively investigated through the game of chess.

\textsuperscript{26} See section 3.2 for a discussion within the framework of cognitive load theory regarding the relationship between familiar and novel elements.
1.4.3 A component model of working memory.

Baddeley and Hitch (1974) provided strong evidence that concurrent digit span impaired performance, but not to the extent predicted by a unitary model of working memory\textsuperscript{27}. Similarly, it was found that concurrent digit span left performance generally unaffected during retrieval from long-term memory (Baddeley, Lewis, Eldridge, & Thompson, 1984a) and had little impact on recency (Baddeley & Hitch, 1977). To account for these findings, and the apparent partial independence of processing functions, Baddeley and Hitch (1974) proposed an alternative working memory model. This model assumed a number of slave- or sub- systems, and it quickly emerged as the dominant theoretical approach (Gathercole, 1997).

Baddeley and Hitch (1974) and Baddeley (1997) proposed at least three major processing components (Figure 2): 1. The *central executive*, a controlling attentional system that is responsible for active processing and limited transient storage. 2. The *phonological loop*, which handles speech-based information. 3. The *visuo-spatial scratchpad*, which deals with visual and spatial information\textsuperscript{28}.

![Figure 2](image_url)  
*A simplified representation of the working memory model*; taken from Baddeley (1997, p. 52).

Most of the evidence for this model highlights the operations of the

\textsuperscript{27} Concurrent tasks were assumed to exhaust storage capacity.

\textsuperscript{28} The visual and auditory components will be reviewed below. The central executive is the component about which least is known (Eysenck, 1986), and will not be examined in this literature review.
phonological loop, in part because of the interest and research generated by speech-based cognition (Baddeley, 1997), and because acoustic coding, including the recoding of visual stimuli, was found to be a principal feature of working memory (Baddeley, 1966b; Conrad, 1964; Conrad & Hull, 1964). Schumacher et al (1996), using positron emission tomography (PET) measures, provided “strong support for a verbal working memory system that is indifferent to the modality of input” (p. 88). This assumption of a phonological loop, comprising a store for speech-based information and an articulatory loop for verbal rehearsal, accounted for our capacity to briefly retain verbal information (approximately 2 s) and to maintain it by ‘inner’ rehearsal.

Together, these mechanisms parsimoniously explained the way in which we process verbal information, including a range of effects such as articulatory suppression: the decline in recall of a word series when a subject repeats an irrelevant syllable aloud during presentation (see for example, Wilding & Mohindra, 1980). This is a robust effect that has been linked to the inability of subjects to access the articulatory loop and therefore maintain a rehearsed item (words) in the phonological store. Neuropsychological evidence offers additional support for these processing functions, one imaging study (PET) indicating that within “the circuitry for verbal working memory, there is a separation between components that underlie storage and those that mediate rehearsal” (Smith & Jonides, 1997, p. 38).

The phonological loop also provides mechanisms that help explain the word length effect: the deterioration of memory when recalling words of longer articulatory duration (Baddeley, 1997). The vocal or sub-vocal rehearsal needed for longer words allows fewer items to be rehearsed and therefore to be maintained in storage. When subjects are denied access to rehearsal (articulatory suppression), the word length effect disappears (Baddeley, Lewis, & Vallar, 1984b). Moreover, by comparing two-syllable
words spoken at either a naturally faster or slower pace, Baddeley, Thomson, and Buchanan (1975) demonstrated that the time necessary for articulatory rehearsal was the key determinant of recall.

Likewise, the Baddeley and Hitch (1974) model provided an explanation for the *phonological similarity effect* (Conrad, 1964), the difficulty of remembering a series of phonetically similar items (e.g., B C P T V; F M N; S X). Because the model assumes that the store uses phonological coding, discriminating between similar sounding (and coded) items is more difficult. Evidence for this view was again provided under conditions of articulatory suppression, Baddeley, Lewis and Vallar (1984b) commenting,

> When material is presented auditorially, then it is automatically registered in this store, regardless of whether or not the subject is allowed to articulate. With visually presented material, however, the store can be used only if the subject is allowed to articulate the items; hence suppression during visual presentation removes the similarity effect (p. 242).

It becomes clear that working memory provides an important means by which phonological coding of verbal information (visual and auditory) can be retained through continual rehearsal. Equally, this same mechanism is subject to either interruption or failure if stimuli compete within the same processes. This initial emphasis of research on acoustic coding somewhat masked what also seemed intuitively plausible: working memory codes information in different ways and according to different sensory mechanisms. The evidence for modality-based coding and the control of these mechanisms is presented in the following discussion.
1.4.4 Modality-based processing.

As reported earlier (see section 1.3), modality-based sensory stores initially receive and select visual and auditory information for transfer to working memory. By extension, “Various types of store may be distinguished in STS [working memory] according to the sensory mode through which the items are presented” (Glanzer, 1972, p. 131). Nonetheless, it was at first uncertain in which ways visual information was coded and whether, as Baddeley and Hitch (1974) contended, visual and acoustic (auditory) stores represented separate slave systems.

Evidence for separate modality-based processes initially arose from the finding that the “retention of verbal material presented auditorially is better than retention of the same material presented visually” (Murdock, 1972, p. 81)\(^{29}\), a phenomenon identified as the *modality effect* (Craik, 1969; Murdock, 1967). In addition, differential recency effects indicated, “that auditory short-term storage is more effective and lasts longer than visual [storage]” (Corballis, 1966, p. 50). At this time, these effects did not necessarily lead to the conception of separate working memory subsystems, although Crowder and Morton (1969) suggested they could be explained by precategorical acoustic storage (PAS). Their system was more perceptual in nature however, and as Penney (1989) laments, it “relegated modality effects to a peripheral place in memory” (p. 398).

A more recent approach taken by Zhang and Simon (1985) was to compare coding characteristics (homophonic and non-homophonic) of Chinese language-based materials. It was assumed when using Chinese characters either phonemically alike (similarity effect) or without familiar names (radicals) that subjects would be unable to

\(^{29}\) Although Murdoch (1972) also adds, “the effect is restricted to the recency part of the curve” (p. 81). The modality effect, as conceived within cognitive load theory, is covered in detail in section 3.3.5.
rely on acoustical encoding. They established that whereas storage for non-homophonic characters was approximately seven chunks, recall for homophonic characters (similar) and non-acoustical material (radicals) appeared to be limited to two or three chunks. Across their experiments, the evidence pointed to a visual or semantic store of limited capacity.

In an earlier investigation, Brooks (1967) undertook four experiments in which a potential conflict between two sources of visual material (e.g., reading and visualisation of spatial relations) was compared to the same materials alternatively delivered via both auditory and visual modalities. In all their experiments the expected conflict between two forms of visual coding was confirmed, indicating that both visual sources competed for the same processing capacity. In contrast, this conflict was absent for the groups who were able to process the information between auditory and visual modalities. From these results, Brooks tentatively concluded, “the internal representation of spatial material uses mechanisms specialized for visual perception” (p. 298).

Allport et al. (1972), in their study, *A Disproof of the Single Channel Hypothesis*, used secondary auditory shadowing tasks that, in combination with primary materials, were designed to exhaust working memory capacity. For their first experiment, primary materials were presented as either words (auditorially or visually) or pictures. As would be predicted (at this time) by a unitary storage system, these tasks should have interfered with performance irrespective of task modality. Contrary to a unitary view, recall was more accurate for visual-auditory presentation than auditory

---

30 This evidence also cast doubt on the claim that Chinese readers encode ideographic characters “directly from visual to semantic form, without going through an intermediary acoustic coding” (p. 196).
31 The same design also tested these effects using nonsense messages to exclude the possibility that reading was an inherently more difficult activity (also see Mousavi et al., 1995, section 3.3.5).
32 The subjects also reported that they had “pictured” the spatial relations.
presentation alone. Their second experiment required subjects to sight read piano music with or without a concurrent auditory shadowing task of either easy or difficult prose passages. Allport et al. found no significant differences between the sessions (with or without divided attention) for either the comprehension of the prose passages, or musically for wrong notes or timing errors. Since the addition of shadowing tasks (auditory) was assumed to consume significant working memory resources, the ability to concurrently sight-read music (visual) was believed to strongly indicate the presence of modality-based processing\textsuperscript{33}.

From the findings above, the conclusion of dual-channel processing is inferred either by way of differential processing demands (Brooks, 1967, 1970), or the differential interference created by competing tasks (Allport et al., 1972). A third approach was to compare capacities more directly. To achieve this, Frick (1984) presented digits (memory span) either visually, auditorially, or for a third group, as a combination of both modes\textsuperscript{34}. He found mixed-mode presentation (visual-auditory) was superior to all other conditions, but only when the auditory material was recalled first (inverted). Based upon a dual-storage explanation, Frick suggested that modality-specific interference might explain the apparent difference between presentations (either normal order or inverted); it was assumed that auditory retention followed by speaking (recall) disrupted auditory memory. Frick concluded that the capacities of the two stores are not additive, but they “are consistent with the hypothesis of independent auditory and visual short-term stores” (p. 514).

\textsuperscript{33} Also of interest, Allport et al. (1972) reported, the “more unpredictable (to the subject) the demands of a particular task…the less the probability that multi-channel or distributed processing of other tasks, even though possible, will be permitted” (pp. 233-234).

\textsuperscript{34} Mixed-mode subjects were “asked to form a picture of the visually presented digits and to remember the sound of the auditorially presented digits” (Frick, 1984, pp. 508-509).
All the studies reported above reinforce the conception of modality-based processing\textsuperscript{35}. In addition to this behavioural research, neuropsychological evidence also “suggests the existence of multiple working memory capacities, each intimately tied to the operation of specific information-processing systems in the brain” (Banich, 2004, p. 344). For example, Warrington and Shallice (1969; 1972) investigated the issues of selective impairments to both auditory verbal and visual short-term memory capacities and found evidence for separate processing mechanisms. Not only do these results reinforce the existence of modality-specific stores, they also offer a confidence in the data that is not possible from either behavioural or neuropsychological evidence alone.

\textsuperscript{35} See also Penney (1989) for a comprehensive review of modality-based research.
1.5 Long Term Memory

1.5.1 Structures and models.

Despite many advances in cognitive psychology, the exact nature and representation of long-term storage structures remains elusive. Tulving (2000) argued that even though the generation of memorial concepts has fuelled much of the debate in relation to the structure of human memory, a greater emphasis should be placed on conceptual analysis as a way of ensuring clear and universal meaning of terms. However, of little dispute is the immense, or for practical purposes, unlimited capacity of long-term memory (Reed, 1992) and the fact that ‘permanent’ storage usually requires from the learner some effort over variable periods of time (Bruning et al., 1999).

Over the past forty years many characteristics of long-term memory (LTM) have been identified, which in turn have helped describe a number of broadly accepted structures. Whether these structures support either a descriptive nomenclature or identify real anatomical distinctions remains an area of conjecture. As Crowder (1985) points out, although spatial representations (stores) initially dominated the way in which memory was conceived, non-spatial models have also accommodated the many findings underpinning the structures of human memory. These non-spatial models have focussed on perceptual activity and the LTM processes of encoding and retention (Craik & Tulving, 1975), a focus inherent in the likes of the Craik and Lockhart (1972) and Anderson (1976; 1996) models of memory.

---

36 Tulving (2000) suggested that this emphasis should be redirected to ‘what’ questions, which “are at least as important as ‘how’ and ‘why’ questions” (p. 42).
Intrinsic to Anderson’s (1976; 1996) *ACT* model\(^{37}\) was a fundamental distinction between *declarative* and *procedural* forms of *knowledge*, a representational difference that remains widely accepted today. Anderson (1996) defined declarative knowledge as a propositional network that represents facts and their relations to other facts. In contrast, he saw procedural knowledge as productions, which describe the way in which facts are used (skills). Farr (1987) remarked that the distinction between facts and skills is also supported by neuroscientific evidence, although as previously noted, it remains uncertain whether neural activations represent separate physical structures (Eichenbaum, 1997). To explain the interaction between these major processes, Anderson also included a transitional phase of *knowledge compilation*, which describes the process by which skills are developed.

Substantial research has been dedicated to declarative memory, especially in relation to Tulving’s (1984a) formalised demarcation between *episodic* and *semantic memories*. Episodic memory deals with autobiographical events (personal referents) and their temporal-spatial relations; an event can be stored “solely in terms of its perceptible properties” (p. 385). Alternatively, semantic memory stores elements of general knowledge (cognitive referents), but without accompanying references to when and where memories were initially stored. Even though important differences have been established between the storage and processing mechanisms of these memory systems (Shoben, 1984), Tulving used these categories only for descriptive convenience “rather than as an expression of any profound belief about structural or functional separation of the two” (Tulving, 1984a, p. 384). Nonetheless, along with the characteristics already mentioned, he identified some important operational differences between these two

\(^{37}\) *Adaptive Control of Thought*: Anderson (1996) embraced the notion that the mind is unitary, which “holds that all higher-level cognitive functions can be explained by one set of principles” (p. 2).
stores:

i. Semantic information is usually encoded as a rich network of relations between items, and it is not reliant on temporal cues for accurate recall.

ii. Episodic information is more vulnerable to forgetting, whereas the elaborate integration of semantic information into existing knowledge structures offers some immunity.

iii. Semantic memory is flexible and is associated with inferential reasoning and generalisation, qualities not available to episodic memory. Therefore, semantic memories can be associated with or adapted to an almost limitless range of other items.

Initially, the weight of research effort favoured episodic memory, a legacy of the earliest experimental paradigms that emphasised recall tasks such as paired-associates and serial learning (Bower, 2000). Recent decades have seen a greater investment of research activity into semantic memory, a by-product of more naturalistic investigations employing advanced multivariate statistical techniques (Bahrick, 2000), and a trend associated with the development of global structural models of memory (Leahey & Harris, 1997). Shoben (1984) commented that the episodic-semantic distinction helps to refine theoretical models, although it “should be emphasized that…[they] are not two fully independent systems that never communicate with each other” (p. 216). Take for example the performance and the observation of music: Even though the knowledge gained from either activity might be recalled from different sub-systems, there is little doubt that semantic knowledge will be strongly influenced by personal experience.

Tulving (1985) not only acknowledged the interrelated nature of the two memory stores, in a later hierarchical rearrangement he proposed that procedural memory subsumed semantic memory, which in turn subsumed episodic memory. This
expanded conception helped allay some of the detailed criticisms put forward by McKoon, Ratcliff, & Dell (1986), including a concern that the qualitative differences between the two stores are not distinctive enough to be empirically tested. Nonetheless, McKoon et al. asserted that the data also supported alternative views that account for descriptive phenomena, and the episodic-semantic distinction did not provide either an appropriate model or theoretical explanation. Taken together, they believed that this distinction is based upon weak experimental and neuropsychological evidence and the dissociations used to explain differences are either too inconsistent or could be explained by either process or single-store models.

Neuropsychological evidence does provide some support for a unitary conception of episodic and semantic memories. According to Eichenbaum (1997), physiological data and computational models indicate that the hippocampus may perform a mediating function that differentiates between different types of memorial associations. Moreover, he concluded,

> The data… indicate that the hippocampus may also interleave patterns within the memory network so as to provide access to the whole knowledge structure from any point. Within this scheme, episodic and declarative memory are not alternative types of memory, but rather are two powerful benefits of the networking of cortical memories supported by the hippocampus (p. 331).

Evidence, also emerging from cognitive neuroscience, has suggested “that most memory tasks likely engage multiple memory systems, in a similar way to the partially independent brain pathways that are engaged in eye-movement control and attention shifts” (Johnson, 1997, p. 136). Which of these (or other) perspectives exactly

---

38 Tulving (1984b) later stated, the “1972 distinction was inchoate: rudimentary, imperfect, incomplete and somewhat disorganized” (p. 224). He did, however, still believe that the two systems were functionally distinct.
represents the mechanisms of LTM storage is a question that requires further investigation.

An additional demarcation within LTM corresponds to *explicit* and *implicit* memory processes. Explicit memory represents conscious recollection, whereas implicit memory attempts to measure unconscious influences; the distinction has “traditionally relied on findings of dissociations between direct and indirect tests” (Toth, Reingold, & Jacoby, 1994, p. 300). “Implicit memory phenomena have been widely demonstrated in both normal and amnesiac subjects” (Parkin, 1989, p. 231), although separating implicit effects from conscious recollection (explicit) is more difficult in practice than it first appears (Eysenck & Keane, 1995).

The dissociation between multiple memory systems refers “to the notion that there are qualitatively different memorial consequences of an experience that are mediated by processes that can, but may not always, function independently of one another” (Schacter, 1985, p. 352). In a major review of implicit memory research, Schacter (1987) remarked that the differences between these memory systems are often ascribed “to the different properties of hypothesized underlying systems” (p. 511), citing the declarative-procedural and semantic-episodic sub-systems discussed above.

For example, Cohen (1985) differentiated between explicit remembering and skill acquisition, a distinction “distinguishable on representational and processing grounds” (p. 421), and which reflects the descriptive differences between the declarative and procedural systems respectively39. In line with this view, there is some evidence that the implicit-explicit divide also corresponds to differences between automatic and effortful processes (Parkin, 1989).

---

39 Likewise, semantic and episodic memories could be interpreted as representing the dissociations between implicit and explicit memories respectively (Schacter, 1987).
In summary, although considerable progress has been made in relation to a functional understanding of LTM, the exact nature of its conceptual or structural basis– processes and/or multiple memory stores–is yet to be fully determined. Nevertheless, the empirical data have provided many details relating to the means by which memories are encoded and retained, issues that are reviewed in the next section.
1.5.2 Encoding and retention.

Any discussion of LTM encoding and retention is deficient if not concurrently considering the activity of learning. For learning to occur, current or familiar knowledge and skills (LTM) provide a scaffold upon which new knowledge elements can be meaningfully processed in working memory (Gagne, Bell, Yarbrough, & Weidemann, 1985, see section 1.4). Indeed, Stern (1985) suggested that the flow of information firstly connects perceptual information to LTM and then to the short-term store. Together, previously learned elements and newly acquired elements are retrieved from, and retained in, LTM respectively. In this way, a large knowledge base–accumulated elements in LTM–supports higher levels of intellectual performance (Sweller, 1993). Consequently, a consideration of encoding and retention is central to the aims and efficiencies of instructional design.

Underlying comprehensive investigations of storage and retention is the universal assumption that meaning is derived from the relations that exist between items in memory (Greeno, 1974), although the manner in which these relations are manifest remains open to fundamentally different interpretations (e.g., Collins & Loftus, 1975; Smith, Shoben, & Rips, 1974). Given these relations, the importance of encoding memories as elaborated concepts (e.g., series, hierarchy, network) or in meaningful ways (Underwood, 1964) is crucial for later retention, and basic to any understanding of LTM.

Analogies generated by process models (information processing) sought to better understand the complex interrelated functions that are descriptive of human memory (e.g., Rumelhart, Lindsay, & Norman, 1972). Arising from this approach, connectionist

40 One of these conceptions–schemas–will be discussed in the following section.
conceptions of LTM have endeavoured to model a more realistic simulation of brain activity, embracing for example the notion of parallel distributed processing (PDP, see Rumelhart, Hinton, & McClelland, 1986; Rumelhart & Todd, 1993). Typical of these models, concepts are represented by nodes and encoded with their relations to other nodes\textsuperscript{41}. Rumelhart et al. (1972) took the position that adults rely more on complex functions such as understanding, problem solving and logical deductions than they do on rote memorisation; as a result these models help to explain how these higher mental operations are applied within everyday life.

Whether the emphasis of encoding (and therefore the recall of) new materials is better placed on either the structure of the stimulus or the processing activities themselves was a key element in relation to the processing models versus multistore debate. Craik & Lockhart (1972) argued for a levels of processing approach where depth of processing provided a better explanation of effective encoding than the overly literal interpretations of multistore models. They stated, “Highly familiar, meaningful stimuli are compatible, by definition, with existing cognitive structures” (p. 676); therefore, these stimuli will be processed more deeply and better retained. Craik & Lockhart viewed processing levels as a continuum of analysis or the degree of stimulus elaboration required by the orienting task.

Hunt, Elliott, & Spence (1979) investigated directly whether the better determinant of effective encoding was either the structure or processing of the orienting task. They employed conditions in which associative meaningfulness was varied for both semantic and nonsemantic processing, and in summary stated,

\textsuperscript{41} Relations may be quite simple or, in many cases, incorporate a rich network of associations that reflect substantial experience and learning.
The results of these studies, as well as previous research involving nonsemantic structural manipulations, unequivocally question any assumption of exclusive control of encoding by orienting tasks and the subsequent inference that a single processing principle provides an adequate conceptualisation of encoding (p. 347).

With respect to encoding, these findings recognised that an interaction between process and structure is more likely to improve the durability of representations in LTM.

Craik and Tulving (1975) also pointed to weaknesses in the earlier Craik and Lockhart (1972) model and preferred, in relation to encoding, either the term *elaboration* or the metaphor spread rather than the vagueness implied by depth of processing. Across a number of experiments, Craik and Tulving (1975) found that performance was dramatically affected when the mental activity of the learner was manipulated, even where other strong determinants such as study time, number of repetitions, intention to learn, recency and associative strength of items were held constant. They strongly proposed, “it is the qualitative nature of the task, the kind of operations carried out on the items, that determines retention” (p. 290)\(^ {42} \). These results clearly reinforce the importance of elaborative encoding that concentrates on deeper rather than superficial learning activities.

A prime illustration of these memory principles—deeper processing and retention—results from students generating meaning rather than simply reading explanations, a phenomenon known as the *generation effect* (Toth & Reed Hunt, 1990). Begg, Snider, Foley and Goddard (1989) remarked, “The benefit of generation is real; it is not an artifact. We have seen dozens of cases in which generated items exceeded items read in pure lists, without a cost for items read in mixed lists” (p. 988). They also suggested that

\(^ {42} \) They added that a minimal semantic analysis is superior to an elaborate structural analysis, and that descriptive coding contexts result in better retention.
the generation effect is not special; any encoding process that helps to distinguish items in context should enhance later retention of those items within the same context.

Encoding is not only dependent on the nature of the task, but also the sense modalities through which task stimuli are received. Schab (1990) demonstrated that odours could also operate as effective memory cues when received at the time of both learning and testing. These experiments were based on Tulving and Thomson’s (1973) *encoding specificity principle*, whereby congruent properties of encoding and recall cues lead to more effective recall. This principle appears to hold for not only the context and sense modality of materials, but also for the subject’s psychological state or mood at the time of encoding and retention (Bower, 1981)\(^\text{43}\). It also explains some aspects of forgetting, where retrieval cues are matched to incorrectly substituted, but otherwise assciatable, information (Bjork, 1978).

Information can be also auditorially encoded for storage in LTM (Baddeley, 1997). For instance, when visual or verbal confirmation is unavailable, we intuitively call on auditory LTM in order to recognise a familiar voice. Meudell, Northen, Snowden and Neary (1980) tested this auditory capacity as part of their study investigating retrograde amnesia. They asked both a control group and subjects with alcoholic amnesia to recognise recordings of famous voices made over the previous 50 years. Even though the control subjects had not heard many of these famous voices for decades, recognition was surprisingly good and remarkably consistent, irrespective of the target decade\(^\text{44}\).

\(^{43}\) Bower (1981) referred to the association between memory and mood as the *mood-congruity effect*, “which means that people attend to and learn more about events that match their emotional state” (p. 147).

\(^{44}\) Performance for the amnesiacs was poorer and it deteriorated for more recent decades.
Paivio (1986) recognised the inherent flexibility of the human processing system as a cornerstone of his *dual-coding theory*, in which a major distinction was made between the functionally independent representation and processing of verbal and nonverbal materials. He believed that the verbal system is adapted for language, and nonverbal processing is conceived as an imagery system that incorporates different sensory modalities. Both systems are functionally distinct and can store stable and discrete long-term memories, however under certain conditions referential connections provide an enhanced capacity to code memories in both systems (Figure 3).

![Dual-Coding Model (Paivio, 1986, p.67).](image)

The use of verbal and nonverbal materials is commonplace in most instructional settings and therefore the idea that combining language and imagery (picture, sound event etc) may improve encoding and retention is of significant import to learning. Paivio (1986) claimed that imaginal coding is essentially stronger in value than verbal
coding and therefore pictures, for example, are a potent aid when combined with verbal information. According to Paivio, imaginal-verbal associations were evident for concrete rather than abstract words, a distinction illustrated by concepts such as “bat” and “honour” respectively.

Paivio’s (1986) ideas are supported by evidence in which the recall of pictures is easier and more accurate than words. Standing (1973) compared recall for either words, snapshot pictures (e.g., a dog) or vivid pictures (e.g., a dog with a pipe in its mouth): It was found that pictorial coding was superior to verbal coding, and the vivid pictures were remembered more accurately than the snapshots. Larkin & Simon (1987) also believed that there were computational advantages when using diagrams, “not because they contain more information, but because the indexing of this information can support extremely useful and efficient computational processes” (p. 99): In essence, representations are situated in a plane, avoiding the processing cost of search when otherwise seeking appropriate relations between adjacent or removed elements.

Many advertisers understand the importance of fully integrating product name (words) and pictures to improve brand recognition. Although pictures may be remembered better than words, simply placing a brand name next to an associated image is not necessarily the best way of ensuring product retention. Lutz and Lutz, (1977) tested this principle by randomly selecting 48 brands from the phone directory and separating them into interactive and noninteractive image-name pairs: Interactive

45 Paivio (1986) suggested “that imaginal and verbal codes are unequal in mnemonic value, perhaps by a 2:1 ratio favouring the image code” (p. 77).
46 What they referred to as a sentential representation of information. These principles assume considerable importance for the instructional designs generated by cognitive load theory.
47 “An interactive image integrates the two items in some mutual or reciprocal action” (Lutz & Lutz, 1977, p. 493).
brands facilitated superior recall, “presumably by increasing the concreteness of the material to be learned, that is, the association of the two items” (p. 497). Furthermore, brands that facilitated stronger images—humorous or unusual—were better remembered than the more functional or prosaic examples.

These encoding principles are often recommended in popular books that espouse the benefits of mnemonic techniques for improving memory (e.g., Lorayne, 1985; O'Brien, 2000). Mnemonics are a deliberate attempt to reorganise memories into meaningful and structured units, usually by attaching new information to memorial structures that already exist in LTM (Leahey & Harris, 1997). Wagner (1978) associates mnemonics with control processes and suggested that although physiological changes in the brain reach asymptote at around 4 or 5 years of age, the observed improvement in memory up until adolescents may be due to the application of these processes.

Mnemonics can be quite effective when attempting to memorise otherwise disconnected facts (Harris, 1978). Equally, these techniques may be ineffective outside quite restricted contexts (Morris, 1978), although Higbee (1978) argued that some perceived restrictions are better viewed as pseudo- rather than valid limitations. Higbee believed that mnemonics are not a substitute for learning principles, but instead they use these very same principles as a consequence of incorporating conventional cognitive strategies: “association, organization, meaningfulness, attention, retrieval cues, deep encoding, and visual imagery” (p. 152). As one example, the keyword method incorporates the use of verbal-visual coding in a way that ties the to-be-remembered

---

48 Harris (1978) argued that the importance of external memory aids (memos, diaries etc) is too often overlooked: He comments that they can provide either encoding cues for further information, or manage information that would otherwise overload storage.
object to current memories. For instance, remembering the musical term *sackbut*\(^{49}\) might involve a keyword such as *sack*, and then the linking of this word to an image: *Santa Claus playing a trombone while standing in his sack* (see Appendix B. for examples of other mnemonic techniques).

Although the primary focus above is related to encoding, many of the same factors are inseparably linked to the conditions associated with superior retention. In studying the effectiveness of retrieval cues, Tulving & Osler (1968) stated that two factors impact on the success of recall: “the amount and organisation of the relevant information…and the nature and number of retrieval cues which provide access to stored information” (p. 593). Both these factors reinforce the principle that retention depends primarily on the degree to which material is originally learned (Underwood, 1964)\(^{50}\), which in turn is better predicted by how that material is learned (i.e., encoding strategies, see Craik & Tulving, 1975). Moreover, “memory storage processes are time-dependent” (McGaugh & Gold, 1974, p. 240); as time between encoding and recall increases, memory is more likely to fail.

Overlearning was one of the earliest identified strategies that reinforced retention, although the point at which something is learned or overlearned (post mastery criteria) is somewhat arbitrary (Farr, 1987). Nonetheless, Bahrick, Bahrick and Wittlinger (1975) and Bahrick (1984) found that the extent (e.g., reinforcement or overlearning) of original learning strongly influenced retention over many decades (25 years or longer): After approximately 5 years, retention may remain unchanged for periods up to 30 years, a portion of episodic memory remaining in what Bahrick (1984)

\_______

\(^{49}\) The sackbut is an antecedent of the modern day trombone.

\(^{50}\) This somewhat contrasts with the emphasis on forgetting, for which the critical variables are meaningfulness, similarity of items and natural ability (Underwood, 1964).
termed the *permastore*. Their results showed “that very significant portions of semantic knowledge remain perfectly accessible for decades without being used at all” (p. 23) and although there were differences in recall predicted by original learning, subjects appeared to suffer from equivalent rates of declining retention.

Following these long-term retention findings, Bahrick (2000) stated, “It should be a high priority for educators and trainers to explore and implement conditions of acquisition that promote permastore stability of content” (p. 353). He suggested that the methods conducive to short-term acquisition (e.g., shorter retrieval intervals) are unlikely to produce the best long-term results; alternatively, strategies such as the *spaced or distributed practice effect* and rehearsal maintenance are ways of facilitating long-term retention (Bahrick, Bahrick, Bahrick, & Bahrick, 1993). For example, gradually increasing the interval at which information is retrieved maximises retention: “This could either be because tests induce greater encoding effort, or because they are more similar to the performance required at eventual recall” (Landauer & Bjork, 1978, p. 631). In addition, appropriately spaced rehearsal improves the retention of unstable knowledge, knowledge that can be recognised but is not easily recalled (Bahrick & Hall, 1991).

Taken together, these strategies serve to reinforce the elaborative or constructive process of encoding memories and the positive benefits this process has for retention. Understanding both the nature and meaning of elements to be remembered and how (and why) they constitute larger conceptions provides coherence (integration), adds structure, and reduces operational memory load, all of which help reconstruct memories and maintain long-term retention (Farr, 1987). Given the importance of storing and

---

51 Given the same total instructional time, a series of shorter and regular sessions is superior to fewer concentrated sessions.
retrieving knowledge within an information-processing system, in which ways are knowledge units and their applications represented and structured in LTM? One conception—schemas—is addressed in the following section.
1.5.3 Schemas.

The information-processing approach to cognition has explained the constructs of knowledge in numerous ways—schemas, concepts, scripts, frames—and although they are “not all synonymous, … Nevertheless, the various concepts are closely enough related that a discussion of any one of them will serve as an introduction to the others” (Rumelhart, 1984, p. 162). Of specific relevance to the theoretical basis of this thesis is the notion of schemas, a term that pre-dates the era of cognitive psychology; this term appears in the substantial work of Jean Piaget (1952) and variously in the earlier work of other psychologists and philosophers52.

As processing constructs, schemas represent knowledge units, allowing “people to classify information according to the manner in which it will be used” (Sweller, 1993, p. 2). According to Rumelhart (1984), schemas not only represent knowledge in the form of data (what), they also facilitate the way in which knowledge is used (how); in essence, a “theory about knowledge…[which] embodies a prototype theory of meaning” (p. 163). Under these conditions, schemas incorporate the range of LTM structures discussed earlier (see section 1.5.1), including those conceived for declarative and procedural knowledge. In addition, schemas underpin processing of tasks, ranging from straightforward recognition and retrieval through to the complex reasoning and sophisticated procedural inferences that are characteristic of skilled or expert performance (see section 1.6).

Rumelhart and Ortony (1977) have made a significant contribution to schema theory and, from a review of the available evidence, they proposed four essential characteristics of schemas:

52 For a comprehensive review of the origin and applications of the word schema, see Marshall (1995), Chapter 1, *Schema roots*. 
i. Schematic structures are inherently variable: A schema represents a generic concept, using slots to accommodate innumerable contextual and situational factors (specific stimuli) that form a probable distribution of expected variations (constraints) for that concept.

ii. Schemas are embedded within other schemas (they are not atomic): Embedding permits the activation of an interrelated network of concepts, allowing quick access to relevant knowledge. Alternatively, activating a discrete schema avoids the minutiae of related knowledge that otherwise would be called to mind.

iii. Schemas operate across all levels of abstraction: As a continuum from perceptual recognition through to complex thought, schemas represent comprehension, including plans and their actions.

iv. Schemas represent knowledge and are highly flexible. “Schemata attempt to represent knowledge in the kind of flexible way which reflects human tolerance for vagueness, imprecision, and quasi-inconsistencies” (Rumelhart & Ortony, 1977, p. 111).

Marshall (1995) emphasised the capacities of these features, including the interconnectivity of knowledge within and between schemas, the flexible access of knowledge from many perspectives, and the unlimited size to which schematic knowledge can be developed. When considered in isolation, a schema represents “relationships among a set of objects that fit into the schema’s slots” (Anderson, Greeno, Kline, & Neves, 1981, p. 207); together, existing schemas provide a

---

53 Rumelhart & Ortony (1977) suggested that knowledge representation is more encyclopaedic than definitional in character.
comprehensive mechanism that can account, or partially account, for new or novel situations. Procedural, interpretative, associative and inferential applications are all part of this mechanism. Moreover, schemas are developmental and continually adapt to new stimuli, primarily using what has gone before to acquire what is new.

The outcome of this schema-driven process efficiently preserves the ubiquity of common concepts (e.g., tree) without the need to necessarily individualise schemas for the almost infinite variation associated with those concepts (e.g., oak, birch, etc). Therefore, without loss of meaning, a schema can apply to a broad range of circumstances—complex or everyday—and is said to be instantiated (Rumelhart, 1981) when any of these situational factors, internally or externally cued, activates matching values from LTM. Take for example, the experience of listening to a Mozart symphony: to a musician, a schema would account for the predicted stylistic features that are also recognisable across many thousands of Classical works\textsuperscript{55}. As a result, schemas allow us to listen with both familiarity (style) and understanding while appreciating the creative elements that make this specific symphony a unique work (flexibility and variation). Over time, if any of these stylistic expectations are consistently violated, our schemas are up-dated, or in many cases, a new schema/s is developed\textsuperscript{56}.

These processes of flexibility, variation and abstraction can be clearly understood in the way we interpret external stimuli. For example, the need to find or

\textsuperscript{54} See also Pichert and Anderson (1977) below, who suggested slots provide a possible explanation for their results.

\textsuperscript{55} The Classical period spanned the years, ca. 1750-1820; stylistic characteristics include symmetrical phrasing, the appoggiatura, diatonic harmony, and an orchestra of strings and double winds.

\textsuperscript{56} In this way, a schema for 19th Century or Romantic music developed to accommodate a change of style.
impose form on what we perceive can be seen as an effect schemas exert on cognition\textsuperscript{57}.

Take for example the words below (Figure 4):

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{A simple example of the flexible way in which schemas help us to interpret variations among otherwise common objects.}
\end{figure}

Although the lettering is either incomplete (poor legibility) or represents significantly different shapes (fonts), few people would have any difficulty with the intelligibility of these words. The schemas we hold for letters, and indeed word groups, have sufficient utility to ensure that any gaps or variation of knowledge can be appropriately bridged in relation to either the remaining or transformed symbols. Thus, schemas are seen as the mechanism by which we construct internal reality and in turn the mechanism by which we interpret and comprehend external reality (Rumelhart, 1984).

The examples above for music and alphabetic symbols provide circumstantial evidence for the application of schemas; in addition, a number of controlled experiments have offered compelling support for this view of cognition. Bransford and Johnson (1972) were interested in the way comprehension or (mis-) interpretation of texts was influenced by either missing or inappropriate schemas. Consider the following extract from one of their studies:

The procedure is actually quite simple. First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many.

\textsuperscript{57} Gestalt theorists referred to this phenomenon as pragnanz.
Without the necessary cues (schemas), the above passage is virtually incomprehensible. Nonetheless, when the appropriate schema for “washing clothes” is activated, this passage can be read with complete understanding.

The above passage is an example of what Rumelhart (1984) refers to as data-driven or bottom-up processing: the data (parts) has to be searched, and if necessary inferences made, in order to establish a context (whole). If an inappropriate context (schema) is applied, or a context is not found, comprehension is either distorted or simply fails. In contrast, when the appropriate schema is present (e.g., washing clothes), understanding can be top-down or conceptually driven, providing an intelligible frame of reference for understanding the constituent parts. For instruction, top-down processing may be more efficient when gradually activating and apportioning meaning to a series of binding variables (Rumelhart & Ortony, 1977).

Equally, when activating top-down schemas there are important consequences for the way in which we interpret reality. Any activated frame of reference (schema) exerts an unconscious influence on what is perceived. This was elegantly demonstrated by Pichert and Anderson (1977) who provided their subjects with a description of a house from either the perspective of a home buyer or that of a burglar58. As predicted, activating different schemas before subjects read the text significantly altered the nature of items they could recall. These activities reinforce the reconstructive nature of memory and indicate that encoding and retention are schema-based processes; schemas serve to interpret what we perceive rather than to faithfully replicate an experience.

A more familiar demonstration of these schema-based processes can be found in the children’s game of Chinese whispers. For each subsequent retelling of a story,

58 These experiments also tested an island story using the perspectives of either an eccentric florist or shipwrecked person. Both stories used control groups.
reconstruction is generated by schematic fragments suggested by preceding cues, which, although often logical or likely, rapidly evolve into a story that is virtually unrecognisable from the original account. Recall is therefore based on activated schemas rather than the actual narrated event\textsuperscript{59}. Alternatively, activating schemas can dramatically alter expectations and override a more logical association between ideas. This of course is the basis for considerable humour (Howard, 1987): The example below from the popular \textit{Far Side} series of cartoons amusingly distorts reality by juxtaposing otherwise conflicting schemas (Figure 5).

\textit{Figure 5.} Taken from The Far Side Gallery (Larson, 1992, pages unnumbered).

Not only are schemas the mechanism by which we comprehend our world, they are the means by which we learn. According to Rumelhart and Norman (1981), this happens in three ways:

i. Accretion – new information is processed via existing schemas.

\textsuperscript{59} Bear in mind the limits of working memory and therefore the need in this instance to ‘fill in’ gaps when attempting to reconstruct a story.
ii. Tuning or Evolution – schemas are modified and refined.

iii. Restructuring – new schemas are created.

They believed that only accretion conforms to the traditional notion of learning as memory, and that for complex learning, effort needs to be focussed on schema evolution and schema creation. Rumelhart and Norman suggested that much of our knowledge is stored in the form of specialized procedural schemas, a proposition that is particularly relevant when solving problems (see section 1.6). Not only is schema acquisition problem specific, but concrete rather than abstract problem representations are more likely to be understood if they tap schematic representations that are consistent with existing knowledge (Howard, 1987).

When all these factors are considered, the implication for instruction is clear: understanding of new concepts is not directed by general rules of inference, but is far better taught in relation to current knowledge (schemas). In the following three sections, these ideas are explored within the domains of expertise and problem solving.
1.6 Expertise and Problem Solving: Implications for Cognitive Mechanisms

1.6.1 LTM processing and (chess) expertise.

From the previous discussion (see section 1.5) it is apparent that LTM is a repository for knowledge in the form of schemas. Although individual schemas can incorporate large amounts of familiar information, working memory remains extremely restricted when handling novel information; under these conditions only a limited number of schemas (basic elements) that require combining in a novel fashion can be managed at any one time. Given these mechanisms, it was reasonably assumed that the highly skilled application of knowledge—expert performance—was primarily attributable to individual problem-solving ability. Somewhat surprisingly, the means by which light was cast on this issue arose from the game of chess and the exploration of expert-novice differences.

Much of the early and systematic research relating to the structure of chess thinking was carried out by De Groot (1978). His classic study included briefly exposing (from 2-10 s) four levels of player (novice, expert, master and grandmaster) to sixteen different board configurations randomly taken from actual master games. All players had the same opportunity to hold what they saw in working memory. He recorded both player introspections and the accuracy with which each player could reconstruct these configurations; the results were not only surprising, they revealed what were to become seminal and influential findings in cognitive psychology: There was little difference between master and grandmaster levels, but a large difference between

60 Simon and Chase (1973) believed that chess was an excellent model environment: “As genetics needs its model organisms, its Drosophila [fruit fly] and Neurospora [genus of fungi], so psychology needs standard task environments” (p. 394).
61 De Groot was not only a Dutch psychologist, but also an avid player of master status!
62 In fact, weaker players were given slightly longer in order to avoid floor effects.
master and non-master levels (expert and novice). He concluded, the “master’s experience enables him to quickly ‘integrate’ the picture of the position and through this to imprint and retain it within a very short period” (De Groot, 1978, p. 329).

As mentioned earlier, expert performance often involved assumptions of superior mental capacity, especially in relation to the depth and speed with which information can be processed in working memory. In fact, De Groot (1978) found no systematic differences in thinking between player levels, attributing any superiority to a substantial accumulation of experience and the resultant extent to which real-game configurations are stored in LTM. This experiential advantage represented more than a large series of positional ‘snapshots’, De Groot commenting, “the essential relations between pieces, their mobility and capturing possibilities, their cooperation or opposition, are often perceived or retained better than the position of the pieces themselves” (p. 331). Although it was found that (grand) masters did not search any deeper, the knowledge they recalled was essentially a complex of information63. In essence, the familiar chunk also gives access to associated information that helps guide the best possible move related to that chunk (Gobet & Simon, 1996b).

These findings were replicated and extended by Chase and Simon (1973a; 1973b), who reasoned that master players were able to encode board configurations into larger relational chunks (more pieces per chunk). To investigate this issue, they included a perception task that asked subjects to reproduce a board configuration while still visible, believing that each glance at the pre-set board during reconstruction represented a chunk. As player skill increased, within glance intervals decreased (perception task), and the size (pieces) and number of chunks increased (memory task). In addition, they

---

63 A master player is likely to perceive more information within 5-10 s exposure than a lesser player can in a quarter of an hour (De Groot, 1966).
found no relationship between levels of expertise and performance for reconstructing random–atypical–configurations of pieces. Together, these results provided little evidence that chess expertise was likely to be based on a superior working memory\textsuperscript{64}. Rather, expertise was based on a vast store of domain specific knowledge, including “hundreds and perhaps thousands of sequences of moves stored away in long-term memory” (Chase & Simon, 1973a, p. 261).

In a further effort to explore the basis of these expert-novice differences, Schneider, Gruber, Gold, & Opwis (1993) compared adult and child chess players\textsuperscript{65}. They found a significant difference for reconstructing board configurations between child experts (8.8 pieces) and adult novices (4.6 pieces), a difference that was maintained until random and unfamiliar patterns were introduced. Moreover, Schneider et al. also administered a WISC digit span task as a control for processing capacity, and as anticipated the digit spans for the children were on average lower (5.8) than those for the adults (7.6). They concluded, in line with earlier results, that for chess experts meaningful domain knowledge “enables them to process information faster and in larger semantic units, which allows them to use different strategies and to remember and learn more than novices in this domain of expertise” (p. 348).

All these studies (Chase & Simon, 1973b; De Groot, 1978; Schneider et al., 1993) strongly indicated that the major determinant of expertise reflects acquired knowledge and skills (Ericsson & Charness, 1994), which for mastery in a particular domain requires at least 10 years of concentrated preparation (Ericsson, Krampe, &

\textsuperscript{64} In a subsequent paper, Simon and Chase (1973) also commented that there was no “evidence that masters demonstrate more than above-average competence on basic intellectual factors” (p. 403).

\textsuperscript{65} This work built upon the classic study carried out by Chi (1978), in which she demonstrated the central role of knowledge as a determinant of memory development.
Tesch-Römer, 1993). Within any domain, this acquired knowledge largely determines how well a person will perform, irrespective of individual and, in some cases, developmental differences (Chi, 1978). For example, the extent to which expert knowledge for chess is represented in LTM was estimated by Simon and Gilmartin (1973) at somewhere between 10,000–100,000 patterns (schemas). Given the possible range of board configurations, these were upper and lower estimates that could explain the performance of chess grandmasters to hold a single configuration in working memory after 5-10 s exposure.

In a recent review of the limited neuropsychological research involving the game of chess, Bart & Atherton (2003) agreed, with some caution, that there was prima facie support for the cognitive psychological findings in this area. As a summary of their results, they stated that activations of the frontal lobe for expert ("professional") chess players suggested “higher-order reasoning and the utilization of expert memory chunks from well-organized chess knowledge memory stores” (p. 8). In support of this storage-based view, Amidzic, Riehle, Fehr, Wienbruch and Ebert (2001) found “that the activation of expert memory chunks produces focal $\gamma$-band activity in the neocortex, whereas amateur players primarily encode and analyse new information” (p. 603).66

The explanation of human expertise and skill acquisition as products of knowledge structures in LTM (Ericsson & Charness, 1994), including the mechanisms associated with chunking theory (Chase & Simon, 1973a, 1973b), has been demonstrated across a diverse range of human endeavours: algebra (Sweller & Cooper, 1985); electronics (Egan & Schwartz, 1979); map reading (Anderson & Leinhardt, 1982).

66 They state, “active memory is indicated by bursts of $\gamma$-band activity in [temporal lobe structures] and other areas of the association cortex” (p. 603); such activity is implicated in the initial formation of long-term memories.
2002); medicine, (Groen & Patel, 1988); reading comprehension (Recht & Leslie, 1988); software design (Jeffries, Turner, Polson, & Atwood, 1981); and text generation (Voss, Vesonder, & Spilich, 1980). Furthermore, in their far-reaching review of expert performance, Ericsson & Charness (1994) proffered a broader view of the implications arising from these findings:

There is no reason to believe that changes in the structure of human performance and skill are restricted to the traditional domains of expertise. Similar changes should be expected in many everyday activities, such as thinking, comprehension, and problem solving, studied in general psychology (p. 745).

In other words, the transition from novice to expert taps cognitive processes accessible to, and descriptive of, any person, a notion probably best summarised by Charness (1992) when examining the impact of chess on cognitive science: “experts are made, not born” (p. 6).

Expert-novice research has offered many profound insights into the study of LTM, and for that matter the study of cognition at large. Chase and Ericsson (1981; 1982) referred to the conspicuous ability of experts to rapidly access and utilise domain knowledge as *skilled memory*. They suggested that “intermediate knowledge states are stored in directly addressable locations” (p. 177) as part of LTM retrieval structures that help to circumvent working memory limitations. Chase and Ericsson (1982) “argued that these ingredients produce an effective increase in the working memory capacity for that knowledge base” (p. 55). What is more, the expert then applies this knowledge base or memory skill with superior monitoring and self-regulation (Glaser & Chi, 1988).

More recent advances have modified the LTM chunking model: Chunk size (for masters) may be higher than originally estimated (from 6-7 to >10, Gobet & Simon, 67 See also the mechanism of automation (section 1.6.3) in which an expert’s access to LTM is relatively rapid in comparison to the conscious processes of working memory.
1996a, 1998) and encoding mechanisms appear to allow, under high working memory load, much of the accessible information to be stored in LTM (Ericsson & Kintsch, 1995). Such a proposition helps to explain the results generated by Charness (1976) in which interference tasks degraded recall of chess positions only marginally (6-8%), leading him to suggest, “With well-structured or meaningful chess material…subjects rapidly fixate the information in LTM” (p. 653).

These advances, arising from the study of expertise, have led to the more recent introduction of new theoretical models, which continue to refine our understanding of cognitive architecture (e.g., Long-Term Working Memory, Ericsson & Kintsch, 1995; Template Theory, Gobet & Simon, 1996c). In particular, the study of skill acquisition across a range of domains has been evident in the area of human problem solving, an area that is pursued in the following discussion.
1.6.2 Problem solving: From puzzles to expertise in complex domains.

“Cognitive skill acquisition has its historical roots in the study of problem solving” (VanLehn, 1996, p. 514), and in turn the ability to solve problems is seen as a highly valuable skill, especially in relation to learning (Sweller, 1988). As noted in the complex domain of chess, grandmasters solve countless problems (games) in order to accumulate the necessary schematic knowledge to achieve expertise. It would therefore appear reasonable to conclude that expertise or skilled memory is largely founded upon the activity of solving problems. Even a casual examination of school curricula and teaching programs will attest to the widely accepted proposition that problem solving is an effective instructional strategy. Despite the plausibility of this proposition, the activities of problem solving and their relations to learning are not so easily reconciled.

The landmark work of Newell and Simon (1972) offered a detailed and dissective understanding of problem solving (Kotovsky, Hayes, & Simon, 1985), no doubt due in part to its information-processing view of cognitive psychology68. They referred to problem solving as operating within a problem space, a metaphorical processing frame within which the major components of a problem could be studied:

i. **Initial state** – represents the problem givens and therefore what is known prior to solving the problem.

ii. **Goal state** – represents a solution, or what needs to be known once the problem is solved.

iii. **Operators** – represent the constraints and limitations imposed by the problem, and the permissible moves necessary to solve the problem.

---

68 See Simon (1990) for a succinct and insightful overview of the relationship between cognitive psychology, information processing and the search for *Invariants of Human Behavior*. 
Greeno (1978) specified three types of problem: *structure inducing*, *transformation* and *arrangement*. In particular, the early study of transformation problems\(^{69}\) proved most helpful in understanding why some problems are more difficult to solve than others. Newell and Simon (1972) suggested that the “size of the problem space provides one key to the estimation of problem difficulty” (p. 93); increasing the number of operators or possible moves quickly taxes an individual’s information-processing capacity. Also associated with problem difficulty is the recognition of problem goals: A goal may be either known or unknown, and either require a single solution that is attainable by an algorithmic process (*well defined*), or permit multiple solutions in which resolution typically relies upon heuristic processes (*ill-defined*).

The ability to solve a problem is determined in large part by its mental representation (Hayes, 1978) and the subsequent knowledge recalled in order to apply an appropriate solution. For example, the classic *Monk* problem usually induces a number of convoluted solutions (e.g., mathematical), and yet if the problem (journey) is represented graphically, the solution is immediately apparent, averting any need for other mental gymnastics (Appendix D.). Unsurprisingly, schemas activated during a task not only affect what we remember (see section 1.5.3), but also exercise a major influence over the approach taken to solving a problem.

This propensity to view a problem, somewhat inflexibly, from either a prevailing or preconceived perspective (instantiated schema), rather than one that would more readily solve the problem, is clearly evident with the phenomena known as *Einstellung* and *functional fixity* (Appendix E.): In brief, once committed to a perspective, it is

\(^{69}\) These problems require the solver to construct a problem space and understand the required operations in order to move from an initial state, through intermediate states, to a goal state. They include many of the classic puzzle problems, e.g., Water Jug; Tower of Hanoi; Missionaries and Cannibals.
particularly difficult to deviate from a perceived course of action (Rumelhart, 1984; Sweller, Mawer, & Ward, 1983). Some of the earliest research into these effects was carried out by Luchins (see Luchins & Luchins, 1959) who studied the well known water jug problems (Appendix F). From these studies it became apparent that the way in which tasks were structured could in effect blind subjects to an otherwise simple solution, leading to some problems being insoluble. Contrary to previous evidence though, Sweller & Gee (1978) were able to readily abolish the Einstellung effect by minor manipulations of perceptual cues, remarking “that under appropriate circumstances problem solving may be more heavily influenced by the manner in which the information is presented...than by the information content itself” (p. 523).

Clearly, successful problem solving involves an accurate mental representation of the problem space, which in turn activates appropriate schemas to attain a solution. Known solution paths operate in a forward direction, however in the absence of appropriate schemas problem solvers usually adopt a search-based strategy, which by contrast often moves in a backward direction (Hayes, 1981). The most common of these search strategies is means-ends analysis, a general technique not only observed when solving transformation problems, (Atwood, Masson, & Polson, 1980; Atwood & Polson, 1976; Greeno, 1978), but also “used extensively by human subjects in many problem environments” (Simon & Lea, 1974, p. 109). Means-ends analysis, through a

---

70 This phenomenon is sometimes referred to as mental set, and can be associated with the mental rigidity resulting from rote learning and the failure to flexibly adapt problem solving strategies.
71 These insights acted as precursors for the fundamental principles associated with cognitive load theory.
72 Hayes (1981) included four search methods: proximity (includes means-ends analysis), trial-and-error, knowledge-based, and fractionation. He refers to these methods collectively as search heuristics.
73 Means-ends analysis forms but one class of problem solving strategies (Bratko, 1991); random search, algorithmic and heuristic processes provide further examples.
series of operators, attempts to reduce the difference between the current state (subgoal) and the final state (final goal). For example, the *Tower of Hanoi* is a typical puzzle problem in which subjects generally employ means-ends analysis in order to find a solution (Figure 6).

![Image of Tower of Hanoi](image)

*Figure 6.* The Tower of Hanoi problem requires moving, one at a time, all discs from the first peg (initial state) to the third peg (goal state) without placing a larger disc on a smaller one. Taken from Gilhooly (1996, p. 34).

In the absence of prior knowledge, the solver will search the problem space for ways of gradually reducing differences between what is known (left discs) and unknown (right discs, Egan & Greeno, 1974). Although this process sounds deceptively straightforward, Owen and Sweller (1985) provided insight into the mental activity subjects are compelled to undertake when using a means-ends strategy:

In order to use means-ends analysis, a problem solver must simultaneously consider the problem goal, the current problem state, relations of the goal state to the current problem state, relations of this relation to allowable operators, and in addition, the maintenance of any subgoal stack that has been constructed (p. 284).

Although a means-ends strategy, under many circumstances, may lead to a solution, Sweller, Mawer, and Howe (1982)\(^74\) argued that it does not necessarily promote rule induction (forward-working strategy), an ability to transfer knowledge from one problem to another (Sweller, 1976, 1980). They proposed a *history-cued* strategy.

\(^74\) Refer also to Mawer and Sweller (1982) and Sweller (1983).
strategy, which helps monitor and retain patterns of past moves, and results in positive transfer to similar problems. “In so far as this transfer represents expertise or knowledge of the problem structure, [they also suggested that this] strategy will result in more rapid acquisition of knowledge” (p. 480). As Sweller and Levine (1982) stated, it is “theoretically possible for a problem solver to attain the goal using means-ends analysis and yet to learn absolutely nothing concerning the problem structure” (pp. 463-464).

These insights led Sweller (1988) to conclude that problem solving via means-ends analysis and learning via schema acquisition are primarily independent activities. “There seems to be no clear evidence that conventional problem solving is an efficient learning device and considerable evidence that it is not” (p. 283). These findings challenged traditional beliefs; indeed, evidence from problem solving in real-world domains also supported the view that schema acquisition underpins expert performance.

In a much studied area, qualitative differences were found between the ways in which experts and novices solved physics problems (Larkin, 1981; Larkin, McDermott, Simon, & Simon, 1980a, 1980b; Simon & Simon, 1978): Experts’ speed and accuracy were based on forward-working strategies, which in turn were indicative of rich schematic knowledge and the ability (as with chess) to very rapidly select from thousands of problem-specific solution patterns. In contrast, novices often resorted to a means-ends strategy in order to find and test possible equations. Likewise, when asked to categorise physics problems according to their mode of solution, novices were inclined to rely on surface features as opposed to the deeper principles characteristic of an expert’s approach (Chi, Feltovich, & Glaser, 1981; Chi, Glaser, & Rees, 1982). Not only do experts apply specialised problem-solving strategies (schemas), these strategies

---

75 They referred to Egan and Greeno’s (1974) notion of working in the past or future when subjects solve a problem.
can be activated after just the briefest exposure to a problem (Hinsley, Hayes, & Simon, 1977)\textsuperscript{76}.

As earlier emphasised, domain knowledge (schemas), irrespective of ability, is “a powerful determinant of the amount and quality of information recalled”, providing a scaffold upon which new information can be quickly recognised, organised and applied in appropriate ways (Recht & Leslie, 1988, p. 19)\textsuperscript{77}. This influence of schema acquisition, involving a shift from novice to expert strategies, also has been observed within many areas of problem solving, including the domain of mathematics (e.g., Schoenfeld & Hermann, 1982; Sweller et al., 1983). In all these cases, the differential strategies used for problem solving reflected the differences between the schemas that had been acquired by either novices or experts (Gick, 1986)\textsuperscript{78}.

From the extensive work in problem solving and expertise, considerable evidence now has been advanced for the importance of instructional activities that promote the acquisition of schemas (see also, Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Gick & Holyoak, 1980, 1983). In other words, learning is not founded upon generalized reasoning skills (Rumelhart & Norman, 1981), but is highly contextualized, knowledge intensive, and progressively attained following considerable time and practice (Chase & Ericsson, 1982). This emphasis on schema acquisition and the cognitive conditions required for effective learning is fundamental to cognitive load theory (Chapter 3). The degree to which schemas can be acquired and, as a result, influence these cognitive conditions is discussed in the following section.

\textsuperscript{76} Using algebra word problems, they found that little more than the initial noun phrase was necessary to initiate schema activation!

\textsuperscript{77} Ironically, without the correct schema (acting as a constraint), domain knowledge “can actually be counterproductive if it leads to search in the wrong part of the space” (Kaplan & Simon, 1990, p. 413).

\textsuperscript{78} Sweller et al. (1983) noted that at least in some cases knowledge development is extremely slow.
1.6.3 Automation of schemas and expert performance.

That the progressive attainment of a skill is in large part due to continual practice would hardly surprise anyone. The more difficult question has always been, by what mechanisms is this improvement brought about? The fact that experts make a complex skill appear effortless, while our attempts at the same skill may be erratic, slow, painstaking and cumbersome, only serves to bring this question into sharper focus. In the previous sections, schema acquisition was seen as the basis of intellectual skill and expertise, however the acquisition and application of schemas should not be seen as a two-way ‘switch’. More accurately, the incorporation and adaptation of schemas represent incremental change (Sweller, 1994) corresponding to a gradual shift from controlled to automatic processes, a continuum “in the attention requirements of mental operations” (Hasher & Zacks, 1979, p. 357).

Schneider and Schiffrin (1977) and Schiffrin and Schneider (1977) investigated this issue and as a result provided a comprehensive basis for understanding the nature of controlled and automatic processing79. They had subjects decide whether one or more alphanumeric target objects (consistent or varied) were among a larger set of distractor objects80. With respect to their findings, Schneider and Schiffrin explained the superiority of the consistent targets condition as a product of the lower processing attention required to detect a familiar object under variable load conditions (time and number of objects). The authors concluded that consistent mapping of a subject’s responses eventually results in the action operating “independently of the subject’s control” (p. 51). This process of automation is learned in LTM over many trials, and can

79 These findings were also incorporated into a two-process theory of human information processing.
80 Alphanumeric configurations included, consonants among digits (or vice versa), or consonants among consonants, or digits among digits.
be unlearned only at the expense of additional work to overcome the effects of negative transfer.\textsuperscript{81}

From these studies, the authors characterised automatic processing (AP) as follows:

i. AP requires little conscious control and consumes negligible processing capacity\textsuperscript{82}.

ii. AP is less flexible than controlled processing, and once initiated progresses in a fashion according to how it was acquired.

iii. The component parts of AP are inseparably linked and only indirectly lead to new learning.

iv. AP is a product of extensive practice, improvement is gradual, and asymptote is reached only following numerous trials.

Although the distinctions between automatic and controlled processes are quite stark, nearly all acts of cognition are likely to be a combination of both processes (Shiffrin & Dumais, 1981). Accordingly, variable processing provides remarkable flexibility in the adaptation of cognitive mechanisms to new tasks. Tasks or sub-tasks can be accomplished more accurately and rapidly once automated, and these elements can be manipulated in parallel with controlled processing, decreasing the risk of exhausting working memory capacity. That is, when more basic mental operations occur automatically, it leaves “maximal resources available for less commonly used and more sophisticated mental operations” (Hasher & Zacks, 1979, p. 360). Viewed in this

\textsuperscript{81} It took 2100 and 2400 trials to automate and unlearn responses respectively.

\textsuperscript{82} Neisser referred to these processes as preattentive; with sufficient practice, many preattentive responses can become \textit{automatisms} (Neisser, 1967).
way, automation is seen as “a major component of skill acquisition in both the cognitive and motor domains” (Shiffrin & Dumais, p. 139). In addition, this mechanism provides further insight into why some problems are so difficult to solve.

Kotovsky et al. (1985) questioned why some versions of isomorphic problems remained intractably difficult despite their having a relatively small problem space. Unsurprisingly, they found that problem representation, and the learning and application of rules were major contributors to problem difficulty (see section 1.6.2). Moreover, they observed that subjects would often explore the problem space for disproportionately long periods and then 'see' the solution, making rapid progress in a two-chunk pattern (two and then three moves). Kotovsky et al. concluded that exploration of the problem space allowed for goal planning (learning), but for this to occur problem rules needed to be sufficiently automated. “Automating the rules removes some of the load of remembering the current state, the goal, the intervening state, and the legality of tests involved in compiling two moves” (p. 287).

Similarly, Neves & Anderson (1981) viewed automation as a series of productions, which through the process of composition are parsimoniously reduced to a single procedure. With continual application of these procedures, the production rules automatically speed up and become more specifically geared to the way in which knowledge will be used. Gick & Holyoak (1983) referred to this continual application as a recurrent task, suggesting automation of these tasks frees available resources in order to efficiently handle non-recurrent components. Over time, processing becomes more

---

83 Kotovsky et al (1985) used versions of the Monster problem (an isomorph of the Tower of Hanoi problem) and estimated this difficulty in relation to node branches and depth of search.

84 This pattern accounted for only 15% of the total problem solving time, although 38% of the total number of moves. They referred to this pattern as final path behaviour.
efficient and selective (Kolers, 1975), although automation “can demand and/or utilize capacity, even if only temporarily” (Shiffrin & Dumais, 1981, p. 115).

Numerous studies of expertise have supported the importance of automation as an essential mechanism of attaining very high levels of expertise (Ericsson & Polson, 1988; Lesgold, Glaser, Rubinson, Klopfer, Feltovich, & Wang, 1988; Staszewski, 1988). In fact, “experts’ knowledge structures function automatically, so automatically that these structures often are not consciously available to the expert” (Britton, Gulgoz, & Glynn, 1993, p. 37). This may explain why at times an expert can find it difficult to deconstruct their understanding when attempting to instruct a novice. It also explains why Glaser and Chi (1988) suggested experts have superior working memories: “This is not because their short-term memory is larger than other humans’, but because the automaticity of many portions of their skills frees up resources for greater storage” (p. xviii).85 The identification of this mechanism has formed an invaluable plank in our understanding of human cognition, and it is this more recent understanding that has directly led to the theoretical framework upon which this thesis is based: Cognitive load theory is comprehensively reviewed in Chapter 3.

85 Recall that this idea was also supported by Chase and Ericsson (1982, see section 1.6.1).
Chapter 2: Cognition and the Processing of Music

2.1 The Nature and Perception of Music Processing

The following sections broadly examine the ways in which music is perceived and processed. Consistent with previous chapters, observations are made from both an information-processing perspective and in relation to the cognitive architecture that defines our intellectual capacities. It should also be noted that the experiments reported in Chapters 5-9 did not seek to explore issues of music cognition directly; rather, they included music materials in order to explore and generalise instructional principles generated by cognitive load theory. Notwithstanding this important distinction, music is a highly specialised domain and it is therefore appropriate that the nature of music processing, and the characteristics it may or may not share with the cognitive mechanisms detailed in previous chapters, are briefly reviewed.

The experience of music is universal and a “musical brain is the birthright of all human beings” (Hodges, 2000, p. 18). Of some significance therefore are the issues surrounding the adaptive purpose of music: Krumhansl (1992) suggested that aesthetic, social and communicative factors serve to emphasise the importance of music and the ways in which it “produces strong emotional effects and engages intricate perceptual and cognitive processes” (p. 198). With this in mind, she believed that the study of musical cognition expands our understanding of human thought, an understanding that is ultimately reliant on identifying the nature of internal or mental representations. In this sense, if not more generally, the processing of music shares a common functionality with broader cognition. Although cognitive psychologists and music theoreticians have not always appeared to share common ground (Agmon, 1990; Hantz, 1984; Walsh, 1984), more recent evidence suggests their conceptions often fall within the same sense
of psychological reality (Zatorre & Krumhansl, 2002).

Music is primarily an auditory phenomenon. Musical sound is either a stimulus for, or by-product of, mental processes, and in turn this complex of mental processes represents what loosely might be termed musical thinking. Karma (1994), in a study involving school students and a group of congenitally deaf subjects, concluded that “sound, physically or mentally, is not a necessary condition for musical thinking” (p. 27). Rather, cognitive processes associated with music are also indicative of broader aptitudes that exist outside the conception of music. Such a finding underscores that music is a reflection of adaptive human capacities and that “responding to a piece of music requires far more than the perceptual skills of the auditory cortex” (Abbott, 2002, p. 12). As Sloboda (1985) comments, “What makes the composer or performer special is his rarity rather than anything fundamentally different about his mental equipment” (p. 3).

A basic question that arises within psychology is whether music taps similar neural processing networks to other complex tasks. More recent neuropsychological evidence, especially in relation to language and music, is starting to shed light on this issue. In defining areas of musical activation, Andrade and Bhattacharya (2003) pointed out that interference to the perception of pitch and timbre is more obvious with lesions in the right hemisphere, and interference to the perception of rhythm and recognising musical sound is more obvious with lesions in the left hemisphere. Peretz and Zatorre (2005) characterised the neural location of musical processing as follows: “a vast network of regions located in both the left and right hemispheres of the brain, with an overall right-sided asymmetry for pitch-based processing” (p. 105). They also remarked that despite the fact that all sounds—musical or otherwise—share most of the same processing stages, “evidence points to a degree of functional segregation in the
processing of music” (p. 90), and to the existence of separate neural sub systems for temporal and pitch relations. With respect to speech perception and melodic patterns, Vroon, Timmers, and Tempelaars (1977) concluded that temporal analysis of both these acoustic patterns occurs mainly in the left hemisphere.

Further support for these findings is gaine from the study of patients with neurological impairment. The observation and comparison of patients suffering disorders such as amusia and aphasia\(^{86}\) have helped clarify task-specific (music and non-music) areas of neurological activation. Peretz, Gagnon, Hebert, and Macoir (2004) believed that such cases “point to the existence of at least two distinct series of processing modules: one for music and one for speech” (p. 377). Furthermore, they indicated that this demarcation has been found in musicians and non-musicians alike. These neurological differences include the processing of auditory music and even extend to a functional distinction between sight-reading music and the reading of either words or Arabic numerals (Peretz & Zatorre, 2005).

Zatorre (1984), after reviewing a wide series of studies, found similar delineated patterns of neural processing (music and language), also commenting that there was no one discrete processing centre for music\(^{87}\). Naturally, this processing dissociation between speech and music (see also Peretz et al., 2004), including any consequent distribution of mental load, has likely consequences for instructional design. Moreover, examining music as a higher order function, as we do with language, may lead to “a more complete understanding of the entire cognitive system” (Zatorre, 1984, p. 197).

---

\(^{86}\) These conditions represent the inability to process music and the inability to speak respectively. Other disorders that have been studied include, agraphia (inability to write), alexia (inability to read), and ideomotor apraxia (loss of coordinated movement; Andrade and Bhattacharya, 2003).

\(^{87}\) Zatorre (1984) also indicated that hemispheric dominance in music is not as pronounced as that (left) found for language.
Before turning to these issues in the following sections (2.2-2.4), it is important to outline the nature of auditory musical elements and the way in which they are perceived and processed as cognitive ‘units’. According to Bigand (1993), research indicates there are two primary questions when investigating music as a phenomenon of auditory cognition: How do listeners transform an agglomeration of atmospheric vibrations into elements “having specific auditory qualities and coherence”;…[how do] listeners manage to perceive relations between sound events separated in time” (p. 231)? Addressing these questions in detail is beyond the scope of this thesis, however some associated comments below and in the following sections might serve to reinforce the nature of auditory sensations and the ways in which they are comprehended.

Musical tones are identifiable by their force (amplitude), pitch (frequency) and quality (peculiarity, Helmholtz, 1954). The brain organises tones into six specific elements: pitch, intensity, time, timbre, tonal memory, and rhythm (see Hachinski & Hachinski, 1994). Not only are these elements fundamental to human cognition, there are certain structural properties of auditory perception that appear to be common to all music (Dowling & Harwood, 1986; Krumhansl, 1992). Among these cross-cultural and universal properties are discrete pitch intervals, octave equivalence, dominant duration (rhythmic) ratios of either 2:1 or 3:1, and limited pitch subsets (4-7 notes) contained within an octave. Together, these elements and the properties of their perception form the basic building blocks of musical cognition, and when combined, create complex interrelationships of sound. “Adding to this complexity…of musical processing is the fact that music, like all sounds, unfolds over time…[and] must depend to a large degree on mechanisms that allow a stimulus to be maintained on-line to be able to relate one element in a sequence to another that occurs later” (Peretz & Zatorre, 2005, p. 95).
So how does the brain find musical coherence when simultaneously processing multiple elements delivered in a continuous stream of auditory events? To investigate this issue, researchers have turned to the *Gestalt* grouping laws that define universal tendencies to organise or group patterns of stimuli (Appendix A). Deliège (2001) states that these laws “formalize a spontaneous and unconscious tendency of human psychological mechanisms to define *units* in the perceptual field based on the properties of objects and the relationships between them” (p. 235). These laws include *pragnanz*, *proximity*, *similarity* and *common fate* (Dowling & Harwood, 1986); some of these laws have become crucial to theories dealing with the way in which we perceive structural organisation in music (see section 3.3).

Deutsch (1975b) refers to some of these auditory grouping phenomena as illusory percepts, many of which have been studied as channelling effects arising from dichotic listening tasks. Sloboda (1985) provided a succinct illustration of grouping tendencies with an example from Tchaikovsky’s *Sixth Symphony*. Figure 7.1 gives the written parts (actual score) and Figure 7.2 shows what most listeners usually perceive, an illusion that holds even when the two violin sections are heard from opposite sides of the stage!

![Figure 7. An example of the scale illusion, taken from Sloboda (1985, p. 157).](image)

Of course, along with the perceptual organization of stimuli, music also involves the conscious holding and manipulation of perceptual units in working memory and, in turn, their storage in and retrieval from long-term memory (Abbott, 2002).
Incorporating these major constructs of our cognitive architecture, Carroll-Phelan and Hampson (1996) generated the following hypotheses in their model of auditory perception: 1. Perceptual analysis is stored as representations of the auditory image in the long-term melodic memory system. 2. These long-term memory images can be activated by direct or associable cues. 3. Auditory images have dissociable rhythm and pitch components. 4. The rhythm component cannot be activated without some activation of neural systems associated with motor activity. 5. Short sequences can be briefly held through activity (rehearsal) in the rhythm or pitch subsystems (of working memory).

These hypotheses accord with the general conception of cognitive mechanisms outlined in Chapter 1, mechanisms which are viewed from a music processing perspective in the following sections.
2.2 Working Memory and Music

Within working memory, the auditory processing of musical tones engages similar neural regions to other domains (Peretz & Zatorre, 2005). Nevertheless, the evidence for dissociable verbal and musical processing subsystems (see section 2.1) also suggests that processing two or more inputs of either music and/or speech may not necessarily induce the same working memory conflicts. For example, Deutsch (1970) demonstrated that when either distractor tones or numbers separated two tones, subjects in the number conditions (recall or no recall of numbers) were able to equally discriminate (same or different) between the two pitches; conversely, “Considerable interference was produced by the intervening tones” (p. 1604). Crowder (1993) referred to the independence of auditory processes, including verbal, pitch and timbral elements, as “a tidy modularity in memory” (p. 124).

This pattern of differential disruption in working memory indicates there is at least partially separate coding of these respective stimuli. Although words and musical sequences are retained equally, music “may perhaps be thought of as being roughly the auditory equivalent of pictures” (Standing, 1973, p. 220). Sharps and Pollitt (1998) suggested that auditory images “appear to provide cues specific to the items to be remembered in the same way that visual images do” (p. 110), speculating that these images in combination with speech may result in dual coding (see Paivio, 1986, see section 1.5.2). Furthermore, Sharps & Price (1992) found that combining visual and auditory images did not lead to improved recall over one of these sources in isolation, a result that implies some sharing of mental resources. Nevertheless, the addition of either
auditory or visual imagery was superior to verbal material alone\textsuperscript{88}. In addition, auditory processing may assume some functional independence between spatial and nonspatial auditory tasks (Arnott, Grady, Hevenor, Graham, & Alain, 2005).

Findings such as these prompted Berz (1995) to propose a modified conceptualisation of working memory (see section 1.4.3, Baddeley, 1997). He believed this revised model (Figure 8) “is important in understanding the significant demands placed on working memory in such activities…in which there would be competition between storage and processing functions” (Berz, 1995, pp. 361-362). According to Berz, an additional slave system exists for the storage and processing of music; it is much less certain whether it is totally or partially independent of the phonological loop.

\textbf{Figure 8.} The proposed theoretical model of working memory by Berz (1995, p. 362), based on Baddeley (1990).

In keeping with this psychological construct of working memory, there have been various investigations into its limitations when processing music. For example, Long (1977) found that memory capacity for a series of tones was between 11 and 15. Pembrook (1986) demonstrated a noticeable decrease of recall from six- to ten- note

\textsuperscript{88} Sharps and Price (1992) referred to visual and auditory images as having mnemonic value (see section 2.3).
(dictated) melodies, and a sharp decrease from ten- to sixteen-note melodies, estimating a short-term memory capacity of between 10 and 16 tones. This capacity is affected by other factors and may quickly decline under some conditions (see Pembrook, 1987, see also section 3.3)\textsuperscript{89}.

Fraisse (1982) has studied extensively the perception of temporal musical phenomena, including the limits imposed when processing groups of notes. He indicated that up to approximately 25 notes could be perceived if they are chunked (e.g., 5 groups of 5) and follow each other in rapid succession (180 ms). Nonetheless, even at this extreme, auditory memory cannot exceed 5 s in total, an upper limit that Fraisse associated with related activities such as reciting the longest lines of poetry (4-5 s)\textsuperscript{90}.

“This duration limit corresponds to what has been called the \textit{psychological present}” (p. 158). Dowling and Harwood (1986) also commented that the evidence, although weak, “points in the direction of a natural pace for psychological events of 1.3 to 1.7 per second” (p. 182), which they proposed accords with an intuitively moderate tempo of 80-100 beats/min.

These estimates reflect the capacity of working memory to process musical events, and are broadly in keeping with earlier reviewed limits (see sections 1.4.1, 1.4.4). As proposed by Miller (1956), working memory can hold approximately 7 +/- 2 chunks of information, a load that generally encompasses the findings above, and may also explain the common use of 5 to 7 discrete scale steps within an octave (itself a perceptual unit) that are used by most cultures across the world (Dowling, 1978).

\textsuperscript{89} The first span of 6-10 notes falls within Miller’s (1956) original estimates; any increase may be due to other factors involving chunking and melodic schemas, e.g. tonality, harmonic constructions etc.

\textsuperscript{90} Of related interest, the perception of “meter can only occur with respect to periodicities in a range from about 100 ms to 6 s” (London, 2002).
Musical units, or chunks, can incorporate and store in LTM either minimal information, or in the case of expertise (see section 2.4), extensive series of elements.

This chunking mechanism may in fact explain the working memory capacity—0-180 s—found by Kauffman and Carlsen (1989), who suggested that the chunking of music may be “highly efficient, thereby allowing for a large storage capacity” (p. 11). Of significance, they used intact music, which comprised a more meaningful stimulus than is often found in research, a move that may have promoted superior recognition. Take for example the activity of improvisation, in which continuous streams of notes instantaneously generate a cohesive melody. Johnson-Laird (2002) believed this creative process can virtually bypass working memory for intermediate computations by calling on procedural rules (algorithm), the memory buffer holding “only the previous note that is improvised, and its current place in the chord sequence, the contour, and the measure” (p. 439).

Further evidence of the processing constraints imposed on musical materials has arisen from the broader psychological effects that helped pioneer much of the early research into working memory (see section 1.4). For instance, Dowling (1973) found a J-shaped serial position curve for recall of brief melodic phrases, demonstrating “the type of recency effect typically found in analogous short-term memory tasks with verbal materials” (p. 39). Similarly, Roberts (1986) tested recall of both music (melodic and harmonic sequences) and linguistic materials, once again generating the expected primacy and recency effects. She also identified a clear modality effect (superior auditory recall), but only for the linguistic materials. Roberts states, “Differential recall for written music and language might occur because written music has faster, more

---

91 Also consistent with verbal materials, Dowling (1973) found evidence highlighting the importance of both encoding strategies (see section 2.3) and rehearsal time.
direct access to its corresponding auditory representation in short-term memory” (p. 152). One reason proffered for this finding was that written music is also heard and therefore is more directly addressable as an acoustic code.

In a further effort to understand working memory mechanisms, Salamé and Baddeley (1989) investigated the unattended speech effect (obligatory access of speech into the phonological store) using conditions incorporating noise, vocal music and instrumental music. There were clear disruptive effects for unattended music, the vocal music causing greater impairment, especially for experienced musicians, than that occurring for instrumental music. They concluded that background music could disrupt verbal working memory, a finding that would likely extend to “other more ecologically important cognitive tasks that also use the phonological store” (p. 119). Of interest, noise did not gain access to working memory, prompting Salamé and Baddeley to propose the possibility of a filter that differentially allowed access to a range of acoustic cues. Moreover, because we can hear and remember sounds that are distinct from speech, they also reasoned that a separate acoustic store might be available for dealing with this material (refer to the Berz, 1995, model above). Despite these at least partially independent functions, Madsen (1987) found that it was not possible to simultaneously concentrate on music and verbal materials, “although it would appear that humans are capable of ‘timesharing’ quite effectively” (p. 325).

Taken together, established working memory effects have been reproduced for music materials across a range of studies (e.g., Bull & Cuddy, 1972; Holahan, Saunders, & Goldberg, 2000; Mondor & Morin, 2004; Taylor & Pembrook, 1984). What is more, 

---

92 Salamé and Baddeley (1989) included a wide range of instrumental and vocal music.

93 It was believed that having gained access, the language component of vocal music might cause added disruption, which may account for the differential effects for vocal and instrumental music.
it has been shown that access to working memory is also differentially subject to the qualities inherent in a musical stimulus: DeWitt and Crowder (1986) asked their subjects to rate pairs of melodies as same or different, each of the pairs separated by either a short or long delay (1 or 30 sec). Their results demonstrated that melodic contour was remembered better after short delays, and that interval information was better remembered after longer delays. In a similar way, Attneave and Olson (1971) established that an auditory image in short-term memory preserved specific pitch values, whereas the auditory image from LTM (familiar tune) “is encoded in terms based upon relations, or intervals” (p. 164). These results appear to implicate the integrated effects of working and long-term memories respectively. It is to the latter cognitive mechanism that we now turn.
2.3 Long-Term Memory, Schemas and Music

Functionally, there is little reason to believe that music cannot be described in relation to the same long-term memory (LTM) structures evident for other areas of cognition (see section 1.5). After all, understanding music requires, like all domains, a large store of declarative knowledge. Due to the experiential nature of this domain, episodic memory also plays a significant role in the encoding, retrieval and storage of information as semantic knowledge. By its very nature and practical emphasis (performing, composing and listening), music also relies heavily on procedural memory, especially for the real-time processing and integrated motor skills necessary for fluent performance. Moreover, experience (passive or active) over many years, ensures that adults “bring a large store of implicit knowledge to bear in listening to music” (Dowling, 1999, p. 620)\(^94\).

The perception of music is a “process that aims to transform a series of unconnected tones into an integrated mental representation in musical terms” (Povel & Jansen, 2001, p. 170). “The ultimate result of our ability to perceive, remember, conceptualize, and act on musical information is the formation of internal frameworks, or schemata, for representing and reproducing more complex musical knowledge” (Dowling & Harwood, 1986, p. 4). For example, Sloboda (1978; 1984) has shown that good musical sight readers organise materials into larger chunks, rely on familiar patterns or stylistic idioms, develop high levels of motor programming skill, and infer or anticipate information rather than use a literal decoding of printed material. Furthermore, when asked to write down a pattern of notes following brief exposure, musicians demonstrate a noticeable perceptual superiority over non-musicians (Sloboda, \(94\) Even young students appear to be more sophisticated listeners than is often believed; trained musicians can overestimate the expertise necessary for these students to carry out a task (Madsen & Madsen, 2002).
These results reflect both the findings of earlier perceptual studies related to chess and the critical importance of schematic knowledge in LTM (e.g., Chase & Simon, 1973a, 1973b; De Groot, 1978, see section 1.6.1).

The building of schemas is a natural process that occurs from the earliest age. A number of studies have investigated the ways in which infants, some as young as six months, perceive, organise and respond to musical stimuli. In one study, Saffran Loman, & Robertson (2000) found that infants could discriminate between novel and familiar musical passages, and that they preferred passages to remain within musical context, a bias not exhibited by a control group. Trainor, Wu, and Tsang (2004) also established that infants remember melodies over long periods of time, providing evidence that both surface (e.g., timbre, tempo) and structural (e.g., relations of pitches, durations) features are retained in LTM. Furthermore, infants rely on categorical distinctions (inter- and intra-), “a fundamental process by which subjects decrease the complexity and the diversity of the social or physical environment by organising it” (Melen & Wachsmann, 2001, pp. 325-326).

Schemas not only represent what has been remembered, they also help determine what will be remembered. A widely exploited example of this phenomenon is readily evident in the area of film music. Boltz (2001) presented subjects with three short ambiguous film clips and paired these clips with positive, negative, and neutral music episodes. Her major finding was “that music influenced the comprehension and memory of filmed material” (p. 444), each contrasting piece of music invoking an interpretative and contrasting framework (schema) and mood for the visual action (also

---

95 The period of retention lasted for at least two weeks following passive exposure to complex pieces of music.
96 The study also included a no music control.
see section 1.5.2, encoding and mood, Bower, 1981). Recall that items or events are more likely to be remembered if congruent with an activated or prevailing schema (see section 1.5.3, Pichert & Anderson, 1977); in Boltz’s study, schemas influenced subjects beyond the more obvious conveyance or manipulation of mood.

Schemas also determine how much will be remembered. Wallace (1994) examined whether music acts, in effect, as a mnemonic. She indeed found that music could render text more memorable under certain conditions: a melody that is simple, symmetrical, easy to learn, and repetitive across verses. Furthermore, Wallace excluded rhythmic elements as a major determinant of memorability; rhythmically spoken text was recalled no more accurately than text alone. Once a melody is stored in LTM it acts as a powerful cue for retrieval of associated text, and also facilitates encoding by affording a structural framework for semantic organisation (chunking). This organisation provides, along with the text, an integrated and enriched memory trace conducive to improved recall. These findings are further supported by the work of Rainey and Larsen (2002): Using popular folk songs, they noted that music acted as an effective mnemonic aid by providing a structure for learning, a retrieval process, and a distinctive memory record.

In line with earlier evidence (see section 1.5.2), encoding and retrieval strategies are significant determinants of success when recalling musical information from LTM. Schulkind, Hennis, and Rubin (1999) showed that older participants’ recall was superior for music of their youth than music from later in life (see section 1.5.2, Bahrick and colleagues, e.g., Bahrick, 2000). Notably, as a determinant of accurate recall, they established that there was a strong relationship between music from this earlier period
and participants’ emotional responses. As a further indicator of improved recall, Delis, Fleer, and Kerr (1978) stated, “one also has better memory for musical passages which have coherent ‘embodied meaning’” (p. 215). They used excerpts of music with either concrete (comprehensible) or abstract (difficult to comprehend) titles: Real world interpretations of music exerted an organisational influence on musical information, facilitating its assimilation with existing schemas.

All these findings support previous evidence (see section 1.5.2) that meaning, usually embedded within familiar contexts, is essential for improving retention and recall. For music, familiar contexts are defined by the typical relations that exist–and are anticipated–between musical elements across a number of musical works; the shared conception of these relations is commonly referred to as style. Bigand (1993) contends that these contexts and relations are necessary in order for a listener to “analyse pieces of music produced by the culture to which they belong” (p. 234), a process that results in a lexicon of schemas.

By this process, musical syntax and grammar is generated by stylistic traditions, and the product of these perceptual regularities are associated with schemas; for Western music this includes knowledge of scale structures, chord functions and key relations (Krumhansl, 1992). Following a study of how contextual factors affect the perception of two contrasting stylistic musical traditions, Kessler, Hansen, and Shepard (1984) reported that there “was less variation among both Western and Balinese groups of listeners when they heard contexts based on music of their own culture, and more within group variation to unfamiliar contexts” (p. 164), enculturation therefore having a

97 Nonetheless, recall (20%) was not all that resilient for music from earlier in the century.
98 Umemoto (1993) states that listeners exposed to a limited style of music will gradually develop specific cognitive structures associated with that style.
significant impact on perception.

Although the nature of schemas varies across cultures, they are often associated with fundamental or common qualities of music. For example, remembering melodies is a musical feature characteristic of all cultures (Dowling, 1978) and considerable research has been undertaken to determine the psychological constructs that underpin their perception. Dowling (1972) found that subjects retain contour (shape of relative pitch movements), and that exact interval size is lost. That is, melodies are thought not to be stored as a literal mental copy, but are likely “stored as a sequence of abstracted chromas” (Dowling & Harwood, 1986, p. 128)\(^99\). Moreover, melodies are the product of not only contour, but also mode, which act as a “framework on which the contour may be hung” (Dowling, 1978, p. 350).

Cuddy (1982) believed that tonality (a modal frame of reference) helps provide cues for “perceptual constancy of frequency ratio” (p. 3), therefore supporting the capacity to simultaneously transform musical elements and retain their musical identity\(^{100}\). She suggested the implications of tonality extended to temporal structure; for tonal music, “notes are not so much events in themselves as they are carriers of direction, progression or movement” (p. 3)\(^{101}\). Deutsch (1980) emphasised the notion of extracting hierarchical structures of tonal sequences that reinforce and help produce a sense of temporal segmentation, in the process alleviating the load on memory. Consistent with this view, Janata, Birk, van Horn, Leman, Tillman, and Bharucha (2002) state, “Central to our ability to hear music coherently are cognitive structures

\(^{99}\) Chroma is the quality of a pitch’s relative place within an octave (Dowling & Harwood, 1986).

\(^{100}\) See Peretz (1993) for a review of neuropsychological evidence supporting the existence of a tonal component in the organization of melodic processing.

\(^{101}\) Furthermore, both trained and untrained listeners seem to respond in qualitatively similar ways.
that maintain perceptual distance relationships among individual pitches and groups of pitches” (p. 2169).

Although music unfolds over time, extracting structural features allows the assignment of meaning to the otherwise temporal fluidity of musical events. Schemas permit a listener to segment music at different levels, the result of a cue abstraction mechanism that captures invariant features or similarities between musical events (Koniari, Predazzer, & Mélen, 2001). Deliège (2001) defines a cue as “a salient element that is prominent at the musical surface” (p. 237). On this basis, listeners are able to “cluster musical materials together into [meaningful] categories such as motives, themes and so on” (Cambouropoulos, 2001, p. 347). As a musical work progresses, listeners are left with a mean of the important features, a process Deliège (2001) refers to as imprint formation.

These processes of cue extraction and imprint formation are schema-based; musical materials can “give rise to different ‘schematas of order’ that are largely dependent on listeners’ previous musical training” (Deliège, Mélen, Stammers, & Cross, 1996, p. 155). In addition to these conscious (explicit) schema-based processes, there is evidence that automatic (implicit) processing may under certain conditions also continue even during the registering of real-time music, a process that gradually adds further detail to the memory trace (Dowling, Tillmann, & Ayers, 2002). Therefore, some aspects of musical thinking appear to exist outside conscious or conceptual awareness, constituting experiential representations of music (Torff & Gardner, 1999).

It is apparent that LTM and the development of schemas provide the foundation for higher order mental processes and, as a result, the skill acquisition that leads to expertise. These processes are examined in more detail in the following section.
2.4 Expertise, Skill Acquisition and Music

Considerations of what constitutes expertise and the ways in which it can be developed are just as central to the study of music as they are to other domains, generating comment from even the earliest of musicologists such as Johann Nikolaus Forkel (Riley, 2003). Although great musicians are accorded almost superhuman abilities (e.g., Mozart, Paganini), their extraordinary skills can still be explained by the processes that have been found to underpin expertise and human cognition at large (Sloboda, 1985). According to Sloboda (1985) these processes include the following: 1. A vast accumulation of musical knowledge that is associated with schemas and the extensive chunking of information. 2. A reliance on a large repertoire of automated motor-cognitive skills, which can be flexibly applied to a range of local problems. 3. The capacity to self-monitor performance. Each of these processes is briefly discussed below in relation to the activities—performance, composition, and listening—that characterise musical expertise.

In accordance with other domains, deliberate practice would be expected to account for a significant proportion of the variance that is evident between the application of novice and expert musical skills (see section 1.6.1). In fact, Ericsson et al. (1993) documented the importance of extended practice (10 or more years) across different domains, including the activities of violin performance, music composition (see Hayes, 1981), and piano performance. Even after extended periods, Ericsson et al. suggested that performance is “not rigidly limited by unmodifiable, possibly innate,

---

102 As applied in the modern sense of the term, Riley (2003) describes how Forkel (1749-1818) developed a theory of expert and inexpert listening practices using rubrics to categorise different types of listeners.

103 With respect to inexpert pianists, the amount of practice undertaken by experts was higher by a factor of 10 (Ericsson et al., 1993).
factors” (p. 366), however the same authors were also mindful that the differences within groups could be attenuated by “additional factors consistent with the skill-acquisition framework” (p. 380)\textsuperscript{104}. For instance, Williamon and Valentine (2000) after having examined the practice habits of pianists concluded, “the content and quality of deliberate practice must be examined before fully understanding the factors which affect the quality of a specific performance” (p. 373). These sentiments were reinforced by Madsen (2004), who found subjects’ memory for the extent of their past practice was poor, and only a weak relationship existed between levels of performance and amounts of practice.

The work of McPherson and colleagues has highlighted the importance of not only quantitative but also qualitative aspects of musical skill acquisition (McCormick & McPherson, 2003; McPherson, 2000, 2005; McPherson, Bailey, & Sinclair, 1997; McPherson & McCormick, 1999): McCormick and McPherson (2003) found a clear relationship between levels of self-efficacy and performance skill, an association they noted is consistent with other academic contexts. As the authors also make clear, self-efficacy “not only implies a self-recognition of being a good instrumentalist, but also explicit judgements about the skills necessary to perform in front of others” (p. 40). Costa-Giomi (1999; 2004), although finding that the contribution of music to cognitive development and academic achievement may be more limited than previously believed, also found that after three years of piano tuition self-esteem measures remained a significant factor.

McPherson et al. (1997) established that a range of musical skills (e.g., sight reading, playing by ear, playing from memory) not only contribute to skill acquisition,
but also that visual, aural and creative (improvisation) forms of performance are interdependent and necessary to coordinate ear, eye and hand. Even though, in relation to rehearsed music, practice was an important predictor of a student’s performance quality, as a measure, “mental strategies was consistently a more powerful predictor for explaining their ability to sight-read, play from memory and play by ear” (McPherson, 2005, p. 27). In short, “better student musicians are likely to possess a sophisticated repertoire of strategies that they use when performing and practicing” (McCormick & McPherson, 2003, p. 38). These strategies constitute appropriate cognitive\textsuperscript{105} and metacognitive skills or schemas (McPherson, 1994), which in the case of improvisation included, as the best predictor of performance, the capacity to self-evaluate (May, 2003). Furthermore, self-monitoring is crucial for effective practice (Woody, 2003), and consistent with the finding that “harder working musicians tend to report higher levels of Cognitive Strategy Use” (McPherson & McCormick, 1999, p. 101).

An extensive knowledge base (declarative and procedural) and the (problem-solving) strategies to apply this knowledge in appropriate ways was seen in Chapter 1 as a crucial and universal characteristic of expert performance. Not surprisingly, Colley, Banton, Down, and Pither (1992) found that an expert-novice comparison of composers revealed that experts possessed a larger knowledge base (e.g., stylistic features of a genre), and possessed a larger range of procedures. Conversely, novices’ recognition was likely to be more superficial, such as in the use of standard cadences. Colley et al. concluded that the novices were “still mastering basic technical skills while the Expert has learnt them sufficiently well to have automated them” (p. 136). Due to the ill-defined nature of composition, it is not surprising to find experts able to break a

\textsuperscript{105} See for example Gromko’s (2004) study involving predictors of music sight-reading ability.
problem down and place appropriate constraints on each task (McAdams, 2004), a strategy that relies on promoting the efficient application of cognitive resources.

Musical performance at high levels represents a complex problem, and in keeping with other domains, Chaffin, Imreh, Lemieux, and Chen (2003) confirmed that an expert pianist demonstrates a deeper rather than superficial understanding of the fundamental principles that define such problems (see section 1.6.2, Larkin, 1981; Larkin et al., 1980a, 1980b; Simon & Simon, 1978). In addition, they concluded that experts’ decision-making was quick and subjects used appropriate cues to anticipate global or future interpretative requirements. In essence, performers not only were capable of seeing the ‘big picture’, they also possessed the mental capacity to use this knowledge in strategic ways. Likewise, Chaffin and Imreh (1997) believed that pianistic memory makes use of multiple and flexible retrieval systems that can rapidly recall information from LTM. Their study purported to “provide the first evidence that the principles of expert memory apply to concert soloists” (p. 334, see Chase and Ericsson, section 1.6.1).

These LTM retrieval structures effectively expand cognitive capacity, resulting in the redirecting of mental resources to novel elements and the self-monitoring of real-time musical performance. Closely associated with expert memory, these same mechanisms have been identified across a wide range of domains and, as documented above, across a range of musical activities. As a further example, Sloboda (1984) has identified the principles of expert memory in the ability to fluently sight-read music. Noteworthy among these principles was what he referred to as “incontrovertible evidence that musical knowledge is implicated in sight reading” (p. 231)\(^{106}\), and that

\(^{106}\) For example, a sensitivity to superordinate structures or groups (Sloboda, 1984).
automation of perceptual processes reduces the load placed on working memory (Sloboda, 1976, see also section 1.6.3).

When taken together, there is substantial evidence across all musical activities that the cognitive mechanisms described in Chapter 1, and their application within the framework of cognitive load theory, described in Chapter 3, are similarly applicable to the instructional design of music materials. These issues are pursued in the empirical work reported in Chapters 5-9.
Chapter 3: Cognitive Load Theory

3.1 Introduction to Cognitive Load Theory

Cognitive load theory (Sweller, 1988, 1989, 1993) is firmly embedded in an understanding of human cognitive architecture and the way in which cognitive structures are functionally organised (Sweller, 2004). In recent decades, broad agreement has been reached regarding these characteristic structures (Sweller et al., 1998), which in turn has provided a more stable basis for hypothesizing effective ways in which to present instructional materials. Consistent with these trends, the principal aims of cognitive load theory (CLT) are to inform instructional design and generate novel applications (Chandler & Sweller, 1991), aims that already have been met with the establishment of many effective instructional techniques (e.g., Chandler & Sweller, 1992; Jeung, Chandler, & Sweller, 1997)\textsuperscript{107}.

Historically, many instructional designs and educational trends have developed without the recognition of, or adherence to, the basic limitations imposed by our cognitive architecture (Sweller, 1993). In contrast, a central tenet of CLT recognizes that the constraints on human information processing are severe and as a result have consequences for the presentation of instructional materials (Sweller & Chandler, 1994). Central to these constraints, when learning and problem solving, are the extremely limited capacity of working memory and the concomitant imperative to use available cognitive resources as efficiently as possible (Kalyuga, Chandler, & Sweller, 2001a). So important are these characteristic features of cognition that Sweller et al. (1998) stated, “Prima facie, any instructional design that flouts or merely ignores working memory

\textsuperscript{107} A number of these techniques are discussed in section 3.3 below.
Although working memory places severe restrictions on cognitive processing, CLT asserts that long-term memory affords the means by which available capacity can be used more efficiently (Sweller, 1993). As previously outlined (see section 1.5), the capacity of long-term memory is immensely large, holding a vast repository of knowledge in the form of schemas. Each schema constitutes a single element and, depending on the learner’s expertise, can incorporate substantial amounts of information. Furthermore, with continual practice these unitary processing structures become increasingly automated and in time consume almost no conscious processing capacity, a mechanism that largely directs skilled or expert intellectual performance (Kalyuga, Chandler, & Sweller, 1998). Considered together, CLT emphasises that the best instructional designs devote working memory resources to the acquisition and automation of schemas (Sweller, 1999). By incorporating otherwise multiple elements within a single schema, a primary source of cognitive load can be managed within working memory limits.

In relation to instructional design, Kalyuga et al. (1998) note that it is insufficient to solely focus on schema acquisition and automation without the corresponding objective of eliminating unnecessary cognitive load. Naturally, what does or does not constitute ‘necessary’ cognitive load becomes a crucial distinction within CLT. In addressing this issue, Paas, Renkl, & Sweller (2003a) identified in relation to instructional tasks three sources of cognitive load108, each of which additively contributes to total cognitive load: 1. *Intrinsic cognitive load* is inseparable from the nature and complexity of a task. It represents the essential elements that have to be

---

108 See section 3.2 for a detailed discussion.
learned and therefore have to be processed in working memory. 2. *Extraneous cognitive load* is superfluous to the processing of essential task elements and is imposed by the unintended or misdirected instructional aspects of presentation, organization and design. 3. *Germane cognitive load* represents any task activities that in order to construct and automate schemas entail investment of mental capacity in the deeper understanding of materials.

Traditionally, the recognition of cognitive load, its constituent parts and their interaction, has contributed little to theoretical understandings of instruction. Too often instructional design has been guided by factors such as visual elegance and common sense, or adhered to the dictates of periodically changing educational fads (Plass, Chun, Mayer, & Leutner, 2003; Sweller, 1999). As a result, it has been neither obvious which factors contribute to task complexity (Sweller & Chandler, 1994), nor which factors reduce or eliminate the deleterious effects of extraneous cognitive load.

An example of the instructional nexus between necessary and unnecessary processing loads was highlighted by Sweller (1988), who indicated that problem solving as a traditional instructional technique was generally assumed to be an effective strategy, and yet the “mechanisms required for problem solving and schema acquisition may be substantially distinct” (p. 284). Because of the impediment to schema acquisition it has been shown that solving problems can leave students unaware of the underlying structural features that determine understanding; problem solving can therefore decrease the likelihood of learning transfer to new problem contexts.\(^{109}\) (Sweller et al., 1982). In the absence of knowledge, learners often adopt search-based strategies such as means-ends analysis, which are an efficient means by which to

---

\(^{109}\) See section 1.6 for a comprehensive discussion of these issues.
achieve a problem goal (solution), but do not necessarily result in establishing the required forward-moving schemas essential for understanding. Moreover, search-based strategies impose a high and extraneous cognitive load that direct processing capacity away from schema acquisition and therefore away from learning (Chandler & Sweller, 1991).

To circumvent a natural tendency towards search-based strategies, Sweller and Levine (1982) employed goal-free problems, each incorporating a limited number of possible problem operators. Without a specific goal, subjects were not required to expend valuable processing capacity searching the problem space and reducing differences between sub-goals. Rather, each new problem move was generated from the previous move in a forward direction, thereby dedicating processing capacity to learning the relations between each move. This of course is precisely the activity required for schema acquisition. As predicted, studying goal free problems was superior to solving conventional problems. This result was attributed to the reduction of cognitive load otherwise induced by problem search, and the redirecting of processing capacity to the induction of schemas. Significantly, this somewhat counterintuitive instructional strategy was generated by CLT, reinforcing the theoretical principles upon which it was based.

Not only has CLT been used to explain why problem solvers fail to perceive features salient to learning and schema acquisition (Plass et al., 2003; Sweller, Chandler, Tierney, & Cooper, 1990), it has been successfully adapted to a wide variety of instructional materials that range from technical disciplines to language-based discursive areas. In addition, CLT has been used to explain a range of instructional phenomena, generate new instructional formats, and to predict the effectiveness of instructional designs, all of which are important justifications for the establishment,

In summary, the essential features of CLT are as follows: 1. It is “a theory that emphasizes working memory constraints as determinants of instructional design effectiveness” (Sweller et al., 1998, p. 251). 2. The primary aim of instruction is to direct limited working memory capacity to the construction and automation of schemas. 3. Information-rich (and automated) schemas accessed from long-term memory reduce the number of elements (intrinsic cognitive load) that require simultaneous processing in working memory. 4. Working memory capacity is maximised by reducing the unnecessary mental activities (extraneous cognitive load) associated with poor instructional design. 5. Excess working memory capacity (germane cognitive load) is redirected to activities that encourage deeper processing. The following sections discuss the nature of cognitive load in more detail and examine how it relates to the design of effective instruction.
3.2 Cognitive Load

3.2.1 Intrinsic cognitive load.

For any learner, all tasks involve an intrinsic degree of difficulty or, in cognitive terms, a mental load. Chandler and Sweller (1996) stated that intrinsic cognitive load is “determined by the mental demands or intellectual complexity of the task” (p. 153) and Paas et al. (2003a) suggested that it cannot be altered by means of instructional design. In other words, it is a constant between the nature of a task and the individual characteristics of the learner attempting to master that task. Fundamental to this relationship between task and learner is the fact that certain materials are inherently complex and require a considerable investment of mental resources, whereas other tasks appear effortless and straightforward.

Task difficulty is not always easy to identify, but it is closely related to the manner in which human cognition manages the processing of task elements (Sweller & Chandler, 1994). Recall that an element is a single item and represents, with respect to an individual learner, the simplest informational structure to be learned (Chandler & Sweller, 1996). That is, all new learning in some way involves the incorporation of one or more new elements. By extension, the number and nature of new elements associated with a task are crucial to determining task difficulty. From this analysis, it seems reasonable to conclude that intrinsic cognitive load should increase in direct relation to the total number of elements, however, as will be seen, the relationship between materials, learner and task difficulty is not so simply defined.

The difficulty associated with learning any task can be conceived in two ways (Sweller, 1993). The first of these is characterised by a large number of individual items (elements), each of which can be learned or understood without reference to other items. Take for example, the numerous performance directions used in music (e.g., maestoso
or \( \mu \)\(^{110}\); in nearly all cases these directions are perfectly comprehensible as discrete elements. Because each element can be readily held and processed in working memory, learning any one of these directions is relatively easy. Nonetheless, learning the sum total of all such musical items presents an enormously difficult task\(^{111}\), not because of our limited processing capacity, but because numerous items need to be retained over time in long-term memory (Sweller & Chandler, 1994). Although collectively materials such as these are difficult to learn, their intrinsic cognitive load is low because alone each item is intelligible without making heavy demands on processing capacity.

The second way in which materials can be difficult to learn arises surprisingly from conditions where relatively few elements are required for understanding. Some semantic aspects of language provide a suitable demonstration of this problem. Consider the following extract:

But some people must wait where other people cannot see the people who are waiting. And people who arrive before other people must wait where they cannot see the other people who arrive after them being admitted before them (Taken from the BBC TV series Yes Prime Minister. Lynn & Jay, 1989, p. 120).

Why do these relatively few words comprising only two sentences present such comprehension difficulty? Sweller (1993) proposed that for complex learning the elements interact, and therefore they are only intelligible when considered together rather than in isolation. In contrast to the previous music example, the total number of elements for this explanation is small, but they cannot be understood or learned in an independent fashion. Instead, some or all the elements must be considered

\( \mu \)\(^{110}\) This symbol describes a musical embellishment known as a mordent.

\( \mu \)\(^{111}\) Musical dictionaries contain thousands of entries, covering an extensive range of terms and signs relevant to the performance of music.
simultaneously, precisely because the understanding of relations between elements is the goal of learning (Sweller & Chandler, 1994). Because working memory capacity is limited, attempting to simultaneously assimilate multiple items expends considerable mental resources and as a result raises intrinsic cognitive load.

As can be seen, cognitive load theory makes a major distinction between the learning that arises from the mental integration of multiple elements, and the learning of elements in isolation\textsuperscript{112}. Isolated elements by the very nature of their limited processing demands are low in element interactivity and thereby impose a low intrinsic cognitive load. By comparison, materials that are unintelligible without first mentally integrating multiple elements are high in element interactivity, and hence high in intrinsic cognitive load. As Sweller (2003) asserts, “learning with understanding imposes a heavy cognitive load” (p. 218); each additional element that has to be simultaneously processed increases the intrinsic complexity of a task.

Viewed in this way, interacting elements represent the appropriate mental integration required to construct a schema, the building blocks of understanding and the basis of expert performance. In fact, “There may be no useful distinctions that can be made between interacting elements and schemas” (Sweller & Chandler, 1994, p. 190). By this reasoning, learning is a process of unifying multiple elements into a single schematic construct. As a consequence, learning with understanding is only relevant when dealing with high element interactivity materials (Marcus, Cooper, & Sweller, 1996).

Where the element interactivity of materials exceeds processing capacity, tasks can be simplified or elements can be initially delivered in isolation (Pollock, Chandler, \textsuperscript{112} That is, rote learning.)
& Sweller, 2002), but as previously noted, intrinsic cognitive load cannot be altered. Be that as it may, the same materials do not necessarily represent identical levels of intrinsic cognitive load for every learner (Sweller, 1993). Inevitably, the development of expertise leads to an increase in the number of domain schemas held in long-term memory and the automation of their inherent processes and actions (see section 1.6). Furthermore, a single schema can be developed or expanded to incorporate virtually indefinite amounts of information. Irrespective of whether schemas are either information-poor or rich, each represents only a single element in working memory. Consequently, acquiring and building schemas helps circumvent the limitations imposed by working memory. As a result, schema acquisition represents a reduction in the total number of elements to be processed and therefore a reduction in the level of intrinsic cognitive load (Sweller et al., 1998).

For experts, the simultaneous manipulation of three or four elements or schemas in working memory may, in contrast, represent what is for a novice overwhelming intrinsic cognitive load. Faced with such a dilemma, novices must find a way of reducing element interactivity—and its associated cognitive load—within processing limits. Although intrinsic cognitive load is fixed, other sources of cognitive load can be manipulated by instructional design. It is these sources of load that are considered in the following section.
3.2.2 Extraneous cognitive load.

If total cognitive load were a product of intrinsic cognitive load alone there would be little point in considering instructional design factors: Specifically, learning difficulty would exclusively arise from the intrinsic complexity generated by the materials. Such a position is not only intuitively implausible, there exists a wealth of evidence to highlight the importance of instructional design factors in facilitating learning (Kalyuga, Chandler, & Sweller, 1999; Moreno & Mayer, 1999b; Sweller, 1989; Sweller & Chandler, 1991; Sweller & Cooper, 1985). Accordingly, “In an educational context, the nature and degree of cognitive load will be determined, at least in part, by the format of instruction” (Sweller, 1999, p. 44).

It was outlined in the previous section that for learning to occur, the acute processing limits of working memory must be directed toward schema acquisition. Logically, the addition of any superfluous or extraneous mental activity may, in combination with intrinsic cognitive load, exceed working memory capacity and redirect cognitive resources away from this primary goal of instruction (Gerjets & Scheiter, 2003). Therefore, extraneous cognitive load results from any instructional design factors—format or activity—that are not essential for learning (Paas et al., 2003a); eliminating these factors permits the reallocation of working memory capacity to the pertinent tasks of learning and problem solving (Sweller, 1993).

In essence, extraneous cognitive load represents the additional mental effort of poor instructional design (Sweller et al., 1998). With this in mind, much of CLT research has been dedicated to generating alternative instructional designs that avoid extraneous processing and the interference this causes to learning (Gerjets & Scheiter, 2003). Still, identifying activities that generate extraneous cognitive load can be deceptively difficult (Marcus et al., 1996), however, the following examples briefly
illustrate the hidden cost of engaging in irrelevant processing when learning\textsuperscript{113}.

You may recollect that conventional problem solving often induces search-based strategies rather than the processes that facilitate the construction of schemas. In effect, searching the problem space is unrelated to the alternative process of negotiating a direct solution. In this sense, search-based strategies must be considered as extraneous activity and therefore extraneous cognitive load. In contrast, when conventional problem solving is replaced with \emph{worked examples} (see section 3.3.1), processing capacity is dedicated to constructing the necessary schemas that direct a forward-moving and efficient solution (Sweller & Cooper, 1985).

A further example is provided when two different, but necessary, information sources must be combined in working memory to facilitate understanding: A diagram and text represent one of the most common instructional contexts in which this occurs. Typically, these two sources are placed apart on a page and the learner must constantly move from one to the other in an effort to locate and integrate associated referents. As a result, this activity usually imposes a high extraneous cognitive load due to the additional processing required when holding and mentally integrating disparate sources of information (Jeung et al., 1997). Sweller et al. (1990) referred to this phenomenon as the \emph{split-attention effect} (see section 3.3.2) and found that when the appropriate referents (text and diagram) are physically integrated, processing capacity can be redirected to the essential aspects of learning.

In both these examples “cognitive effort is directed to a search process that is unrelated to learning...because of the way the material is structured” (Chandler & Sweller, 1992, p. 234). Because intrinsic and extraneous cognitive loads are additive,\textsuperscript{113} See section 3.3 for a detailed discussion.
the extent to which the latter interferes with learning is of course dependent on whether working memory capacity is exceeded (Sweller et al., 1998). Section 3.3 examines in greater detail possible sources of extraneous cognitive load and the means by which they can be avoided through instructional interventions. By reducing extraneous cognitive load, processing capacity can be concentrated on maximising schema acquisition, a feature that is central to the following discussion.
3.2.3 *Germane cognitive load.*

Within any learning context, understanding should not be conceived as a unidimensional concept that is achieved by a unitary act of cognition. As outlined in Chapter 1, schemas represent a knowledge continuum where all new learning contributes to increasingly knowledge-rich constructs in long-term memory. The extent to which we understand any new concept is governed by numerous factors that are germane to the activities of schema acquisition. That is, although certain instructional activities might increase the mental effort or *germane cognitive load* required by the learner, the trade off is deeper understanding\(^{114}\). This investment of processing capacity in schema induction is influenced by instructional design factors (Paas et al., 2003a), and by a learner’s self-monitoring or metacognitive strategies (Valcke, 2002).

According to van Merriënboer, Schuurman, de Croock, and Paas (2002), activities that raise germane cognitive load are “effort demanding, inductive processes that typically require ‘mindful abstraction’ from the learners” (p. 12). Accordingly, germane cognitive load suggests a level of processing beyond what is necessary to simply activate and integrate simultaneous elements in working memory (Gerjets & Scheiter, 2003). Bruning et al. (1999) believes that some consensus now exists regarding the deeper processing of complex information: It is a highly constructive process that views and manages information in a variety of different ways. This view is of course entirely consistent with the earlier discussion of long-term memory encoding strategies (see section 1.5.2), which emphasised the qualitative nature of tasks, and in particular semantic factors, as crucial to retention and understanding (Craik & Tulving, 1975).

\(^{114}\) This additional effort of course relies on available processing capacity.
Employing instructional strategies such as task elaboration highlights an increasing interest in the beneficial effects of germane cognitive load (Bannert, 2002). For instance, de Croock, van Merriënboer, and Paas (1998) used contextual interference—defined as contextual factors that interfere with skill acquisition—to study transfer performance when diagnosing complex system failures. They found that although subjects from a high contextual interference condition needed longer practice time and made more practice errors, their far transfer performance was superior. De Croock et al. explained this superiority in terms of the increased investment of mental effort and the resultant acquisition of applicable schemas; in effect the imposition of higher, but germane, levels of cognitive load.

Van Merriënboer et al. (2002) also concluded that high contextual interference yielded instructional benefits. These techniques included varying presentation, provoking group discussion and questioning at deeper levels, all of which provided viable alternatives when seeking to generate superior transfer performance. Notably, they stressed the twin goals of lowering extraneous cognitive load and replacing any excess capacity with activities that raise germane cognitive load. Naturally, this inverse relationship is only relevant where there is sufficient capacity remaining between intrinsic and total cognitive loads.

From a constructivist standpoint, Valcke (2002) rhetorically enquired how CLT accounts for the ways in which learners construct external representations: He believed that the more recent emphasis on germane cognitive load may help to direct

---

115 For this study the high contextual interference condition had to apply a new case type procedure for each diagnosis, whereas the low contextual interference condition used a block design whereby the same case type could be reapplied to each new trial.

116 These comments were offered as part of a review that examined theoretical developments in relation to CLT.
processes in ways that optimise training efficiency. Valcke also made a case for metacognitive load as a subset of germane cognitive load, the self-monitoring of cognitive processes influencing the critical interaction between short- and long- term memories. He cites the study of Stark, Mandl, Gruber, and Renkl (2002) in which the difference between surface and deep levels of example elaboration may indicate germane metacognitive load, although he acknowledges that the measuring of such load may prove problematic.

Clearly, any cognitive activity that contributes to schema construction and a deeper and developing understanding of materials should be encouraged. For this reason, cognitive load should not be viewed as a variable ‘volume’, but rather the nature of the task and the learner’s expertise must be balanced against the instructional interventions that maximise cognitive effort and minimise cognitive inefficiency. Short of instructional trial and error, or for that matter intuition, the degree to which cognitive load can be identified and manipulated rests with the extent to which it can be measured, the topic of the following discussion.
3.2.4 Measuring cognitive load.

In the preceding sections, the concept of cognitive load was discussed in some detail, and defined as a multidimensional construct (Paas, 1992; Paas & van Merriënboer, 1994a) that is central to generating instructional applications within CLT. To this end, implicit manipulations of cognitive load have yielded several robust and effective instructional techniques over nearly three decades, yet this “key concept…has largely eluded our attempts of direct measurement in authentic learning situations” (Brünken, Plass, & Leutner, 2003, p. 53). Without reliable measurements of cognitive load, the theoretical predictions upon which findings are based remain open to alternative interpretations and to potential confounding by unrelated factors.

The measurement of mental effort is considered the core element of cognitive load117 (Paas, 1992) and can be divided into three broad categories (Eggemeier, 1988): Taking into account recent developments, these categories consist of performance- and task- based measures, subjective measures, and psycho-physiological measures. Brünken et al. (2003) conceived the measurement of cognitive load as an interaction between two dimensions (Figure 9): On one dimension, cognitive load can be classified as either direct or indirect, and on the other dimension, classified as either objective or subjective.

![Classification of Methods for Measuring Cognitive Load Based on Objectivity and Causal Relationship](image)

**Figure 9.** Dimensions for measuring cognitive load (taken from Brünken et al., 2003, p. 55).

---

117 Paas (1992) states, “the intensity of effort is considered to be an index of cognitive load” (p. 429).
Before 1992, the measurement of cognitive load relied almost entirely on performance- and task- based measures (Sweller et al., 1998), and they remain the most common means of determining differences between two or more groups following instructional interventions. These measures are more commonly derived from a primary task, or from either the primary or secondary task within a dual-task paradigm. Differences between primary measures are usually represented by either, or a combination of, correct number of solutions (e.g., Yeung, Jin, & Sweller, 1997), study or solution times (e.g., Cooper, Tindall-Ford, Chandler, & Sweller, 2001), or number of errors (e.g., Ayres, 1993).

With respect to CLT, the wide employment of indirect measures has proved invaluable for supporting theoretical assumptions and for generating associated instructional applications. However, in a more recent article, Brünken et al. (2003) made a detailed case for the dual-task approach as a direct and objective measure of cognitive load, arguing that this technique overcomes some of the inherent weaknesses associated with other indirect subjective techniques. Although commonly employed for working memory research, the dual-task or secondary technique has been applied only spasmodically to the measurement of cognitive load (e.g., Brünken, Steinbacher, Plass, & Leutner, 2002; Chandler & Sweller, 1996; Marcus et al., 1996; Sweller, 1988). It is a measure of cognitive capacity in which “decreases in performance on a secondary task are used to measure increases in use of cognitive capacity on a primary task” (Britton, Glynn, Meyer, & Penland, 1982, p. 52).

Task-based indicators represent yet another method of measuring cognitive load, and include an estimate of the relations or processing steps required to understand a specific task. By calculating the number of individual elements that need to be simultaneously considered in working memory, these indirect measures attempt to
objectively quantify the intrinsic difficulty posed by a learning activity. Recall that the extent to which elements interact is crucial for predicting whether instructional materials can be processed within available working memory capacity. From this perspective, estimates of element interactivity highlight processing complexity and the mental effort required to complete a task.

These task-based estimates, in combination with other measures, are commonly used to support experimental predictions and/or when hypotheses involve a comparison of low- and high-element interactivity materials (e.g., Chandler & Sweller, 1996; Tindall-Ford, Chandler, & Sweller, 1997). They have proved a useful way of reflecting required mental effort, although as a single estimate for each task they assume that students possess equivalent levels of domain knowledge. Naturally, for even very homogeneous populations, the extent to which schemas have been acquired, and therefore instructional elements will interact, is subject to at least some individual variation. Nonetheless, in relation to randomised treatment conditions, domain knowledge (on average) is assumed to be equal.

Some measures of cognitive load have sought to consider in more detail both task and learner characteristics. For instance, as an estimate of the respective cognitive loads generated by goal-free and goal-specific problems, Sweller (1988) developed a computational model (Appendix G.) that tracked the number of productions required in order to reach a goal state\textsuperscript{118}. He defined a production system as “a set of inference rules that have conditions for applications and actions to be taken if the conditions are satisfied” (p. 264). The computational model accounted for the number of elements within a single working memory, and Sweller cautiously proposed that these models

\textsuperscript{118} For this study, calculating the number of steps for goal-specific problems was assumed to incorporate a means-ends analysis process.
might have more general applicability.

By far the most common technique for measuring cognitive load has been the subjective rating scale pioneered by Paas (1992). It asks learners to introspect on the perceived intensity of mental effort expended during a task\textsuperscript{119}. According to Gopher and Braune (1984), learners are quite capable of self assessing and rating their cognitive processes within a relative scale of perceived difficulty, concluding “that psychophysical scaling may provide a powerful measurement approach to quantify the subjective experience of workload” (p. 529). Typically, these scales provide gradations of seven - (e.g., Kalyuga et al., 1998; Marcus et al., 1996) or nine- (e.g., Paas & van Merriënboer, 1994b; Tabbers, Martens, & van Merriënboer, 2004) statements\textsuperscript{120}, and require students to indicate the statement that most accurately reflects the difficulty they experienced while completing the task.

Paas et al. (1994) closely examined the rating-scale technique and found that as a unidimensional scale it is a sensitive and reliable measure of cognitive load\textsuperscript{121}, and that it is “useful for obtaining information on the invested amount of mental effort in instructional research” (p. 428). Moreover, they believed that this type of scale is easy to use in the classroom and constitutes an important supplement to primary performance measures. Rating scales are a widely applied technique, although, in common with other methods, rely on satisfying certain key assumptions if measures are to be considered accurate. For example, subjective ratings do not necessarily indicate the specific nature of cognitive load (intrinsic, germane or extraneous); a high rating potentially represents

---

\textsuperscript{119} See Appendix B for the rating scale used in Experiment 1, and section 5.2 for an explanation of its use within this experiment. This subjective rating scale was used for all experiments reported in chapters 5–9.

\textsuperscript{120} A typical scale may range from \textit{extremely easy to understand}, to \textit{extremely hard to understand}.

\textsuperscript{121} Cronbach $\alpha = 0.90$ and 0.82.
either the negative effects of extraneous and/or intrinsic cognitive load, or the positive
effects of germane cognitive load. Likewise, a low rating might also reflect the
deliberately reduced mental effort of a task that is perceived as far too difficult by the
learner.

Based on Ahern and Beatty’s (1979) conception of efficiency\textsuperscript{122}, Paas and van
Merriënboer (1993) devised an innovative way of unifying rating scale and performance
data into a single measure (Figure 10). They referred to this technique as \textit{relative
condition efficiency}, and defined it as “the observed relation between mental effort and
performance in a particular condition relative to a hypothetical baseline condition, in
which each unit of mental effort equals one unit of performance” (p. 739). Accordingly,
performance measures and cognitive load subjective ratings are standardised (\textit{z} scores)
in order that a unit of performance can be equated with a unit of cognitive load.

$$E = \frac{z_{\text{Performance}} - z_{\text{Mental Effort}}}{\sqrt{2}}$$

\textit{Figure 10.} \hspace{1cm} Formula for the calculation of relative condition efficiency ($E$).

Relative condition efficiency helps overcome the potential problems of
differential performance, yet equal levels of actual load, or likewise, differential levels
of actual load, yet equal performance. If a learner’s position within the group (all
experimental treatments) remains unchanged across the two measures, their relative
efficiency is seen as neutral or in equilibrium. In contrast, the higher a learner’s relative
performance and the lower their relative cognitive load rating, the higher is their

\textsuperscript{122} Their efficiency view “is consistent with the notion that the component processes of any task might
individually require less capacity if they are better learned or more automatic” (Ahern & Beatty, 1979, p.
1290). This view is fully compatible with the principles of CLT.
instructional efficiency\textsuperscript{123}. This method has been widely applied to CLT research (e.g., Camp, Paas, Rikers, & van Merriënboer, 2001; Carlson, Chandler, & Sweller, 2003; Cerpa, Chandler, & Sweller, 1996; Kalyuga, Chandler, Tuovinen, & Sweller, 2001b; Paas & van Merriënboer, 1994b) and, given the reliability of objective performance and subjective rating-scale measures, remains a dependable and straightforward method of assessing cognitive load under a range of instructional conditions\textsuperscript{124}.

More recent technological advances have helped promote psycho-physiological techniques as a means of measuring cognitive load, and current indications suggest that some of these measures may eventually play a more prominent role within instructional research. These measures encompass a range of both physiological and neurological techniques, however up until now their impact within CLT has been minimal, due in large part to the impracticality and intrusiveness of these methods, inaccessibility of equipment, and the poorer reliability of measures. A brief review of these areas follows, although a broad discussion of techniques is well beyond the scope of this thesis.

Physiological measures include heart rate, brain activity (event-related potentials) and eye activity (blink rate, papillary dilation), all of which have been used as indirect objective measures of cognitive load (Sweller et al., 1998). They are based on the assumption that changes in one of these physiological factors are related to changes in the mental effort expended on a particular task. In one evaluation of physiological measures, Paas et al. (1994) examined spectral analysis of heart rate variability, in which the time between heart beats was determined by three different feedback mechanisms. One of these, variation in blood pressure, is linked to mental

\textsuperscript{123} That is, higher levels of achievement (assessment) are gained at a lower ‘cost’ or level of mental effort.

\textsuperscript{124} As an indicator of cognitive load, relative condition efficiency has been widely reported in the empirical chapters below.
effort and can be isolated within the mid-frequency band of measurement (from 0.07-0.14 Hz). Although such targeted measurements appear attractive as an accurate assessment of psychological states, Paas et al. concluded that both the reliability and sensitivity of this measure are low, and “that this technique is not appropriate for application in instructional research” (p. 429).

Task-evoked papillary response refers to another physiological measure related to cognitive load. Small changes in the size of a pupil are observed in response to processing loads presented by tasks of varying difficulty. Ahern and Beatty (1979) found that this technique appears “to reflect differences in central, rather than peripheral, brain processes” (p. 1291), and therefore provides an index of processing load when subjects are attempting a task. Their results also indicated that papillary changes for more able students suggested a decrease in brain activation rather than supporting alternative explanations such as the capacity to increase mental effort or mental power. Though not directly related to CLT, these findings support the notion that superior performance is associated with the reduced mental load of acquiring and automating domain-based schemas125.

A more promising line of research has emerged from direct and objective neuroimaging technologies (Brünken et al., 2003) such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). These technologies “allow researchers to observe the degree to which a brain region in a neurologically intact individual is activated by a task, so that its contribution to task performance under normal circumstances can be directly observed” (Banich, 2004, 125 In fact, Ahern & Beatty’s (1979) study set out to examine “the relation between ‘intelligence’ and capacity demands during mental activity” (p. 1289), however they concluded that the “present data do not bear upon the…neurophysiological basis of intelligence” (p. 1291).
pp. 79-80). Even though these techniques are more usually employed in areas such as working memory research (Smith & Jonides, 1997), it is conceivable that their applications will continue to grow in cognitive neuroscience. The difficulties with these methods at present relate to the practicalities of use and inaccessibility of expensive and sophisticated equipment. Certainly, their use within a more realistic classroom or instructional setting is unlikely in the foreseeable future.

We turn now to a number of instructional techniques that have used CLT to design appropriate interventions, and which have incorporated various direct and indirect measures of cognitive load to support claims of improved instructional efficiency.

\[\text{126 The application of these technologies is often awkward and therefore places restrictions on the subject that are unrelated to the normal task and performance environments.}\]
3.3 Instructional Implications and Applications of CLT

Together, the instructional techniques described below provide substantial evidence for the theoretical claims made by CLT. In each case, the theory was used to establish viable applications and, in turn, generate compelling explanations for their instructional efficacy. Supported by a variety of cognitive load measures, there now exists substantial evidence for both the theoretical interpretations, and the generalizability of CLT to other disciplines. In large part, this thesis has built upon the techniques reported below, and the current sound empirical basis of CLT, in order to apply theoretical principles within the instructional context of music. In view of this research goal, the instructional effects of split attention (see section 3.3.2) and dual modality (see section 3.3.5) are given special emphasis; this approach is in accord with the nature and procedures of the experiments reported in chapters 5-9.
3.3.1 Worked examples.

Earlier (see section 3.1) it was pointed out that under some circumstances learning was better facilitated by goal-free rather than conventional problems. This counterintuitive theoretical application arose from the work of Sweller and colleagues (Owen & Sweller, 1985; Sweller et al., 1982; Sweller et al., 1983), which recognised that search-based strategies associated with solving conventional problems failed to promote schema acquisition and, as a consequence, minimised learning. In response to these early insights, CLT was used to generate instructional techniques that fulfilled two primary aims: First, to identify procedures that reduced unnecessary cognitive load, and second, to ensure that cognitive resources were directed towards the acquisition and automation of schemas.

Worked examples appeared to be a technique that satisfied both these principles; in effect, worked examples provide an expert model laid out as a step-by-step procedure for solving specific problems (Atkinson, Derry, Renkl, & Wortham, 2000). Traditionally, worked examples are used as an adjunct to other instructional materials, and usually provide a limited demonstration of a to-be-learned concept, which is then followed by a series of conventional problems. In many disciplines, initial worked examples are considered an essential preliminary step for grasping new concepts. Indeed, evidence suggests that for students, worked examples are more compelling than written instructions (LeFevre & Dixon, 1986). Although the instructional application of worked examples has enjoyed a long history, their use within CLT departed quite radically from their previous sparing application as instructional examples. Instead,

---

127 Recall that reducing extraneous cognitive load effectively maximises available processing capacity (see section 3.2.2).

128 Note that conventional problems require a solution without the learner receiving any direct assistance.
large numbers of worked examples were now advanced as an alternative to problem solving and as a means of promoting initial schema acquisition (Sweller, 1999).

As a primary method of schema induction, Sweller and Cooper (1985) applied the principle of worked examples to students learning algebra. Learning materials were divided into introductory and acquisition phases, students given either a series of conventional problems, or a series of worked example-conventional problem pairs. Notably, when studying a large number of worked examples, problem pairs were believed to be a necessary inclusion in order to avoid either demotivation or waning engagement with instructional materials. Following the application of this strategy, the predictions for the worked example condition were largely realised: Students studied the acquisition problems for substantially less time and yet solved the subsequent test problems more rapidly.

These findings supported the assertion that worked examples help novice learners acquire the necessary schemas to successfully solve problems; Sweller and Cooper (1985) noted though that the advantages did not extend to dissimilar problems. Nevertheless, they believed that the evidence gathered across their five experiments confirmed the instructional value of worked examples and “the specificity of schemas used in knowledge-based problem solving” (p. 83)\(^{129}\).

According to CLT, the application of schema-based knowledge is not only domain-specific but, to varying degrees, also demands cognitive resources until fully automated (see section 1.6.3). This process of automation is relatively slow and quite gradual, a continuum upon which certain activities eventually draw upon little more than minimal processing capacity. This considered, Cooper and Sweller (1987)

\(^{129}\) They also state (following Experiment 4), “the schemas acquired were easily forgotten” (p. 83).
hypothesized that although initial schema acquisition would be expected to facilitate the solving of similar problems, dissimilar problems would exceed processing capacity until the necessary problem rules were also automated\textsuperscript{130}. They argued, “because problem-solving operators are automated, the problem solver has greater cognitive capacity available to deal with those aspects of the new problem that are unfamiliar” (p. 348).

To achieve at least partial automation of schemas, Cooper and Sweller (1987) simplified the earlier requirements by reducing both the number of problems and the number of problem rules required for solution. Following this modified approach, they found that large performance differences favoured the worked example condition, but this time for dissimilar (transfer) rather than similar problems. These results effectively reversed those found by Sweller and Cooper (1985), a predictable empirical pattern when the degree of schema automation was controlled by factors such as the complexity of materials and the repetitive use of worked examples. It was assumed that the resultant automated schemas redirected increased processing capacity to manipulating unfamiliar elements, precisely the cognitive precondition necessary when learning from high element interactivity materials.

Further evidence for the instructional application of worked examples was provided by Zhu and Simon (1987), who conducted a number of field experiments\textsuperscript{131} using mathematics materials in which the exclusive use of worked examples (no direct instruction) was compared with regular classroom teaching. They reported that the worked example groups were as, or more, successful than the conventional instruction groups, and that understanding was achieved in a shorter time. That advantages for

\textsuperscript{130} The term similar refers to \textit{isometric} problems, and the term dissimilar to \textit{transfer} problems. Transfer problems also lie on a continuum that ranges from near to far transfer of learned concepts.

\textsuperscript{131} In one example, three years of curriculum materials were completed in two years.
worked examples were evident, despite the stark contrast in teacher input, is a strong recommendation for their practical application. Clearly, the extended period within which, not to mention the authentic learning setting over which, this study was conducted strongly supported worked examples as an effective instructional strategy.

Carroll (1994) sought to extend research of worked examples beyond either laboratory findings, or those environments that would not be expected to necessarily generalize to an urban U.S. high school\(^{132}\), in the process acknowledging CLT as a primary rationale for the research findings that had been generated to this point. For Carroll’s experiments, students studied how to translate word problems into algebraic equations by completing either conventional problems, or a mixture of conventional problems and worked examples. The results of the two experiments were fully consistent with the studies reported earlier: Of note, Carroll emphasised that the advantages of worked examples were equally applicable to low ability students, including some students classified as learning disabled.

Advantages of worked examples, in particular those for lower ability students, were reinforced by the work of Quilici and Mayer (1996), who asked students, following instruction, to sort statistics word problems into appropriate categories. Not only did worked example groups perform at a significantly higher level than their no-example counterparts, they were more likely to categorise problems according to structural rather than surface features. Awareness of problem structure was indicative of deeper understanding and of a greater capacity to sort problems into appropriate categories. Unsurprisingly, they also found that this attention to problem structure was further enhanced by worked examples that emphasised those very features.

\(^{132}\) As represented by the Sweller and colleagues’ experiments (above), and the Zhu and Simon (1987) experiments respectively.
Significantly, lower ability students demonstrated the greatest gains, presumably because worked examples furnished the necessary cognitive support when available schemas were inadequate to manage the required number of problem elements.\(^\text{133}\)

These studies (Carroll, 1994; Quilici & Mayer, 1996) indicated that worked examples were more likely to benefit learners when tasks imposed high levels of cognitive load (also see Pillay, 1994, orthographic projection tasks); hence, the obvious benefits to those subjects with lower levels of domain knowledge. Somewhat related, van Gerven, Paas, van Merriënboer, and Schmidt (2002) found that worked examples were similarly effective for elderly learners who had experienced age-related declines in cognitive abilities. They concluded, “with regard to the elderly, worked examples make a more efficient use of the available working-memory capacity than conventional problems” (p. 103). As a result, they found that, in relation to a younger cohort, the elderly demonstrated greater gains in training efficiency and training time. In keeping with the interpretations of previous studies, it was believed that high levels of cognitive load were, to some extent, ameliorated by the use of worked examples.

Clearly, when designing effective instruction, the critical relations between worked examples, domain knowledge and resultant cognitive load must be fully considered. To address this issue, Renkl and Atkinson (2003) proposed a fading procedure that gradually reduced the reliance on worked examples. They assumed that increasing levels of skill acquisition lowered cognitive load such that problem-solving activities no longer exceeded mental capacity. In this way, search-based strategies (extraneous cognitive load) are gradually replaced with schema-based problem solving, which effectively raises the productive or germane cognitive load of these learning

\(^{133}\) These findings also reaffirm the centrality of schematic knowledge in relation to the development of expertise (see section 1.6).
tasks. In order to support this transition, Renkl and Atkinson incorporated self-explanation activities during the intermediate stages of fading as a means of students explicating important problem elements.

Worked examples not only facilitate schema induction, their effectiveness is based upon the qualitative way in which they engage the cognitive resources of learners. By also incorporating self-explanations, Chi et al. (1989) found that Good\textsuperscript{134} students were able to process worked examples more strategically to justify and refine their actions, find relations, use specific references, and monitor comprehension. In effect, these are activities that potentially optimise germane cognitive load in order to build deeper understanding, a goal that relies on engaging excess capacity in meaningful ways. For this reason, self-explanation strategies raise germane cognitive load and therefore comprise but one method of predicting effective learning and problem-solving transfer from worked examples (Renkl, 1997; Renkl, Stark, Gruber, & Mandl, 1998).

Other techniques in relation to worked examples have also generated ways of increasing germane cognitive load: Paas (1992) tested a completion strategy in which one group was given partially, rather than fully, worked-out examples. He found that presenting worked examples in both partial and regular formats was superior to conventional problem solving, supporting the hypothesis that these instructional techniques provided a more cognitively “efficient knowledge base for solving transfer problems” (p. 433). In essence, completion problems provide an efficient transition from study to application, the increased mental effort (cognitive load) and repetition of this strategy facilitating a deeper and more automated schematic base for future learning and problem solving. In a similar approach, Paas and van Merriënboer (1994b) found

\textsuperscript{134} Good and Poor students were “defined post hoc, using their problem-solving successes” (Chi et al., 1989, p. 158).
that a high variability condition of worked examples likewise encouraged productive mental activities conducive to schema acquisition\textsuperscript{135}.

As can be seen, the application of worked examples has been empirically supported across numerous studies; nonetheless, not all research has found this technique superior to conventional problem solving. Charney, Reder and Kusbit (1990) found problem solving more effective when learning a spreadsheet computer application (VisiCalc) than either tutorial instruction (including worked examples), or a learner exploration condition (discovery learning). Subjects in the problem solving condition devoted more time to training than the other conditions, “but produced faster and more successful performance at test” (p. 335), the authors noting that differences may have been due to the amount of expended effort. Even so, these results not only contradicted the findings previously reported, they also opposed a CLT interpretation that suggests problem solving and discovery-based learning impose excessive levels of cognitive load\textsuperscript{136}.

In response, Tuovinen and Sweller (1999) also compared direct instruction (worked examples-problem solving) with discovery learning (problem solving-exploration), however they employed equalised learning times, a confounding variable identified with the previous study\textsuperscript{137}. On this experimental basis, Tuovinen and Sweller found worked examples superior, but only for students who had no previous experience

\textsuperscript{135} High variability raises cognitive load, however, when incorporated within a worked example condition, the load is germane to a deeper understanding of the way in which new and old problem elements can be combined (Paas & Van Merriënboer, 1994b).

\textsuperscript{136} Similar studies (Charney & Reder, 1986; Charney, Reder, & Kusbit, 1986) also found that solving problems was superior to guided practice and the independent goal setting of exploratory learning.

\textsuperscript{137} Charney et al. (1990) also acknowledged training time as a possible explanatory factor. Otherwise, the methodology of the two experiments was very similar.
using a database program. They explained these findings by suggesting that with sufficient domain knowledge, the cognitive efficiency of worked examples becomes irrelevant, established schemas providing both a basis for further exploration and the means for lowering excessive levels of cognitive load. Such an explanation is in keeping with Charney et al. (1990), who noted, the “poor performance of our exploration group…may be due to the subjects’ minimal prior knowledge of the domain” (p. 337).

To summarise, worked examples are an instructional technique applied and modified according to the principles of CLT. They reduce cognitive load by facilitating mental processing germane to the building and automating of schemas. Where learners already possess a knowledge base sufficient for understanding, worked examples may raise cognitive load and reduce the degree to which mental resources can be devoted to further exploration of the problem space. Notwithstanding the important contribution worked examples have made to our practical and theoretical understanding of instructional design, they are also subject to a broader principle, one which strongly emphasises the avoidance of extraneous cognitive load. This issue is dealt with in the following section.
3.3.2 Split-attention effect.

Worked examples proved an effective alternative to conventional problems because they redirect mental resources to the essential cognitive activities associated with learning. By reducing sources of extraneous cognitive load, sufficient working memory capacity is made available for acquiring schemas. The demonstrated advantages of worked examples were generated by CLT, which was based on the relatively recent advances that have been made in cognitive science (Sweller et al., 1990). Based upon these same principles, it was reasoned that problem search (e.g., means ends analysis) was unlikely to be the only source of extraneous cognitive load, the evidence for alternative sources unexpectedly arising from subsequent investigations into the effectiveness of worked examples.

Tarmizi and Sweller (1988) failed to find any advantage for worked examples over conventional problems when they applied this technique to formats dealing with Euclidean geometry. The conventional format, of diagram followed by problem statements, contained two separate sources of information in which neither was intelligible without the other. By the very nature of this design, the learner was forced to hold in working memory elements from one source (statements), while searching for appropriate referents from the other source (diagram). As a result, understanding was contingent on the mental integration of two otherwise discrete sources of information. It was therefore reasoned that this additional processing would quickly exhaust working memory capacity.

With this in mind, Tarmizi & Sweller (1988) proposed that the learning of these geometry materials would be assisted by the reduction of search normally associated with displaced, but mutually referring, elements. By physically integrating problem statements with the geometric diagram, the materials were reformatted into a unitary
design, avoiding the extraneous cognitive load that results from otherwise splitting attention between diagram and text. Over a series of experiments, Tarmizi & Sweller established the superiority of these integrated worked examples, and demonstrated that learning from the traditional split-source formats was no better than learning from conventional problems. They concluded, “the critical factor is not the surface format of the problem and its presentation but rather the deeper, cognitive implications of its presentation format” (p. 435). Accordingly, the split-attention effect forced learners to mentally recast materials in order that they assumed a form more suitable for processing (Sweller, 1999). Unfortunately, the price of such mental reformatting is frequently too high, diverting limited processing capacity away from activities essential for learning. On the other hand, by physically integrating split-source materials, related information components are efficiently processed as single elements.

From this work with geometry materials the mental activity associated with split attention was identified as unrelated to learning. As a result, Ward and Sweller (1990) sought to extend the idea that instruction must be consistent with lowering cognitive load and directing attention to problem states and their associated moves. They tested this hypothesis with worked examples in the areas of optics and kinematics, the latter a discipline (physics) where previously the use of worked examples had generally been unsuccessful. Moreover, these studies were carried out in an authentic instructional setting, a normal classroom environment with a regular class of students\textsuperscript{138}. The use of integrated formats as worked examples proved highly effective under these conditions, reinforcing the theoretical principles that were used to explain the split-attention effect. In addition, Ward and Sweller suggested that the reformatting of material in line with

\textsuperscript{138} The success of worked examples under these conditions is also supported by the earlier reported work carried out by Zhu & Simon (1987, see section 3.3.1).
cognitive processes is not only essential, but also achievable for most split-source materials.

At first, the split-attention effect was pursued and confirmed with the use of both worked example and conventional problem formats. If the explanation for this effect was to remain theoretically viable, it was argued that similar advantages should also be observed with explanatory materials. Sweller et al. (1990) remarked that explanatory instruction often required students to simultaneously attend to multiple sources of information. Once again, these materials often assumed the form of text and diagrams. Testing some of these principles in the areas of coordinate geometry and numerical control programming, Sweller et al. found powerful effects that favoured integrated formats, irrespective of whether split attention occurred in either explanatory or problem phases.

Clearly, physically integrating mutually referring instructional elements reduces the processing burden on working memory. Nevertheless, integration was believed to be necessary only when two or more individual elements were unintelligible unless processed simultaneously. Where elements could be understood in isolation there was little theoretical justification for artificially integrating split-source materials. Chandler and Sweller (1991), in a study extending over a 12 week period, explored this assumption using materials dealing with the wiring of electrical light circuits. Consistent with previous findings, the positive effects of integration persisted over this prolonged period. In contrast, no advantage was observed for integrated elements that could otherwise be learned in an independent or self-sufficient fashion (Figure 11). In this case, bringing these elements together simply generated additional, but irrelevant,

---

139 Explanatory materials introduce fundamental concepts and/or procedures, and they usually precede the use of worked examples and conventional problems.
mental activity. According to CLT, any additional, but unnecessary, processing increases extraneous cognitive load\(^{140}\).

**Internal wiring for intermediate switching**

1. The active wire goes from the active to the common of switch 1
2. In this type of switching we use an additional switch called the intermediate switch
3. The wires connecting switch 1 to the intermediate switch and to switch 2 are called strap wires
4. The switch wire goes from the common of switch 2 to the light
5. The neutral wire goes from the light to the neutral
6. Under no circumstances is the gauge of wire used in this type of circuit to be broken

*Figure 11.* The diagram was self-sufficient for understanding. Accordingly, no advantage was gained by integrating the diagram and accompanying statements (conventional group, Chandler & Sweller, 1991, p. 305).

To provide further evidence for the increased cognitive load associated with split attention, Chandler and Sweller (1991) designed an elegant experiment in which they artificially manipulated the integration of split-source materials. Two groups were presented with identical materials in which the electrical wiring diagram could be fully understood without the accompanying textural information. The implicit group was simply told to study the instructions. Alternatively, the explicit group was directed to

\(^{140}\) See the Redundancy effect, section 3.3.3.
mentally integrate the textural and diagrammatic materials, a task unnecessary for understanding. The explicit group performed at a lower level and took approximately twice as long to study the materials, a result that supported a CLT explanation. Although unnecessary, it was argued that the attempt to mentally integrate displaced sources of information expended working memory resources otherwise required for understanding the essential aspects of the materials. By comparison, it was assumed that the implicit group directed their cognitive resources to the self-sufficient diagrammatic information.

As with all empirical findings, their importance resides in the theoretical advances and practical applications that flow from a complex of overlapping results. Although the split-attention effect was commonly found between text and diagrams, there were strong theoretical grounds for extending the principle to any two sources of mutually referring information. To this end, Chandler and Sweller (1992) used empirical reports in psychology to show that even where both information sources were text, split attention interfered with learning. As predicted, students reading integrated reports spent less time studying the materials and performed better on subsequent test questions. Although split-source text is less frequently found as explanatory material, these results underscore the need for instruction to accurately direct processing capacity to the appropriate integration of essential elements.

Similar advantages were also found for two-step arithmetic word problems where Year 3 students, after being read a story, studied either integrated or split-source worked examples (Mwangi & Sweller, 1998). Students were asked to generate self-explanations during the study phase, a process hypothesized to facilitate the generation
of problem solving schemas\textsuperscript{141}. Accordingly, it was suggested that “studying integrated worked examples optimizes the inference-making processes” (Mwangi & Sweller, 1998, p. 193). Additionally, the content of self-explanations offered an objective measure of cognitive load: Splitting attention between information sources raised this load and imposed constraints on the inferential processes required for generating explanations. Conversely, the cognitive load for the worked example group was reduced by the processing efficiencies of integrated formats, facilitating the generation of the required explanations.

The impairment to learning associated with the split-attention effect has been widely reported by Mayer and colleagues (Mayer & Anderson, 1991; Mayer & Gallini, 1990; Mayer & Simms, 1994; Moreno & Mayer, 1999a) in relation to their research in multimedia learning. They refer to the advantages of integrating either pictures or animations with explanatory text as the contiguity principle. Consistent with CLT, the underlying explanation of this principle emphasises “that the effectiveness of multimedia instruction increases when words and pictures are presented contiguously in time or space” (Moreno & Mayer, 1999a, p. 358). In accordance with this proposition, the phenomenon of instructional integration is related to either the spatial (Mayer & Gallini, 1990; Mayer, Steinhoff, Bower, & Mars, 1995; Moreno & Mayer, 1999a) or temporal contiguity effects (Mayer & Anderson, 1991, 1992; Mayer, Moreno, Boire, & Vagge, 1999; Mayer & Simms, 1994). Both these instructional effects are explained by Mayer’s (2002) cognitive theory of multimedia learning, which is also closely related to Sweller’s CLT (1988; 1989; 1993) and the limited capacity assumptions of human information processing (Mayer & Moreno, 2003). In line with these assumptions,

\textsuperscript{141} Recall that the same strategy was used to deepen understanding of worked examples (see section 3.3.1).
contiguity effects are attributed to a decrease in the processing load imposed by cognitively efficient (contiguous) materials.

In summary, the split-attention effect is an extensively reported instructional phenomenon, which emerged from the theoretical explanation that accompanied an ineffective application of worked examples. It was suggested that any instructional format is deficient where a learner is forced to engage in mental activities that are not associated with learning (schema acquisition). As a consequence, worked examples are only effective where physically and/or temporally separated elements are fully incorporated into a unitary design. By extension, this instructional imperative applies to any multiple sources of information that must be simultaneously considered in order for understanding to occur.

The strength of instructional designs derived from the investigation of the split-attention effect was in large part due to the support for hypotheses generated by CLT. Both the theoretical predictions and practical applications of the worked-example and split-attention effects were crucial for extending the validity and efficacy of this theoretical base. What is more, mounting evidence for the principles underlying this theory helped generate, with increasing confidence, predictions further removed from intuitive instructional practices of the past. Such a prediction is discussed next.
3.3.3 Redundancy effect.

The interference to learning associated with split attention is evident when a learner must perform additional processing in an effort to mentally unify disparate sources of information. Under these circumstances, searching for (extraneous load) and simultaneously processing (intrinsic load) multiple elements may together impose an excessive cognitive load. As a further consideration, materials might also include elements that replicate information either already present, or already understood. In relation to these issues, studies investigating the split-attention effect raised some important questions: During learning, does mental processing differentiate between essential and unessential instructional elements? More specifically, where unessential or redundant elements are also processed, do they interfere with learning?

Two previously discussed studies (see section 3.3.2) found prima facie evidence that the inclusion of redundant material has implications for instructional design. First, Ward and Sweller (1990) added elaborative statements, which were unessential for understanding, to the core information provided by a self-explanatory diagram. At that time, the inferior performance of this condition was attributed to split attention\footnote{See Experiment 5, Split Attention condition.}. Second, Chandler and Sweller (1991) demonstrated that the performance of learners compelled to integrate both essential and redundant elements (explicit condition) was inferior to their counterparts who were free to exercise some choice over which elements to study (implicit condition). These differences between explicit and implicit conditions arose despite both groups being presented with identical materials. Looking back, both studies provided indirect evidence that redundant material was processed at a cost to understanding.
Chandler and Sweller (1991) compared the split-attention and redundancy hypotheses directly in a series of experiments using either mixed electrical circuit or biology materials. These materials were fully comprehensible when studying a diagram only format, which was predicted to outperform both the integrated diagram-text and conventional split-source formats. As expected, the diagram only condition was significantly quicker processing instructions, an indication of reduced cognitive load; students studying this condition also performed better on subsequent test problems. The results of these experiments were generated and explained by CLT: processing redundant—and therefore additional—material interfered with learning. Notably, redundancy affected learning even when materials were presented in an otherwise cognitively efficient integrated format.

These findings were crucial for a more complete understanding of the characteristics that contribute to effective instructional design: Although the extraneous cognitive load associated with split attention is avoided where multiple sources of information are reformatted into a unitary source (Sweller & Chandler, 1994), the processing of redundant material, irrespective of instructional format, also expends limited cognitive resources. As a result, integrating materials in order to reduce extraneous cognitive load is only beneficial where all information (elements) is essential for understanding. Consequently, integration also requires the careful elimination of redundant elements.

Bobis, Sweller, and Cooper (1993) replicated the redundancy effect using paper-folding activities with Grade 4 students. Of theoretical interest, redundancy was generated by either equivalent diagrammatic and textual information, or two equivalent sources of diagrammatic information. Although diagram and text represent a common instructional format, these results also suggested that any two equivalent information
sources might generate the negative effects of redundancy. Consequently, Bobis et al. concluded, “the usefulness of additional information must outweigh the consequences of having to process it” (p. 14).

Inherent in the utility and validity of any instructional theory is the capacity to generate effective designs based on theoretical predictions, even when these predictions stand contrary to accepted and plausible past practices. Activities such as conventional problem solving, studying adjacent diagrams and text, or incorporating additional but unnecessary information\textsuperscript{143}, have all assumed an important role in traditional instructional materials. Despite this, each of these practices has been shown to be deficient in relation to an understanding of cognitive mechanisms and to maximising the efficiency of mental processing. The explanations for these deficiencies are consistent with CLT, as are the modifications that avoid their instructional deficits.

In the spirit of these counterintuitive findings, Sweller and Chandler (1994) and Chandler and Sweller (1996) hypothesized that initially learning complex software operations from a self-contained manual may be more efficient than using a manual and computer together\textsuperscript{144}. They found that a self-contained manual, with simulated screen captures and integrated explanations, significantly enhanced learning outcomes. In comparison, referring information between a computer and either a traditional manual, or integrated manual, generated split-attention and redundancy effects respectively. Both increased extraneous cognitive load. Moreover, these studies also predicted, and found, that split-attention and redundancy effects would be a problem only when using

\textsuperscript{143} These activities involve the negative effects of problem search (see section 2.6.2), split attention (see section 3.3.2), and redundancy, respectively.

\textsuperscript{144} Both these experiments made use of computer-aided design (CAD) and computer-aided manufacture (CAM) systems.
high element interactivity materials. As expected, differences between groups were either significantly reduced, or disappeared, under low element interactivity conditions.

Although early research into both split attention and redundancy produced consistently strong results, cognitive load, as the theoretical basis of these empirical findings, was usually inferred rather than directly tested. To provide direct support for the effects of cognitive load, Chandler and Sweller (1996), as part of the experiments referred to above, incorporated a concurrent secondary task as an indicator of the load associated with the to-be-learned materials (primary task)\textsuperscript{145}. Secondary task performance was entirely consistent with theoretical predictions: First, low element interactivity materials produced equivalent performance between conditions. Second, recall accuracy was almost halved for the conditions that experienced either redundancy or split attention. Third, the self-contained integrated manual suffered only minor decreases in recall accuracy.

From these results, Chandler and Sweller (1996) suggested that in relation to understanding software applications, “under some circumstances, the removal of computing equipment during critical phases of learning may provide considerable benefit” (p. 168). Although this statement is based on sound theoretical assumptions, it may at first appear controversial as an instructional recommendation. Nevertheless, it does not imply that all software training is best carried out in the absence of computers! By integrating the same high element interactivity materials on screen, Cerpa et al. (1996) were able to reproduce the earlier pattern of results found for the self-contained manual condition. In all these studies, whether generated under real or simulated conditions, flawed instructional designs failed to consider the cost of processing

\textsuperscript{145} See Measuring cognitive load (section 3.2.4) for an explanation of the dual-task paradigm.
redundant elements.

More recently, Mayer and colleagues (Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Mayer, Heiser, & Lonn, 2001; Moreno & Mayer, 2002) have reported the redundancy effect within the framework of Mayer’s (2002) cognitive theory of multimedia learning. They adopted a more restricted definition of redundancy than that applied to CLT, in which they referred to eliminating interesting but irrelevant information as weeding, and reserved the term redundancy for a duplication of essential information (Mayer & Moreno, 2003). Neither definition conflicts with CLT, both subsumed by the notion that any material unnecessary for understanding should be removed; additional processing, regardless of content, increases extraneous cognitive load.

What constitutes material “unnecessary for understanding” is often more difficult to identify than at first appears. By providing a comprehensive approach to a wide range of learners, virtually any textbook will contain, at least for many students, additional and therefore redundant materials. Indeed, following three experiments, Mayer et al. (1996) found, “there was no instructional treatment, including a 600-word passage with summary, that proved to be more effective in promoting retention and transfer than a summary” (p. 72). In effect, these results supported the adage, less is more. From a cognitive perspective, the authors attributed this advantage to the reduced load of processing smaller amounts of text. Moreover, the processing efficiency of a multimedia summary (text and diagram) proved to be more effective than an

146 The next section, Learner experience and expertise (see section 3.3.4), argues that relevance of elements is a product of learner experience.

147 Hence, the title of this study, When Less is More: Meaningful Learning From Visual and Verbal Summaries of Science Textbook Lessons.
extended prose explanation. Any additional or unnecessary processing potentially overloads working memory, regardless of its perceived efficacy as explanatory material.

Mayer, Heiser, and Lonn (2001) tested redundancy in a multimedia environment more directly: Duplicating narration, either exactly or in summary with on screen text, interfered with learning. They explained this finding in terms of the competing visual stimuli of words and animations overloading the limited visual processing capacity of working memory. As an extension of these investigations, Moreno and Mayer (2002) found no redundancy effects for on screen text and narration, providing no other visual stimuli were present, or, the animation preceded the redundant explanation. It was argued that working memory capacity was sufficient in both these instructional variations to avoid cognitive overload. With respect to CLT, it could be argued that neither format generated sufficient element interactivity to exhaust working memory capacity. Moreno and Mayer’s results also confirmed that computer-based instructional materials, which incorporated both diagrams and redundant visual-auditory verbal information, were less efficient than the equivalent diagram and narration only formats.

Sweller (1999) noted that the phenomenon of redundancy pre-dates the split-attention effect, although recognition of this effect has ebbed and flowed over many years. For instance, Reder & Anderson (1980; 1982) found that summaries from college textbooks were superior to studying the actual chapters, and that these advantages persisted over test intervals of up to 12 months. They indicated that spaced practice and

148 The modality effect (see section 3.3.5; using both visual and auditory working memories to increase processing capacity) and sequential processing respectively were used to explain these results.
149 It should also be noted that preceding explanations with animation creates split attention and, for high element interactivity materials, would likely overload working memory.
150 Moreno and Mayer (2002) explained this superiority in relation to the split-attention effect.
the absence of details were possible causes for these advantages. For reasons of cost efficiency and genuine scholarship, many textbooks tend to take a comprehensive and one-size-fits all approach. Too often though they offer little guidance to higher order objectives (Driscoll, Moallem, Dick, & Kirby, 1994), or to the ways in which materials can be tailored to meet individual needs, issues that no doubt extend to the inclusion of redundant material.

Sweller (1999) suggested the main reason for the intermittent awareness of redundancy was the absence of any coherent theoretical explanation for its effects. Likewise, the effect was in part ignored because, intuitively, additional or duplicatory information was seen as either helpful, or at worst, learning neutral. Nonetheless, the seemingly anomalous idea that redundant material interferes with learning is easy to accept from the perspective of CLT and its emphasis on limited working memory capacity: The additional processing generated by attempting to integrate redundant material is assumed to raise extraneous cognitive load and therefore inhibit the integration of essential elements.

From this discussion, it can be readily appreciated that recognizing whether elements are either essential or redundant becomes pivotal to the design of effective instruction. In effect, the answer to this dilemma relies, at least in part, on what we already know, or don’t know; that is, the identification of a student’s level of domain knowledge. This important issue is covered in the following discussion.
3.3.4 Learner experience and expertise.

As discussed in the previous section, including additional material, even if it is not essential, is not an uncommon practice when designing instruction. At first glance, reinforcing some material appears a reasonable strategy when attempting to emphasise and support the learning of certain core elements. Implausible as it may seem, the evidence indicates that the redundancy this strategy generates should dissuade instructional designers from providing these unnecessary materials. Even so, differentiating between redundant and essential elements is not necessarily the perfunctory task it might first appear. In many instructional circumstances, the decision to include, or exclude, material inevitably pivots on an understanding of the individual learner.

Recall that intellectual skill is dependent on the number and nature of schemas held in long-term memory (see section 1.6), which in turn determines our capacity to simultaneously process elements in working memory. Therefore, what constitutes a series of elements for one learner might be little more than a single element–described by a schema–for another learner. On this basis, the same materials may be readily understood by some, and yet totally incomprehensible to others, the number of elements either manageable or excessive respectively. Take for example a medical diagram or schematic that describes an interactive diagnostic procedure (high element interactivity): Experienced doctors very familiar with such information would likely find the diagram self-explanatory and any accompanying information redundant. On the other hand, a medical student (relative novice), less familiar with these materials, might require the additional information in order to process appropriate elements in the correct sequence and with the appropriate meaning.

These issues, from a cognitive load perspective, directly impact on whether
some information should be integrated to prevent split attention, or eliminated to avoid redundancy (Sweller, 1999); in all cases the experience of the learner is the mediating factor. Yeung et al. (1997) examined this nexus between split-attention, redundancy and learner expertise: They used reading passages that included additional information (integrated or split-source) designed to assist text comprehension and understanding of vocabulary. As predicted, low ability English as a second language (ESL) readers benefited from the integrated explanations for comprehension (semantic), but not in relation to vocabulary meanings (lexical): For comprehension (high element interactivity), the split-source format raised extraneous cognitive load and interfered with learning, whereas vocabulary meanings (low element interactivity) were learned best when they were separated from the text, which otherwise acted as a redundant source of information for the integrated condition.

The same experiments failed to find any advantage for the integrated design when administered to high ability ESL readers. These students already understood the meaning of the integrated explanations (semantic), which meant this text was unavoidably processed as redundant material. Under these conditions, split-source materials enabled the more experienced readers to comprehend the text without resorting to redundant information. That is, the redundant material was easier to ignore. Yeung et al. (1997) explained the interaction between format and learner experience as a product of cognitive load: For less experienced students, integrated materials may be indispensable because they provide the elements essential for understanding, whereas the same materials may force experienced learners to process redundant material, unavoidably raising extraneous cognitive load.

The relationship between expertise and instructional materials was comprehensively explored by Kalyuga, Chandler and Sweller (1998; 2001a).
first series of experiments, using a combination of diagram and text, they extended the investigation of learner experience to electrical engineering materials, using integrated, split-source, and diagram only formats. Initially, inexperienced trade apprentices performed at higher levels when learning from integrated materials (split-attention effect). Following two sessions of further training the advantage incrementally shifted to the diagram only condition, and by the end of the experimental period this format was clearly superior to the integrated materials (redundancy effect).

These findings were also readily transferred to a comparison between worked examples (direct instruction) and a less guided exploratory-based environment. Consistent with earlier findings, Kalyuga et al. (2001a) found that for complex learning worked examples were superior for inexperienced trainees, although the difference between formats disappeared following further training. With increased experience the exploratory-based learners improved more than their worked example counterparts. For the same experiments there were no differences between conditions where the materials consisted of simple tasks (low element interactivity).

In relation to complex learning the findings above indicated that inexperienced learners benefited from direct guidance, the schema-deficient search of exploratory learning imposing an additional cognitive load. With further experience, learners became more effective when negotiating an exploratory environment, presumably due to the constructed schemas that permitted a more efficient manipulation of task elements\(^\text{151}\). In a similar investigation, experienced learners performed better when solving conventional problems than they did studying worked examples (Kalyuga et al., 2001b). These learners were instructionally more efficient when exploring a problem

\(^{151}\) See also Tuovinen & Sweller (1999), worked examples (see section 3.3.1).
than when studying the redundant material provided by worked examples. With appropriate schemas, problem solving effectively raises germane cognitive load, a result of adapting problem-specific schemas to new variants within the problem space.

In rather an unorthodox application of CLT, Cooper et al. (2001) and Ginns, Chandler and Sweller (2003), looked to exploit the instructional possibilities of domain experience and its associated schema-rich base for learning. They asked students to imagine the content of materials—including the relations between various elements—as a way of better studying and understanding the content of worked examples. For experienced students, imagining the relations between elements facilitated the automation of appropriate schemas. Not only was this hypothesis confirmed, but as predicted, less experienced students found studying worked examples more helpful than the imagining technique (imagination effect). According to the authors of both studies, without sufficient schema-based knowledge, inexperienced learners found the load of manipulating complex elements in working memory (imagining) too great.

From these experiments and their theoretical perspective it becomes abundantly clear that the consideration of both learner experience and the design of instructional activities are inseparable. Be that as it may, Kalyuga, Ayres, Chandler and Sweller (2003) commented that the knowledge levels of learners are not always given the explicit attention their impact on learning deserves. Furthermore, in a recent review they indicated “that a large number of cognitive load theory (CLT) effects that can be used to recommend instructional designs are, in fact, only applicable to learners with very limited experience” (p. 23). Indeed, findings have consistently demonstrated that with additional training specific CLT effects not only disappear, but also eventually reverse, prompting the application of the overarching term, the expertise reversal effect.

In one series of experiments this effect was observed for trade apprentices
following instruction on the cutting speed nomogram (Kalyuga, Chandler, & Sweller, 2000). The findings from the first experiment supported dual-modal learning (diagram and auditory explanation) over visual only (diagram and visual explanation) and visual-auditory conditions (diagram and auditory-visual explanations), and as expected the diagram only condition was the least instructive for inexperienced learners. After training however, the former dual-modal advantages disappeared. In a second experiment, additional training sessions were undertaken and, in a reversal of the previous findings, learning from a diagram only format was superior to the dual-modal format. Consistent with previous studies experienced learners found the diagrams perfectly sufficient. When presenting their instructional recommendations, Kalyuga et al. (2000) noted that when redundant explanatory text is presented in the auditory mode, it “should be able to be easily turned off or otherwise ignored” (p. 135).

Identifying expertise is crucial, especially when schemas need to be sufficiently developed to enable the simultaneous processing of new elements in working memory. Somewhat paradoxically, Pollock et al. (2002) noted that for inexperienced learners the cognitive burden of processing elements may interfere with the construction of the very schemas necessary to understand those elements. As a result, Pollock et al. suggested that complex materials could be initially learned as isolated elements in order to allow inexperienced learners to construct the schemas required for later holistic understanding. This isolated-interacting elements technique was superior to a conventional format that required students to study all the elements together. What is

\[152\text{ In relation to machine drilling, a “cutting speed nomogram is a quick and reliable method of determining the rpm to run a drill of a given diameter (in mm) at a given cutting speed (in meters per second)” (Kalyuga et al., 2000, p. 129).} \]

\[153\text{ See the modality effect (see section 3.3.5).} \]
more, in subsequent experiments the advantages of this instructional technique for experienced learners failed to materialise, replicating in relation to expertise a similar pattern of results to other CLT effects.

In a similar vein, Mayer and Chandler (2001) also tested a two-stage technique for multimedia materials, both stages incorporating animation and narration. Inexperienced learners (in the field of meteorology) who were permitted to negotiate materials in a part-to-whole technique, self-paced to normal speed, performed better on transfer problems than their corresponding whole-to-part learners. A second experiment, comparing part-to-part with whole-to-whole presentations demonstrated a similar advantage for the self-paced condition. The instructional benefits of at first learning elements in manageable units ensure that when learners are presented with a complex cause-and-effect system they “can devote more cognitive resources to building connections among system components rather than also trying to understand how each component works” (Mayer & Chandler, 2001, p. 396).

These findings are consistent with CLT and the suggestion that learning proceeds best when instruction facilitates the construction of schemas. Furthermore, the most appropriate instruction matches, and therefore optimises, task complexity with learner expertise. In addition, Renkl and Atkinson (2003) assumed that the transition from inexperience to experience corresponds to a decrease in intrinsic cognitive load, and suggested that instructional activities should vary in accordance with a learner’s capacity to explore the problem space in appropriate ways. For example, during the intermediate stages of skill acquisition, worked examples with accompanying self-explanations represent germane cognitive load, however, during the later stages of skill

154 Part-to-whole techniques also allow elements to be learned in virtual isolation.
acquisition, “When skills should be optimized (in terms of speed and accuracy) and automated, problem solving represents germane load because it directly contributes to these learning goals” (p. 19).

The literature provides additional evidence for the effects of learner expertise, although as indicated for redundancy, these effects have not always been explained in relation to CLT. For example, McNamarra, Kintsch, Songer, & Kintsch (1996) used science materials to investigate text coherence, and found high-knowledge readers benefited more from minimally coherent text than readers with little domain knowledge. The authors attributed these differences for high knowledge readers to the compensatory processing required when making the necessary inferences for comprehension. Alternatively, they suggested that minimal text enabled these high knowledge learners to make use of surplus processing capacity in order to infer deeper meaning. Both explanations are consistent with CLT.

When considered together, the findings reviewed for the instructional implications of CLT (see section 3.3) point to the inescapable conclusion that the interaction between learner experience and mental processing determines not only the level of total cognitive load, but also the relative levels of intrinsic, germane, and to some degree, extraneous cognitive loads. Once these factors are considered, a learner’s cognitive capacity can be matched to the effective load (intrinsic and germane) generated by instructional materials. In essence, these factors maximise the cognitive capacity of learners. Under certain conditions, this capacity can also be increased by the method of instructional delivery. These conditions are explored in the following discussion.
3.3.5 Modality effect.

It is impossible to discuss instruction without considering the one or more sensory modalities (e.g., visual, auditory) in which it is received. As a consequence, the medium by which instruction is delivered (e.g., book, computer) has also been viewed as fundamental to the way in which we learn. Over the past twenty years we have become acutely aware of instructional media, the personal computer revolution making multimedia presentation attractive, pervasive and highly accessible. Nonetheless, Sweller (1999) remarks that even though these presentational techniques are “now cost effective compared to several years ago, it does not follow that the use of innovative multimedia is educationally effective” (p. 145). He also makes a point that is too often either overlooked or ignored by educators: Multimedia techniques considerably pre-date the widespread application of computers and yet their impact on education has been minimal!

In reviewing overall media effects, Mayer (1997) similarly concluded that differences between presentation mediums, such as computers (e.g., animation and narration), and textbooks (e.g., illustrations and text) were minimal. He noted that the same instructional effects are often manifest whether learning from either a book or computer, suggesting the emphasis is better placed on how learners construct meaning rather than on media effects per se: As a case in point, recall that split-attention effects were found across a range of instructional media. In relation to multimedia formats, this emphasis on cognition has shifted the focus from what Mayer referred to as delivery media (e.g., computer, book) to a more relevant emphasis on presentation modes (e.g., pictures, words) and sensory modalities (e.g., visual, auditory).

The evidence for favouring mixed mode delivery has been mounting for several years and was strongly supported in a recent and comprehensive meta-analysis carried
out by Ginns (2005). The discussion of modality in Chapter 1 (see section 1.4.4) cited numerous sources that provide support for at least partially separate and independent visual and auditory working memory processors. In addition, it has been well established that for many explanatory materials, combining visual media (either diagrams or illustrations with visual text) is more effective than visual text alone (Iding, 2000; Levin & Mayer, 1993; Mayer, 1989, 1993), and that a combination of either diagrams or illustrations with narration (visual-auditory) is superior to either format alone (Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990)\(^{155}\).

Given the abundant evidence for separate processing channels, Mousavi et al. (1995) reasoned that dual-modal instruction may provide advantages over uni-modal formats by effectively increasing working memory capacity. According to CLT, the limitations of working memory are the central determinant of instructional design, and therefore any increased capacity associated with dual-modal processing should provide an observable advantage to learning. Using geometry worked examples they compared a conventional split-source (visual-visual) format with dual-modal (audio-visual) formats, the text delivered as either auditory or auditory-visual (simultaneous) material. The initial results confirmed the predicted superiority of audio-visual processing\(^{156}\).

Mousavi et al. (1995) believed that these results provided strong support for an expanded working memory hypothesis, although they noted, an alternative explanation could not be excluded: The split-source format compelled successive referral between diagrammatic and textual information, whereas dual-modal materials could be attended to at the same time. Therefore, in two further experiments they presented students with either uni- or dual-modal materials in either simultaneous or successive formats. As

\(^{155}\) Effectiveness was measured by tests such as superior understanding, transfer, and generative solutions.

\(^{156}\) See experiments 1 and 2.
expected, the superiority of the dual-modal condition was just as marked in either format\textsuperscript{157}. On this basis they argued that these results supported an expanded working memory hypothesis rather than the differential effects of searching for, and holding, associated referents.

The \textit{modality effect} was subsequently reported by Tindall-Ford et al. (1997), who gathered further evidence in relation to dual-modal processing expanding working memory\textsuperscript{158}. The first of their experiments incorporated high element interactivity electrical engineering materials; they estimated that understanding would be severely affected by the high intrinsic and extraneous cognitive loads of a split-attention format. They also included dual-modal and integrated visual formats, reasoning that by expanding working memory and lowering extraneous cognitive load respectively, both formats would better facilitate the management of high element interactivity materials. Not only were these predictions confirmed, it was assumed that they supported CLT interpretations of the instructional advantages provided by dual-modal and integrated visual formats.

Although these initial findings were promising, Tindall-Ford et al. (1997) sought direct evidence that the modality effect was attributable to cognitive load. In subsequent experiments they recorded subjective ratings, calculated efficiency measures (see section 3.2.4), and in the third of three experiments, directly compared low and high element interactivity materials. For these materials, subjects were asked to learn both a series of 30 electrical symbols\textsuperscript{159} (low element interactivity) and the explanations of three electrical circuits in which multiple elements had to be assimilated simultaneously.

\textsuperscript{157} There was no significant interaction.
\textsuperscript{158} In a study aptly entitled, \textit{When Two Sensory Modes Are Better Than One}.
\textsuperscript{159} Each described by a simple term that could be learned without reference to the others.
(high element interactivity). They found that learning from dual-modal formats lowered cognitive load, produced a corresponding improvement in test performance, and resulted in overall higher measurements of instructional efficiency. Conversely, there were neither cognitive load nor performance differences between formats for low element interactivity materials. In other words, modality effects were only evident when the materials placed a burden on working memory, providing compelling evidence for the rationale offered by CLT: Using two sensory modes expands processing capacity.

To exhaust alternative explanations for dual-modal advantages, Mousavi et al. (1995) tested the possibility that visual elements, such as reading, were more demanding than listening to the corresponding auditory material. They included a comparison of textual explanations that were either heard (auditory only) or read (visual only), and hypothesized that both conditions would demonstrate equivalent understanding. As expected, reading was found to be no more cognitively demanding than listening, Mousavi et al. concluding that in fact the reverse may be true. Across all six of their experiments, and in conjunction with cognitive load measures and performance differences from a variety of materials, these findings were interpreted as lending considerable support for the increased processing capacity provided by combining visual and auditory processors.

For the modality effect the prediction of effectively increasing working memory capacity stands in contrast to most other CLT techniques, in which advantages are usually attributed to the reduction of extraneous cognitive load. Even so, working memory remains subject to cognitive overload, irrespective of modality, once available capacity is exhausted. Using materials intrinsically high in cognitive load, Kalyuga et al.

---

160 See Experiment 6.
compared dual-modal and visual split-source formats, the sequential animations concurrently provided with either auditory or visual explanations respectively. As a third condition, the same animations were accompanied by simultaneously delivered auditory and written text. According to CLT, the extraneous cognitive load associated with the search for referents (split-source format), and the processing of redundant materials (repeated text format), was expected to overload working memory. Indeed, both split-attention and redundancy effects were found. In contrast, the conventional dual-modal group, without the additional load of extraneous processing, generated superior performance and superior instructional efficiency measures.

In support of these CLT studies, both Mayer and Moreno (1998) and Moreno and Mayer (1999a) confirmed modality effects under similar experimental conditions, comparing split-attention and concurrently delivered dual-modal formats. Furthermore, Moreno and Mayer also compared an integrated visual format (onscreen text and animation), a split-source visual format, and concurrently delivered dual-modal materials, in an attempt to separate contiguity effects (see section 3.3.2) from those associated with modality. They found consistent modality effects across all measures\textsuperscript{161}, commenting “that mixed modality presentations are superior to the most integrated text and visual presentations” (p. 366). In particular, concurrent delivery of audio-visual materials should be preferred to the conventional separation of visual materials (e.g., diagram and text). Although Mayer and colleagues explained modality effects within a dual-coding framework (see section 1.5.2), they also attributed the superiority of these effects to an increase in effective working memory capacity.

As can be readily seen, there now exists considerable evidence for the

\textsuperscript{161} The three measures were, verbal recall, visual-verbal matching, and problem-solving transfer.
instructional advantages of dual-modal delivery. Nevertheless, the advantages reviewed thus far have almost exclusively involved a combination of textual and diagrammatic information; unsurprisingly, these presentation modes are also extremely common across an enormous number of domains. If however the theoretical explanations underlying dual-modality effects are correct, it is reasonable to assume that they should apply irrespective of the nature of auditory and visual inputs. In pursuing this line of reasoning, evidence was gathered using a variety of materials, each providing similar instructional conditions to those already described.

In the first attempt to replicate the modality effect using different formats, Mousavi et al. (1995) combined either sentence-sentence or sentence-diagram formats in either audio-visual or visual-visual modes\(^{162}\). Although text-diagram formats were superior to text-text formats (*multimedia effect*), a modality effect was confirmed, irrespective of presentation mode. Likewise, Tindall-Ford et al. (1997) reported a strong modality effect in conjunction with high element interactivity materials that incorporated text and tables\(^{163}\): The differences (performance and cognitive load measures) between groups were large, and demonstrated that the central mechanism behind the modality effect was not the mode of presentation, but the capacity of working memory to simultaneously process mutually referring sources of information.

In addition to research that has replicated and explained the modality effect, a number of studies have also investigated various ways in which dual-modal materials can be delivered more effectively. In effect, these studies have specifically considered modality effects in combination with other principles such as split attention and contiguity. Bagget (1984), in an early study of contiguity, provided evidence that recall

---

\(^{162}\) See experiment 5.

\(^{163}\) See Experiment 2.
deteriorates, especially after a delay, as the time period between audio and visual segments widens (increments of 7 s, 14 s and 21 s). Under more realistic instructional conditions, Mayer and Anderson (1992) and Mayer and Simms (1994) compared simultaneous and successive presentations of dual-modal materials that explained the operations of complex interactive systems (e.g., automobile braking system). They also found clear advantages attributable to simultaneous presentation (contiguity effect), each of these results providing evidence for the temporal contiguity principle and the negative interference of temporal split-attention (see section 3.3.2).164

Clearly, dual-modal advantages can be further enhanced if extraneous cognitive load is also reduced by allied strategies such as simultaneous presentation. Even so, not all materials are easily modified or guarantee the total exclusion of visual-based search in order to coordinate mutually referring visual and auditory referents. In these cases, despite dual-modal presentation, extensive visual search is likely to exceed working memory capacity unless compatible strategies help circumvent excessive processing loads.

Using materials that incorporated high visual search, Jeung et al. (1997) hypothesised that visual indicators, in the form of electronic flashing, would more efficiently direct working memory resources, and as a result, lower extraneous cognitive load. Over a series of experiments they presented geometry materials that required either minimal or extensive search. Three conditions, comprising a conventional visual-visual split-attention format, a typical dual-modal format, and a dual-modal format with flashing, were compared. Under conditions of extensive search, the flashing format offered the predicted advantage, an advantage attributed to the reduced search provided.

164 It is assumed that the advantages of simultaneity occur for the same underlying reasons that explain the advantages of spatially integrating visual sources (see section 3.3.2).
by the visual indicators. This conclusion was considerably strengthened by the findings for the minimal search materials: both dual-modal conditions performed at an equivalent level and were superior to the visual-visual condition (split-attention). With reduced levels of extraneous cognitive load, both audio-visual formats demonstrated a modality effect (increased processing capacity) for the low-search materials. Given the breadth of this study, Jeung et al. summarised these results as providing “ongoing support for the central role of cognitive load factors in instructional design” (p. 341).

In a similar vein, Craig, Gholson, and Driscoll (2002) also found that the sudden onset of corresponding onscreen animation with associated explanations was as effective as viewing an entire animation: Both these conditions were superior to a static-picture condition. According to the authors, “both the sudden-onset and animation conditions improved performance by directing the learner’s attention to specific elements of the visual display as they were discussed in the narrative” (p. 433). From a CLT perspective, the use of these techniques, as a way to accurately target elements for integrated processing, is assumed to make a more efficient use of available cognitive resources.

A somewhat related, albeit more complex, strategy was investigated by Moreno, Mayer, Spires, and Lester (2001) using an onscreen pedagogical agent\textsuperscript{165} that provided supportive narration in relation to the onscreen task of assembling a plant (visual-auditory). In comparison to the same information presented as onscreen text and diagram (visual-visual), the dual-modal condition was superior for transfer performance tests, but not for retention\textsuperscript{166}. In addition, a modality effect was evident in further

\textsuperscript{165} They used a life-like bug named Herman, which was part of a multimedia program entitled Design-A-Plant.
experiments where either the agent (narration) was present (either fictional character or human face) or not.

Pedagogical agents help direct cognitive processing to essential tasks. Atkinson (2002) provided further support for the advantages of these agents under dual-modal conditions. He incorporated a monologue-style of instruction, a form that was previously deemed inferior to the more personalized style of instruction. In spite of this inclusion, Atkinson reported, “example processing and problem-solving performance can be improved by incorporating the use of a dual-mode of presentation in example-based instruction (Mousavi et al., 1995)” (p. 426). Modality effects clearly extend to an interactive environment: Moreno et al. (2001) believed that multimedia programs “can result in broader learning if the visual materials are combined with auditory explanations of agents, especially when the student is a participant rather than an observer of the learning environment” (p. 210).

The dual-modality effect is of considerable importance to the field of instruction; there would be few authentic learning environments that would not regularly incorporate the delivery of auditory and visual resources. Although the two sensory modes are not mutually inclusive, multimedia techniques more than ever before offer the convenience, adaptability and technical means of presenting dual-modal materials. Whether this method of instruction is widely and effectively practised within real learning environments is far less certain. Reviewing the instructional use of technology, Iding, Crosby and Speitel (2002) commented that teachers “appear to be unaware of educational software that could be helpful in their teaching, and the majority do not use

---

166 These tests would presumably define a relative difference between high and low element interactivity environments.
technology in many teaching-related tasks” (p. 164).

Although there appears to be considerable work needed to ensure the frequent and appropriate use of dual-modal techniques, they are virtually an inescapable consequence of working in some domains: Music demands the extensive use of auditory materials with, as one possible option, written text as verbal explanation. Therefore, dual-modal techniques offer potential formats in which learning might tap more cognitively efficient ways of processing information. Nevertheless, music also presents considerable challenges in relation to established applications of modality: As one example, both music and oral explanatory text are processed through the auditory mode, presenting a potential conflict between the two sources of information. This and other issues are of central interest to the empirical investigations reported in the following chapters.
4.1 Background to Empirical Investigations

Over the past twenty years CLT has provided broad empirical support for the design of effective instruction; this support is documented in detail throughout Chapter 3. Although the findings and instructional formats generated by this theory have proved robust across a variety of instructional domains, there have been few if any attempts to apply this theoretical framework to the materials of music. As is arguable for many domains, music brings a distinctive combination of factors to the area of instructional design, not least of which is the inclusion of both auditory (sound) and visual (notation) representations of music itself. Given these characteristic features, the exploration of music instruction through the application of CLT was seen as a valuable research goal.

From a practical standpoint, there seem few reasons to believe that music is not subject to the same cognitive constraints and design imperatives more broadly relevant to other domains. Although many of the previously reported investigations (Chapter 3) used technical materials, this “has been more for reasons of convenience rather than theoretical imperative” (Sweller, 1999, p. 152). Nonetheless, the artistic and affective milieu in which music is taught can in part overshadow the more practical considerations of instructional design. Moreover, there can be a natural antipathy towards explaining the learning of music in ways that are perceived as overly clinical or emotionally detached. The research reported below, however, recognises that conceptual and artistic understandings of music are mutually inclusive and indivisible goals. It is assumed that these goals will only be strengthened if they are also embedded within a cross-discipline approach, which in turn is based on the same principles that articulate and constrain all human cognition.
Nearly all music instruction inevitably demands the interaction between auditory and visual modes: Either auditory or visual explanatory text, auditory musical excerpts and visual musical notation are the staple mediums by which instruction is transmitted to, or more importantly perceived by, a learner (Figure 12). A comprehensive understanding of music is unlikely without in some way bringing together these vital sources of information.

Bars 1 and 2 give two possible note combinations that equal the value of 6 quaver beats to each bar

\[
\begin{align*}
2 \text{ beats} &+ 1 \text{ beat} + 1 \text{ beat} + 1 \text{ beat} + 1 \text{ beat} &+ 3 \text{ beats} &+ 1 \text{ beat} &+ 2 \text{ beats} \\
&= 6 \text{ beats} & & &= 6 \text{ beats}
\end{align*}
\]

Figure 12. An example of the simultaneously delivered format of visual text, musical notation and auditory music from Experiment 4.

Too often, the configuration of these three information sources is directed more by experience, intuition, or expediency, than by any underlying set of theoretical instructional principles. For the current investigations, CLT was believed to offer a cohesive framework upon which to investigate ecological and effective ways of structuring auditory and visual music materials.
4.2 Cognitive Architecture, CLT and Music Instruction: An Overview

The first three chapters provide a comprehensive review of the theoretical principles that guided the empirical investigations reported below (Chapters 5–9). Central to these principles was an understanding of human cognitive architecture and the ways in which we process information to learn. Based on this understanding, CLT suggests that many instructional designs are ineffective because they ignore universal and fundamental aspects of cognition (Sweller et al., 1998). By extension, it was assumed that the effective design of music instruction was no less reliant on these same cognitive imperatives. With this in mind, the following overview provides the key theoretical elements that both directed and helped interpret the experiments that follow.

Intellectual skills, in the form of schemas, reside in an almost inestimably large long-term memory. Schemas are dynamic constructs that organise information on the basis of how that information is used (Sweller, 1993). Although schemas represent a single cognitive element, the amount of information they contain can greatly vary. On this basis, human expertise is largely determined by the extent to which schemas have been acquired within a particular domain. Typically, expertise requires very large numbers of schemas (Simon & Gilmartin, 1973) acquired over many thousands of hours; in turn, schemas are highly specific to a domain (Ericsson & Charness, 1994).

It follows, that the primary goal of instruction is to facilitate the construction of schemas (Sweller, 1994). To achieve this, new information must be consciously processed through a limited capacity working memory (Miller, 1956). Only a small number of individual elements, possibly as few as two or three, can be manipulated in working memory at any one time (Sweller et al., 1998). In one sense, this self-limiting

---

167 Any expression of expertise includes both the number of schemas and the extent to which each schema represents a series of otherwise individual elements.
feature acts in a way that checks the extent to which information (schemas) is brought to bear when faced with a new problem\textsuperscript{168}, and in so doing focuses processing capacity on matters of perceived relevance (Sweller, 2003). Through repeated and appropriate application, schemas become automated, which helps circumvent working memory constraints (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) by reducing processing loads and by redirecting conscious control to the management of new elements. The ramifications of these mechanisms for learning are crucial to instructional design.

Learning with understanding cannot take place unless multiple elements are integrated within working memory (Sweller, 1994); the mental integration of these elements is a characteristic requirement of high element interactivity materials. If the number of elements exceeds processing capacity, understanding will fail. Because schema development varies between learners, the number of elements represented by instructional materials cannot be viewed as fixed. Consequently, what may appear as several elements for one learner may constitute only a single schema, and therefore element, for another (Yeung et al., 1997). As a result, all materials must be pitched at an appropriate level, permitting the required number of elements to be integrated within a limited working memory. In contrast, rote learning necessitates only a very limited number of elements to be learned at a time, a process that by definition is not subject to the same cognitive constraints when learning with understanding (Sweller, 1993).

According to CLT, the mental integration associated with high element

\textsuperscript{168} When faced with a novel problem, an almost endless and random array of elements could be recalled (long-term memory) if not moderated by an appropriate cognitive mechanism (limited working memory). Sweller (2003) provides a detailed case for how evolutionary processes might have resulted in these mechanisms.
interactivity materials imposes high levels of cognitive load (Sweller, 1993). Cognitive load can be categorised as intrinsic, extraneous, or germane, depending on the nature and structuring of instructional materials. The aim of instructional design is to either reduce or eliminate extraneous (alterable) cognitive load, leaving maximum processing capacity to accommodate the intrinsic (unalterable) load of materials (Paas et al., 2003a). Moreover, where excess capacity remains, instructional designs should direct the processing of elements in meaningful ways that are germane to deeper schema acquisition (Gerjets & Scheiter, 2003). In relation to learning, germane cognitive load leads to more distant transfer of concepts from one context to another (De Croock et al., 1998).

These design principles (CLT) have generated a number of effective instructional formats, and in so doing identified several factors that interfere with the process of learning. For example, in the absence of problem-specific schemas, search-based strategies, such as means-ends analysis (Newell & Simon, 1972), raise extraneous cognitive load and divert valuable cognitive resources away from the central role of acquiring schemas (Chandler & Sweller, 1991). Similarly, the redundancy of requiring learners to process duplicatory or additional material also raises extraneous cognitive load and interferes with understanding (Bobis et al., 1993).

Of significance to the experiments in this thesis is the extraneous cognitive load associated with split attention: the unnecessary separation (time or space) of otherwise mutually referring sources of information (Tarmizi & Sweller, 1988). Although adjacent sources of information may appear aesthetically pleasing and structurally balanced, holding in working memory information from one source, while searching for associated referents from another source, unnecessarily expends limited processing capacity.
The split-attention effect (physical and temporal) is highly relevant to music instruction, where commonly two or three sources of information are physically separated and/or successively delivered. Take for example the material below (Figure 13): In order to understand the selected concepts, a learner must read or listen to the explanatory text and refer these statements to the appropriate parts of the accompanying written notation. It is not unusual to find instruction (e.g., textbooks) that provides large tracts of explanatory text separated from equally large excerpts of music\textsuperscript{169}. In these cases, not only is the quantity of information of concern, it is assumed that comprehending musical notation—interpreting and translating a dynamic language—imposes similar processing loads to those experienced when comprehending equivalent sources of visual and auditory text. Together, these activities place considerable processing demands on working memory.

\textsuperscript{169} Because the length of music necessary to understand some concepts may be quite large, the degree to which information has to be related between sources may be greater than represented in the example below.
Study Example 2: Follow the sequence of letters, a to c below.

a. The upper number of a time signature indicates the total value of beats to each bar.

b. Bars 1 and 2 give two possible note combinations that equal the value of 4 beats to each bar.

\[
\begin{align*}
\text{note 1} &= 2 \text{ beats} \\
\text{note 2} &= 1 \text{ beat} \\
\text{note 3} &= \frac{1}{2} \text{ beat} \\
\text{note 4} &= \frac{1}{2} \text{ beat} \\
\text{note 5} &= 1 \text{ beat} \\
\text{note 6} &= \frac{1}{2} \text{ beat} \\
\text{note 7} &= \frac{1}{2} \text{ beat} \\
\text{note 8} &= 2 \text{ beats}
\end{align*}
\]

Note 1 + Note 2 + Note 3 + Note 4 = 4 beats
Note 5 + Note 6 + Note 7 + Note 8 = 4 beats

Note 1 + Note 2 + \frac{1}{2} \text{ beat} + \frac{1}{2} \text{ beat} = 4 \text{ beats}
\frac{1}{2} \text{ beat} + 1 \text{ beat} + \frac{1}{2} \text{ beat} + 2 \text{ beats} = 4 \text{ beats}

Where the lower number indicates the type of note value by which all other notes will be measured. For example:

Lower Number 4 = quarter note or crotchet
Therefore, in this example, the crotchet is used to measure the length of all other notes.

Figure 13. An example of split-source materials taken from Experiment 2 (stave appears smaller than the original example).

Physically integrating these two adjacent sources of information provides one method of lowering the extraneous cognitive load associated with the split-attention effect\textsuperscript{170}, and in so doing redirects processing capacity to the mental integration required

\textsuperscript{170} If the text were auditory and delivered either before or after viewing the notation (temporal split attention), a complementary strategy for lowering extraneous cognitive load would be the simultaneous delivery of both sources.
for understanding. In the past however, other factors have likely contributed to the infrequent application of this method to the materials of music: First, in the absence of obvious trends or contrary evidence, instructional designers will usually resort to traditional practices\textsuperscript{171}. Second, it is difficult to combine these materials without the outward appearance of an otherwise congested and visually unappealing format. Third, reading or perceiving music is viewed as a continuous process that would be disrupted, rather than enhanced, by the close addition of text. According to CLT, none of these explanations warrants the retention of split-source materials where they can be otherwise integrated within a unified format\textsuperscript{172}.

Also of importance to this thesis are instructional applications of dual-modal materials, which are common formats associated with learning music. It is assumed that auditory and visual sources of information are processed through partially independent working memory channels (Baddeley & Hitch, 1974) and therefore dual-modal presentation increases effective working memory capacity. The dual-modality effect has been investigated using a range of materials. In these cases, two modes proved superior to the same materials presented in a visual only format (Mousavi et al., 1995; Tindall-Ford et al., 1997).

Unlike the previously reviewed dual-modal materials (see section 3.3.5), music does not necessarily offer a strictly visual-visual alternative to dual-modal formats. The very nature of auditory music often compels at least one information source to be delivered auditorially and therefore requires any other source, such as explanatory text,

\textsuperscript{171} A certain authority is no doubt communicated by familiar academic designs.

\textsuperscript{172} These comments do not suggest that listening to larger excerpts of music should always be avoided. However, where listening with understanding is the goal, appropriate schemas need to have been acquired.
to be presented in either the auditory (uni-modal) or visual modes (dual-modal). These choices effectively reverse the typical experimental comparison between formats, from visual-visual and visual-auditory, to a choice between auditory-auditory and auditory-visual. Furthermore, the inclusion of musical notation provides an additional visual source; together, these considerations create new dimensions for the study of dual-modal instruction (Figure14).

3. Smooth shape….

![Smooth shape](image)

angular shape

![Angular shape](image)

**Figure 14.** An example taken from Experiment 5. The audio was delivered either with the text and notation (simultaneous) or with the notation following the text (successive).

It was therefore an aim of the experiments below to investigate the conditions under which three sources of information (explanation, musical notation, auditory music) might be effectively delivered. As outlined above, the instructional formats and materials incorporated in this thesis substantially differed from those previously reviewed in the literature. As a consequence, the resultant investigations of split attention (especially temporal) and dual modality contributed potentially new understandings to the instructional applications of CLT.
4.3 Overview of Experiments

The five experiments reported in the following chapters broadly correspond to three overlapping phases. With respect to CLT, each phase increasingly diverged from the more familiar conditions and formats of past studies. Together, this body of empirical work had two primary objectives: The first objective was to establish ways in which CLT might explain and determine more effective instructional designs in music. To this end, each study incorporated authentic materials and conditions commonly found for music instruction. The second objective was to further investigate CLT principles in a domain where instructional materials were substantially removed from those employed in previous studies. Therefore, the nature of these music materials extends the application and understanding of CLT and further supports its relevance as a generalizable theory across all instructional domains.

For the first phase, Experiments 1 and 2 established the conventional application of split-attention and dual-modality principles to the materials of music. Although the use of musical notation (in place of the familiar diagrammatic information) introduced a new instructional medium, the structure of presentation formats was otherwise in keeping with previous studies. Under these conditions, it was expected that CLT predictions would equally apply to the domain of music. Hence, experimental conditions distinguished between physically integrated, conventional split-attention and dual-modal formats. All formats required the integration of two otherwise unintelligible sources of information.

The second phase further explored the dual-modality effect by introducing the instructional medium of auditory music, as a result reversing the modality of explanatory text (auditory to visual): Rather than the more conventional delivery of visual diagram and auditory text, the dual-modal formats for Experiment 3 combined
visual explanatory text with auditory music. Furthermore, the uni-modal conditions were auditory rather than visual, consisting of both musical excerpts and explanatory text. This was a marked deviation from previous CLT formats, including those incorporated in Experiment 2. Naturally, the application of CLT to these formats is of considerable import to instructional research in music; few concepts within this discipline can be validly learned without substantial reference to auditory music.

For the third and final phase, attention was turned to combining three sources of instructional information into a unified format. These formats included visual musical notation with the dual-modal arrangement of visual text and auditory music outlined above. Across three experiments\textsuperscript{173}, either auditory music (Experiment 4) or visual musical notation (Experiments 5 and 6) provided the third and additional information source, and these formats were compared with either visual-visual or visual-auditory two-source designs respectively.

These instructional configurations represented not only ecologically valid ways in which music is taught, the merging of explanatory text, auditory music and musical notation was also seen as an effective strategy for delivering complex materials. For this reason, musical notation was included as a means of supporting the referral of information between explanatory text and auditory music; holding auditory excerpts in working memory while finding their appropriate referents was assumed to substantially raise cognitive load. On the other hand, providing all three sources, either simultaneously or successively, potentially overloaded working memory and raised intrinsic and extraneous cognitive loads respectively. This dilemma is not dissimilar to some of the issues faced by Pollack et al. (2002) when investigating their isolated

\textsuperscript{173} The final experiment was a double experiment, denoted as 5 and 6.
interacting elements technique\textsuperscript{174}.

In summary, these phases formed a cohesive series of studies that examined, within a CLT framework, two major principles underpinning instructional design: First, dual modality, a central issue of Experiments 2 – 6, incorporated formats consisting of either two or three information sources. Second, all experiments compared conditions in which materials were either temporally and/or physically integrated or delivered in split-source formats. In all cases, the formats designed for these experiments were generated by CLT, which in turn generated the findings reported below.

\textsuperscript{174} Although the replication of music in visual (notation) and auditory forms may appear redundant, for younger and inexperienced learners (lower levels of aural acuity), the abstract nature of auditory music may in itself prove very challenging when attempting to mentally integrate explanatory text with aspects of musical examples.
Chapter 5 Experiment 1

5.1 Introduction

Experiment 1 was designed to investigate whether two mutually referring information sources in music were better understood when either integrated or discretely placed. Sweller et al. (1990) and Tarmizi and Sweller (1988) established the superiority of spatially integrated materials over the same information split between two sources; a number of studies since have replicated these findings across a range of disciplines (Bobis et al., 1993; Chandler & Sweller, 1991, 1996; Sweller & Chandler, 1994; Yeung, 1999; Yeung et al., 1997). Textual information has been variably combined with diagrams, tables and other text, the results consistently reinforcing the importance of facilitating the mental integration of related instructional elements. Nevertheless, the CLT principles that generated these findings have not, to my knowledge, been applied to the materials of music.

The purpose of this experiment was to test whether the separation of explanatory text from the musical notation to which it related would interfere with understanding in comparison to the same information presented as a unitary source. According to CLT, processing capacity is quickly exhausted when working memory is required to hold information from one source while simultaneously searching for referents from another source. This phenomenon is referred to as the split-attention effect (Sweller, 1999) and is avoided if the two sources of information are placed together as part of an integrated design.

Two conditions were used in this experiment. A conventional split-attention format placed musical notation above a written explanation; to understand the instructional sequence subjects had to alternate their attention between the two sources.
of information. In contrast, an integrated format placed each explanatory statement directly adjacent to its associated item of musical notation, forming a singular source of instruction. Both conditions required that students attend to the statements and notation; neither could be ignored if the materials were to be fully understood.

Experiment 1 hypothesized that integrating written explanations with musical notation (integrated condition) would be superior to the same materials placed apart (split-attention condition). The design of this experiment was important because even though explanatory text is common to most instructional situations, the written language of music (notation) is in some respects appreciably removed from the nature of previously tested materials (e.g., diagrams). By incorporating music’s idiosyncratic language within a CLT framework the corroboration of previous findings gain greater import, strengthening both the theoretical and practical implications for these instructional designs.
5.2 Method

5.2.1 Participants.

Fifty Year 7 students (approximately 11-12 years of age) from an academically selective boys’ high school in Sydney took part in this experiment. Note that for the experiments reported in this thesis, all students (Year 7) were drawn from the same high school; however individual participants usually varied from one experiment to the next as a result of the empirical work being completed over a number of years. All students were in some way either actively or recently involved in music, independent of school curriculum requirements. These students were categorised as musically experienced in relation to their peers, which was a necessary step to ensure current levels of understanding were sufficient to complete the problems once additional study was undertaken. An emphasis was placed on selecting students who played a musical instrument; in this way participants were not only conversant with the language of music, but also had, to some degree, automated schemas for reading music.

As part of the primary school curriculum in NSW all students are provided with a formal program of music education experiences. However, these musical experiences can be differentially augmented for some by extra curricular activities and/or by private instruction. In order to ensure students’ base level understanding was sufficient to commence instruction in this experiment, only students with extended instrumental experience (gained through either extra-curricular or external school tuition) were chosen. In addition, a survey was developed to determine students’ instrumental experiences, formal accreditation, broader level of domain knowledge, and skill development in music (Appendix H.). Screening of recent performance indicators (e.g., class tests common to all participants) was also carried out to confirm that students’
current levels of understanding were sufficient to complete the problems once additional study was undertaken.

All students in this study were academically able, as measured by the criteria for selective school placement in NSW. If placement is sought, each student is required to sit an entrance examination that measures aptitudes in mathematics, English language and general ability. Students are ranked according to academic merit from a combination of their composite performance on these tests, moderated primary school English and mathematics results, and comments by Primary school principals, and then progressively offered places in one of the 30 selective high schools across the State.

Due to these factors of selection, academic abilities of the students in this experiment describe a relatively homogeneous population, however the specific abilities subject to school selectivity are not variables under consideration in this study. It should also be noted that ability in music is not a criterion for selective school entry and, irrespective of other considerations, the disparate nature of Year 7 students’ musical experiences warranted the steps outlined above in order to identify participants for this study. Given the nature of the student population and the process of participant selection, the sample available for Experiment 1 enabled a fair distribution of knowledge and skills between experimental conditions.

5.2.2 Materials and procedure.

All students were given three distinct phases of experimental materials (each immediately followed by the next), which in sequence consisted of five instructional examples, six acquisition problems, three sets of similar and three sets of transfer test problems (thirty problems in each set). For the instruction and acquisition phases, two sets of materials differed by experimental treatment (integrated and split-attention) and
introduced the basic principles of understanding time signatures and their manipulation. The test problems were identical for all participants and both groups of students completed the experiment together in one session (all experiments were administered in dedicated sessions rather than normal music classes). All examples and problems were developed by the experimenter and presented on A4 sheets of paper. The materials formed a logical extension of work undertaken as part of the regular music program for Year 7 students at the chosen school.

The instruction phase consisted of five examples that explained time signatures and the values of notation in relation to a given unit of beat. Each instructional sequence was placed on a separate page. The first example gave a series of basic note values and their beat equivalents in relation to the crotchet. Students had already rote learned and were familiar with crotchet beat values, but did not understand the associated indication of a time signature’s lower number. The second example provided two bars of rhythm to demonstrate the relationship of variable note values to the given time signature. The third example repeated the format of the first, although note values were studied in relation to a quaver beat; the fourth and fifth examples repeated the format of the second, incorporating however time signatures of \( \frac{3}{4} \) and \( \frac{5}{4} \) respectively. Nineteen minutes were allowed for the explanatory phase (5 min each for pages 1 and 3, and 3 min each for pages 2, 4 and 5). The timing for each phase was established in a pilot study consisting of similar instructional sequences and test problems. Students were asked to turn their page when the allocated time expired, and any reference to earlier or later pages was not permitted.

A complete explanation of both lower and upper time signature numbers and the transfer of both bar and beat values to different compound metres built upon prior Year
knowledge and the experience students had gained as practising musicians. For students in their first year of secondary school the extension of time signature principles was high in element interactivity. Once a unit of beat was established (lower number), students had to relate each note value within a bar of rhythm to that unit. Concurrently, they were obliged to be aware of the metrical division (upper number) and the theoretical and/or practical ramifications it had for the rhythmic organization of music. In summary, students simultaneously managed three variables: a unit of beat, a metrical division and, within these constraints, a very large possible number of note value configurations. All these elements had to be held and manipulated in working memory if their relations were to be appropriately understood.

The acquisition phase followed the instruction phase. It consisted of six acquisition problems, which were divided into three problem pairs, alternating between a worked example and conventional problem. The worked examples provided a sequential solution process that for each type of problem appropriately manipulated the previously learned instructional elements. Each of the previously outlined rhythmic elements (note values, metre, unit of beat) was in turn rendered incomplete and could only be resolved when differences between known and unknown problem states were successfully considered.

The first pair of problems presented a 9\text{\textfrac{4}{4}} time signature and an incomplete bar of rhythm. A single note value was required to complete the bar according to the given time signature. The second pair of problems supplied only the lower number of a quaver-based time signature (8\text{\textfrac{4}{8}}) and a complete bar of rhythm. The upper number of the time signature was required after considering the given unit of beat and combined note values. The third pair of problems provided 4\text{\textfrac{4}{4}} and 8\text{\textfrac{4}{8}} time signatures respectively.
and two complete bars of rhythm without the bar line in between. A bar line had to be inserted once the appropriate value of notes and given time signature were considered.

It was anticipated that students would find the understanding required for these problems quite difficult and consequently acquisition materials, including worked examples, were included in this first experiment as a means of fully explaining the appropriate use of problem solving operators. It was conceivable that the instruction and acquisition phases could have been merged, retaining some use of worked examples and abandoning conventional problems. In the final analysis, it was believed that for the chosen students the incorporation of worked examples helped to promote the necessary schemas for a comprehensive understanding of the materials and the subsequent application of knowledge across a range of related test problems.

Students were allocated 2 min 30 s to study and complete each of the worked examples and conventional problems respectively (15 min in total). Each problem was presented on a separate page, which was turned on instruction once the allocated time had elapsed. Access to the instruction phase pages was available throughout this phase, but any reference to earlier or later acquisition phase pages was not permitted. All conventional problems were of two bars duration and for the first problem required the insertion of two notes (one each bar) to complete the answer. The second and third conventional problems required only one intervention (time signature and bar line respectively) for a correct solution. One mark for each correct note or intervention was allocated, giving a possible total of 4 marks for the three acquisition problems. All questions were objectively marked and no marks were allocated for an incorrect answer. The number of correct solutions was recorded.

Following the final acquisition problem, students were issued with a subjective rating scale and asked to reflect on the difficulty of understanding (mental load) the
previous learning materials. They were required to tick a box corresponding to a 9-point scale that ranged from *Extremely easy to understand* to *Extremely hard to understand* (Appendix I.).

Cognitive load ratings were collected to investigate, both in isolation and in conjunction with test performance (efficiency measures), the mental resources required across different experimental conditions in order to achieve the desired learning outcomes. These ratings were collected following either acquisition problems or instruction phases throughout all experiments in this thesis rather than, as indicated by Paas and van Merriënboer’s (1993) original conception, following the test period. Sweller and colleagues (see section 3.2.4) have applied as standard practice this variation of the relative condition efficiency formula on the basis that it is of primary interest to establish the cognitive load of learning rather than of testing, the latter providing only an indirect measure of the mental effort devoted to any instructional episode. In comparison with more direct and laboratory-based methods (e.g., heart rate, neurological scans), subjective ratings are considered to be an effective, straightforward and non-invasive way of indicating working memory load (Paas et al., 1994). When combined with test performance an accurate estimate of cognitive load is achieved and confounding factors such as effort, fatigue and motivation are minimised.

The split-attention and integrated groups had 24 and 26 participants respectively. Students were randomly assigned to one of two instructional groups, with approximately equal numbers of students drawn from each of five music classes. Both conditions received identical explanation and manuscript examples by way of differing formats for both the instruction and acquisition phases. The split-attention condition was given the musical notation and associated statements in a split presentation format (Figure 15). The statements were provided below the notational examples to which they
referred. The notation and sequential solution statements formed two adjacent but
 discrete components on the page.

**Split-Attention Format**

Solution:

a. Determine the total value of beats to the bar, \( = 6 \)

b. Determine the value of the beat = quaver.

c. Allocate the quaver (beat) value of each note and total.

\[
\begin{align*}
\text{note 1} & \quad + \quad \text{note 2} & \quad + \quad \text{note 3} \\
2 & \quad + \quad 1 & \quad + \quad 2 & \quad = \quad 5
\end{align*}
\]

\[d. \quad \text{Calculate the difference between the time signature value (6) and the bar total (5).}\]

\[6 - 5 = 1 \quad \text{in} \]

*Figure 15.* An example of the split-attention format used in Experiment 1.

The integrated condition received identical materials, however, all components were physically integrated into a unitary source of information (Figure 16). The students studying the integrated format had the opportunity to process instructional elements simultaneously, an advantage not available to the students studying the split-attention format. By design, students studying split-source materials were required to hold explanatory statements in working memory while mentally relating them to their physically displaced notational referents.
Integrated Format

Solution:

- **a.** Determine the total value of beats to the bar, $= 6$

- **b.** Determine the value of the beat = quaver.

- **c.** Allocate the quaver (beat) value of each note and total.

- **d.** Calculate the difference between the time signature value (6) and the bar total (5)

$$6 - 5 = 1$$

Figure 16. An example of the integrated format used in Experiment 1.

A test phase followed the acquisition phase. Written test materials were identical for both conditions and they were divided into two major sections consisting of three sets of similar problems (A, B and C) and three sets of transfer problems (D, E and F). Each set contained thirty problems; sets A, B and C corresponding to the three types of acquisition problems and sets D, E, and F were three equivalent sets of transfer problems. Three minutes were allocated to each section (18 min in total) and once this time limit had expired the next section was presented without explanation.

Similar problems made use of only quaver-based time signatures ($$8$$) and transfer problems included a range of time signatures that varied between minim and semiquaver beat values ($$2 \cdot 16$$). The time signatures associated with transfer problems were at no time incorporated into either the instruction or acquisition phases; accordingly, a deeper understanding of the explained principles underlying time signatures and rhythmic manipulation was tested by these problems. Students were not allowed access to either the instruction or learning phase pages while completing test problems.
Students were asked to attempt as many of the thirty problems as possible (six each page) for each of the similar and transfer sections. The number of problems was beyond most students’ capacity to complete in the allocated time and therefore any potential differences between conditions reflected not only levels of mastery, but also the rate at which problems were solved. All problems were of two bars duration and therefore sets A and D required the insertion of two notes (one each bar) to complete the answer. One mark for each correct note was allocated to these sections giving a possible total of 60 marks for each set (30x2). The remaining test sections required only one intervention (either time signature completion or bar line) and one mark was allocated for each question, giving a possible total of 30 marks for each set. Together, each series of similar and transfer problem sets were marked out of a possible 120 marks. All questions were objectively marked and no marks were allocated for an incorrect answer. The number of correct solutions was recorded.
5.3 Results

The variables under analysis were number of correct solutions during acquisition, and the number of correct similar and transfer test problems completed. Means and standard deviations are displayed in Table 5.1.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Condition</th>
<th>Max Score</th>
<th>Split-Attention M</th>
<th>SD</th>
<th>Integrated M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition Problems</td>
<td>Split-attention</td>
<td>4</td>
<td>3.54</td>
<td>0.78</td>
<td>3.96</td>
<td>0.20</td>
</tr>
<tr>
<td>Similar Test Problems</td>
<td>Split-attention</td>
<td>120</td>
<td>58.04</td>
<td>21.07</td>
<td>58.88</td>
<td>12.08</td>
</tr>
<tr>
<td>Transfer Test Problems</td>
<td>Split-attention</td>
<td>120</td>
<td>33.83</td>
<td>17.48</td>
<td>35.85</td>
<td>19.09</td>
</tr>
</tbody>
</table>

Note. n = split-attention 24; integrated 26

Independent \(t\)-tests indicated a significant difference between conditions for acquisition problems that favoured the integrated format, \(t(48) = 2.66, d = 0.74\). An examination of means indicated that from a possible perfect score of 4, most students readily achieved the understanding required to successfully manipulate the problem-solving operators. Both conditions enjoyed high levels of achievement, however, the very narrow dispersal of marks within the integrated condition indicated there was almost uniform understanding. No significant effects were found for either similar or transfer problems, \(t(48) = 0.18\) and \(t(48) = 0.39\) respectively. The observed variation within the integrated condition was once again narrower for similar problems.

\(^{175}\) The level of significance throughout this thesis is \(p < .05\) unless otherwise stated; the effect size is calculated as \(d = (M_1 - M_2)/\sigma_{\text{pooled}}\), Cohen (1977).
Otherwise, these data indicated equivalent performance between conditions on all test problems.

Means and standard deviations for cognitive load and efficiency measures are displayed in Table 5.2.

<table>
<thead>
<tr>
<th>Table 5.2 Cognitive Load and Efficiency Measures for Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
</tr>
<tr>
<td>Phase</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cognitive Load</td>
</tr>
<tr>
<td>Efficiency Measures</td>
</tr>
</tbody>
</table>

Note. n = split-attention 23; integrated 26

An analysis of cognitive load indicated there was no significant effect, \( t(47) = 1.91 \). Mean levels of cognitive load were very low, and on average fell between Extremely easy to understand and Very easy to understand, suggesting that by the end of the acquisition phase neither condition found the learning taxing.

The explication of cognitive load as a multidimensional construct has been a central feature of more recent research (Brünken et al., 2003; Paas, Tuovinen, Tabbers, & van Gerven, 2003b) and therefore efficiency measures were also calculated to further investigate factors of mental load. The variables under analysis for instructional efficiency were subjective ratings and performance measures. An analysis of efficiency measures indicated there was no significant effect, \( t(47) = 0.35 \). Means and

---

176 The formula for instructional efficiency is \( E = \frac{(z_{p} - z_{m})}{2^{0.5}} \). See section 3.2.4 for a comprehensive discussion of cognitive load and efficiency measures.
standard deviations are displayed above in Table 5.2. The lower mean level of cognitive load recorded for the split-attention condition counterbalanced the higher mean levels of performance scores for the integrated condition, resulting in little overall difference for instructional efficiency. The results of determining relative cognitive efficiency are clearly seen when standardised performance and cognitive load measures are plotted on a Cartesian plane (Figure 17). For Experiment 1, both conditions lie near the diagonal line, representing neutral efficiency, and in different quadrants, representing differences of individual measures.

**Experiment 1**

![Efficiency Measures](image)

*Figure 17.* Relative condition efficiency as conceived by Paas and van Merriënboer (1993).
5.4 Discussion

In accordance with a cognitive load hypothesis the acquisition results demonstrated a learning advantage attributable to the integrated format. This advantage only held for the learning phase and did not extend to similar and transfer problems, where both conditions performed at an equivalent level. The interacting levels of acquisition performance and cognitive load strongly suggest that all students found the learning quite straightforward and that by the test phase understanding was already secure.

It can be hypothesized that some students from the split-attention condition initially found the necessity of mentally integrating related but displaced sources of information harder to process. By the third pair of acquisition problems (worked example and conventional problems) there was a smaller difference between group means, suggesting that both conditions had by this point achieved similar levels of understanding. From this converging pattern of acquisition means it is not surprising that both conditions performed equally throughout the test problem phase. It should also be noted that consistent with the direction of acquisition results, performance scores for the similar test problems remained less variable for the integrated condition, and for the transfer test problems these students recorded a non-significantly higher mean.

The very low mean subjective ratings suggest that lower and declining levels of mental effort were all that was required to understand the worked examples and complete the conventional problems. Ceiling effects in Experiment 1 most likely resulted in subjective ratings that were less sensitive to measuring differences in mental effort. The collective data suggest, for all participants, that materials in Experiment 1 were relatively easy to learn, although for the integrated condition they were relatively
quicker to learn.

The amount of time allocated to learning tasks is fundamental to any discussion of instructional methods. Regardless of environment, understanding is always qualified by both performance goals and the time taken to achieve them. In the time allowed for Experiment 1, element interactivity was initially high enough in the acquisition phase to afford a learning advantage to the integrated condition. Plainly, the longer the time allocated the more likely a group of participants will learn the given material. The 19 min and 15 min taken for the instruction and acquisition phases respectively may have been overly generous and as a result facilitated the necessary learning irrespective of instructional format. It is assumed that students studying the split-attention condition did not make the necessary referential connections with the same speed as the integrated condition; however, given a surfeit of acquisition time they attained equivalent performance outcomes by the end of that phase.

Repetition, like time, is an instructional principle relevant to nearly every learning environment. All other factors being equal, repeated instruction eventually leads to the desired performance. In the current experiment, the similar nature and recurrent structure of the paired acquisition problems reinforced rather than expanded the targeted learning principles. The total effect of this learning sequence was to dilute any early learning advantage that may have been conferred by integrated instruction. Taken together, the compounding effects of time, repetition and low intrinsic cognitive load undermined the impact of the learning materials and therefore limited the extent to which differences in performance were manifest throughout the experiment. It was predicted that by modifying these variables in Experiment 2, higher levels of cognitive load would explain differential performance in both acquisition and test problems.

In summary, Experiment 1 provided some evidence for a split-attention effect
when musical notation was separated from its related explanatory statements. However, as outlined above, the length of time allowed for instruction and the ease with which students were able to learn the materials most likely resulted in ceiling effects.

Experiment 2 attempted to rectify these problems.
Chapter 6 Experiment 2

6.1 Introduction

Experiment 2 was designed, in part, to provide further evidence for the hypothesis tested in the previous study. The results of Experiment 1 partially supported the superiority of integrated over split-attention formats, but failed to extend these findings to the test phase. Given indications across the data that students found the instruction too easy, Experiment 2 modified these materials in order to increase overall learning difficulty. Without learning difficulty cognitive load effects are unlikely to be found. In line with earlier recommendations instructional time was considerably reduced and in order to avoid repetition, some of the previously included explanatory statements were deleted. Furthermore, the number of acquisition problems (including worked examples) and allowable working time were also reduced. Specifically, only one of the three problem types employed in the previous experiment was included in the current materials. As a consequence, the corresponding test problems consisted of only one set each of similar and transfer problems (replacing the three different problem sets previously included).

In addition to the split-attention hypothesis, Experiment 2 aimed to investigate whether delivering explanatory statements and notation in two modes (auditory-visual) was superior to presenting statements and notation in a conventional split-attention format (visual-visual mode). Working memory is comprised of at least partially independent processors for visual and auditory inputs, which under certain conditions confer instructional advantages (Mayer & Moreno, 1998; Moreno & Mayer, 1999a). Moreover, Sweller and colleagues (Jeung et al., 1997; Kalyuga et al., 1999; Mousavi et al., 1995; Tindall-Ford et al., 1997) have gathered considerable empirical evidence for
their theoretical position that dual-modal instructional formats effectively expand working memory capacity. They identified this phenomenon as the modality effect, replicating its instructional benefits over multiple experiments that found dual-modal presentation ameliorated the adverse effects of a split-attention design.

To test both the split-attention and dual-modality hypotheses, three conditions were used in this experiment. The split-attention and integrated conditions were as outlined in Experiment 1. In addition, a third condition was introduced that received the musical notation in an identical visual format to the integrated and split-attention conditions, however the explanatory statements were delivered as auditory text. To ensure intelligibility of the materials, all conditions required that students attend to both sources of information (as for Experiment 1).

Experiment 2 hypothesized that either integrating mutually related sources of information or delivering the same information between visual and auditory modes would be superior to the same materials delivered in a split-source format. Given the relevance of auditory input and dual-modal delivery to music it is important to emphasise that the design of this experiment, as with the first, incorporated materials that maintained ecological validity. Combining either visual or auditory explanations with the written language of music (notation) sought to extend our knowledge of the ways in which these essential materials are best presented in a typical classroom setting.
6.2 Method

6.2.1 Participants.

Seventy-nine Year 7 students from an academically selective boys’ high school in Sydney took part in this experiment. Students were selected as musically experienced in relation to their peers, which was a necessary step to ensure current levels of understanding were sufficient to complete the problems once additional study was undertaken. All other participant characteristics and the procedures for selection were as outlined in Experiment 1.

6.2.2 Materials and procedure.

Experiment 2 materials were closely related to those described for the previous experiment. Once again they were presented to all students in three sequenced phases, however for this experiment these phases consisted of four instructional examples, four acquisition problems, one set of fifty similar test problems and one set of fifty transfer test problems (each immediately followed by the next). The first two phases differentiated between three experimental treatments (integrated, split-attention and dual modality) that introduced the basic principles of time signatures and their manipulation. Test problems were identical for all participants. The experiment was administered in two sessions; the written conditions and the dual-modality condition were completed as separate groups. All examples and problems were developed by the experimenter and presented on A4 sheets of paper. As with the previous study, the materials were a logical extension of the regular Year 7 music program at the chosen school.

The instruction phase consisted of four examples (one each page) that explained time signatures and the values of notation in relation to a given unit of beat. Each instructional sequence was placed on a separate page. The first example outlined the
relations between note values and a crotchet beat, and the second example explained, in relation to a \( \frac{4}{4} \) time signature, the functions of the upper and lower numbers (Appendix J). Students had previously rote learned crotchet equivalents of basic note values and therefore the first example was presented as a preparatory step before further instruction. For this reason, all three conditions received this first example in a written format as integrated notation and explanatory statements. The second example, which explained the relationship of note values to a time signature, and all examples thereafter differed by experimental treatment. The third and fourth examples highlighted the relationship of note values to quaver and minim beats respectively, the sequence of materials similarly structured to the first example.

In comparison to the previous experiment a number of steps were taken to increase learning difficulty. The instruction phase extended the explanation of note relations to minim beats (Example 4), increasing possible beat variants to three (crotchet, quaver and minim). Unlike the materials of Experiment 1, the sequential explanation of time signature numerals and their relationship to note values was included only once following the preparatory example of crotchet beat values (Example 1). This was a significant reduction over the three similar examples previously provided (one for the crotchet beat and two different examples for a quaver beat). As a result, students were expected to draw greater inferences from the more familiar crotchet example when relating the principles of time signatures to quaver- and minim-based beat values. That is, the transfer of time signature principles from the more familiar \( \frac{4}{4} \) to the unfamiliar \( \frac{8}{4} \) and \( \frac{2}{4} \) time signatures was far less explicit in Experiment 2. As a consequence, transfer distance increased when students were required to make the necessary mental connections from one instructional example to another.
Reducing the time allocated to study the materials further increased instructional difficulty. Students were given 5 min 30 s for the instruction phase (1 min for the first example and 1 min 30 s each for the second, third and fourth examples). When all modifications to the materials were viewed together, it was apparent that for this experiment students still had to simultaneously hold and manipulate in working memory a number of elements, however, they were given less time and fewer examples upon which to establish and support schema acquisition. The relations between elements had to be understood from the crotchet example (second), and then appropriately applied to new contexts (beat values, Examples 3 and 4).

The acquisition phase followed the instruction phase. It consisted of four acquisition problems, which were divided into two problem pairs, alternating between worked examples and conventional problems. The worked examples manipulated previously learned instructional elements and provided a sequential solution process. A significant deviation from the previous materials was the retention of only the first of three problem types (completing a bar of rhythm) for both worked examples and conventional problems. As a result, excessive reinforcement of the same instructional principles and the use of lower element interactivity problem types were avoided. Completing a bar of rhythm according to a given time signature required a greater number of problem moves in order to reach solution. Additionally, the time signature was varied from the worked example to the equivalent conventional problem, requiring a less formulated or imitative response within the problem space.

Students were allocated 1 min 30 s to study each of the worked examples and 1 min to complete each conventional problem (5 min in total). Each problem was presented on a separate page, which was turned on instruction once the allocated time had elapsed. Access to the previous instruction phase pages was available throughout
the acquisition phase, however students were not permitted to review earlier or later acquisition pages during this phase. The reduction of problem-solving time in comparison to Experiment 1 effectively tested students’ capacity to apply a schema-based solution process rather than falling back on search-based strategies. In view of the changes outlined above, it was predicted that any format that reduced extraneous cognitive load and dedicated greater working memory resources to schema development (germane cognitive load) would prove more effective.

The conventional problems were of two bars duration and required the insertion of two notes (one each bar) to complete the answer. One mark for each correct note was allocated, giving a possible total of 4 marks for the two acquisition problems. All questions were objectively marked and no marks were allocated for an incorrect answer. The number of correct solutions was recorded.

Following the completion of the acquisition problems students completed a 9-point subjective cognitive load rating scale, which was used as a basis to measure cognitive load and instructional efficiency during the acquisition phase (as per the explanation for Experiment 1).

The two written conditions had 26 participants and the dual-modality condition 27 participants. Students were randomly assigned to one of three conditions, with approximately equal numbers of students drawn from each of five music classes. Three groups received identical explanation and manuscript examples by way of differing instructional formats for both the instruction and acquisition phases. The split-attention and integrated groups were formatted as previously described (Figures 15 and 16). The visual-auditory condition was supplied with identical notational examples to those in the visual conditions, except that students heard identical text in an auditory form rather than reading written explanations. All auditory material (explanatory text) was delivered
on compact disc. For all groups alphabetical markers identified the specific areas to which explanations referred, and specific musical notes were identified in accordance with their corresponding statements (Figure 18).

![Figure 18. An example of the visual-auditory dual-modality format used in Experiment 2.](image)

The split-attention condition compelled students to hold elements in working memory while searching for the necessary referents elsewhere on the page. It was assumed that this format would raise extraneous cognitive load, consume additional mental resources, and therefore reduce available capacity for schema acquisition. Conversely, by placing mutually referring elements together on the page, it was believed that the integrated condition would avoid the extraneous cognitive load of mentally integrating adjacent elements. The benefits of the dual-modality condition were expected to arise from both an expanded working memory and the simultaneous processing (temporal synchronicity) of elements.

A test phase followed the acquisition phase. The written test materials were identical for all three groups and were divided into two sections of fifty similar and fifty transfer problems. Similar problems used a variety of quaver- and minim-based time signatures ($\frac{8}{2}$) as per the instructional examples, and transfer problems consisted entirely of semi-quaver based time signatures ($\frac{16}{6}$). Similar and transfer test problems were allocated 3 min each (6 min in total) and followed without further explanation. Students were not allowed access to either the instruction or learning materials during
the test phase.

Students were asked to attempt as many of the fifty problems as possible (six each page) for each of the similar and transfer test sections. The number of problems was beyond students’ capacity to complete in the allocated time. Therefore, any potential differences between conditions reflected not only levels of mastery, but also the rate at which problems were solved\textsuperscript{177}. All problems were of two bars duration and required the insertion of two notes (one each bar) to complete the answer. One mark for each correct note was allocated to these sections, giving a possible total of 100 marks for each set (50x2). All questions were objectively marked and no marks were allocated for an incorrect answer. The number of correct solutions was recorded.

\textsuperscript{177} Only a relatively small number of problems were expected to be completed in the allocated time, however the large number of presented problems ensured even exceptional performance would not exceed the total number of given problems.
6.3 Results

The variables under analysis were number of correct solutions during acquisition, and the number of correct similar and transfer test problems completed. Means and standard deviations are displayed in Table 6.1.

Table 6.1 Acquisition and Test Problems for Experiment 2

<table>
<thead>
<tr>
<th>Phase</th>
<th>Max Score</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Split-Attention</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Acquisition Problems</td>
<td>4</td>
<td>2.19</td>
</tr>
<tr>
<td>Similar Test Problems</td>
<td>100</td>
<td>11.92</td>
</tr>
<tr>
<td>Transfer Test Problems</td>
<td>100</td>
<td>13.50</td>
</tr>
</tbody>
</table>

Note. n = split-attention 26; integrated 26, dual-modality 27

An analysis of variance (ANOVA) with orthogonal planned contrasts was used in all phases to test the primary theoretical prediction that the integrated and dual-modality conditions would prove superior to the split-attention condition. A further contrast was performed between the integrated and dual-modality conditions to test whether their predicted advantages in relation to working memory processing would lead to equivalent performance. The hypotheses were supported across experimental phases. A combination of the integrated and dual-modality conditions performed at a significantly superior level on the acquisition problems in comparison to the split-attention condition, \( t(76) = 2.72, d = 0.67 \). Moreover, there was no significant difference between the integrated and dual-modality conditions \( t(76) = .66 \), indicating that they performed at an equivalent level for the acquisition phase. In contrast to Experiment 1,
students found that the acquisition problems were more difficult; consequently, the earlier identified ceiling effects were no longer evident.

A combination of the integrated and dual-modality conditions also solved a significantly higher number of similar and transfer test problems in comparison to the split-attention condition, $t(76) = 2.52, d = 0.59$, and $t(76) = 2.95, d = 0.67$, respectively. As predicted, there was no significant difference between the integrated and dual-modality conditions for either test problem sets, $t(76) = 0.35$, and $t(76) = 0.33$, respectively. In short, a consistent pattern of hypothesised differences was found across all problems, including problems in which students had to transfer their learning to previously unattempted manipulations of problem solving operators.

Means and standard deviations for cognitive load and efficiency measures are displayed in Table 6.2.

<table>
<thead>
<tr>
<th></th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Split-Attention Integrated Dual Modality</td>
</tr>
<tr>
<td>Phase</td>
<td>M</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td>3.44</td>
</tr>
<tr>
<td>Efficiency</td>
<td>-0.42</td>
</tr>
</tbody>
</table>

*Note. n = split-attention 26; integrated 26, dual-modality 27*

Mean levels of cognitive load across all participants were higher in Experiment 2 compared to Experiment 1 (falling around *Fairly easy to understand*), but once again failed to indicate any significant differences between groups: split-attention vs. combination of integrated and dual-modality conditions, $t(76) = 1.11$; integrated vs. dual-modality conditions, $t(76) = 0.08$. Nevertheless, when performance scores and cognitive load ratings were standardised and combined, a combination of the integrated
and dual-modality conditions was found to be more instructionally efficient than the split-attention condition, \( t(76) = 2.24, d = 0.56 \). As hypothesized, there were no significant differences for instructional efficiency between the integrated and dual-modality conditions, \( t(76) = 0.34 \). Figure 19 plots on a Cartesian plane the relationships between conditions for instructional efficiency. The non-significant differences in mental effort (horizontal plane) show a marginally wider deviation for the split-attention condition and a significant difference in acquisition performance (vertical plane) that favours the combined integrated and dual-modality conditions. This combination of factors was sufficient to discriminate between groups across the hypothetical axis of relative condition efficiency (diagonal line).

**Figure 19.** Relative condition efficiency for Experiment 2 (see Paas and van Merriënboer, 1993).
6.4 Discussion

These data are in accord with the hypothesis that integrated and dual-modal instructional formats facilitate superior understanding of materials that otherwise would require attention to be split between mutually referring sources of information. Moreover, these predictions were generated by CLT, which suggests the central reason for the demonstrated learning advantages rests with both formats better utilising working memory resources. Such an explanation is further strengthened when it is remembered that crucial modifications were made to the current materials based upon a cognitive load interpretation of the findings for Experiment 1. In comparison to the previous study learning difficulty was increased, and as a result differences clearly attributable to instructional format were manifest across all problem sections.

When processing integrated information students avoided the extraneous cognitive load otherwise imposed when required to hold one source of information in working memory while searching for its associated referent. In this way, adjacent and mutually referring instructional elements permitted scarce working memory resources to be directed towards constructing schemas rather than searching for the necessary relations between elements. In addition, restructuring the same information into a dual-modal format made available for learning the expanded processing capacity of combined visual and auditory working memories. The conclusions drawn for both integrated and dual-modality conditions are supported by a large body of research that points to working memory as the fulcrum by which mental leverage can be exerted to ensure maximum cognitive resources are devoted to a given problem (Sweller, 1999; Sweller & Chandler, 1991; Sweller & Chandler, 1994; Sweller et al., 1990).

In their work on multimedia learning, Mayer and associates have identified a
series of effects that also resonate with the findings of this experiment. They demonstrated that coordinating the presentation of visual and verbal materials is superior to presenting the same information as otherwise discrete sources, referring to this phenomenon as the spatial contiguity effect (see section 3.3.2, Moreno & Mayer, 1999a). Moreover, they proposed that the building of referential connections between textual explanation and visual images is central to promoting greater understanding; misaligning information sources induces wasteful search and results in the consumption of limited cognitive resources (Mayer & Moreno, 2003). Viewed from this perspective, the integrated condition in Experiment 2 achieved spatial contiguity by providing a unitary and sequential solution process in which statements and notation were offered as inextricably linked instructional explanations.

Mayer & Moreno (2003) have also investigated the advantages of dual-modality multimedia formats, offering as a solution to the problem of split-attention (spatial dissociation) an approach they referred to as off-loading: In this way, the processing of words as text is redirected to words as narration. They explained this step as a way of reducing the excess processing capacity of one channel (visual) in order to achieve a tempered cognitive load across two channels (auditory-visual). Consistent with this explanation, Experiment 2 data support the advantages of distributing information between two working memory processors. From a cognitive load perspective this off-loading not only distributes information between sensory processes, it effectively expands working memory by allowing more information to be held and manipulated than would be possible from one processor alone (Mousavi et al., 1995).

In relation to dual-modal instruction, ensuring mutually referring elements are delivered at exactly the right time can be procedurally delicate, even more so when auditory input must be aligned with the scanning of visual materials. Mayer & Anderson
(1991; 1992) established the importance of presenting visual and auditory information simultaneously. The synchronisation of dual-modal inputs is referred to by Mayer and Moreno (2003) as the temporal contiguity effect, an effect that they refer to as an instructional “technique for reducing cognitive load” (p. 50). By providing alphabetical markers (visual) in Experiment 2, attention was directed to the appropriate notation at precisely the right moment of narrated explanation, thus ensuring the temporal alignment of both visual and auditory elements\textsuperscript{178}.

From a series of identified multimedia effects over several years, Mayer and colleagues have placed their findings within a theoretical framework that emphasises the limitations of the human processing system and the importance of avoiding cognitive overload (Mayer & Moreno, 2003). Their fundamental explanations arise from a CLT perspective and emphasise the importance of efficiently combining auditory and visual processing channels in order to avoid overwhelming working memory. The results from Experiment 2 are consistent with Mayer’s (2002) cognitive theory of multimedia learning, which in turn acknowledges cognitive load theory as a basis for its conception.

Not only is Experiment 2 important for the empirical evidence that it provides in relation to CLT, but also for the ways in which it can inform instructional design in music. As earlier stated, there is little direct evidence to-date for the efficacy of cognitive load principles in relation to the specific materials of this domain. Nonetheless, the breadth of (CLT) research data so far collected and the characteristics many of these learning environments share with music raise expectations of the utilitarian nature of cognitive load principles. For example, Mousavi et al. (1995) and

\textsuperscript{178} See Jeung et al. (1997) and their use of a flashing technique to achieve a similar outcome (see section 3.3.5).
Tindall-Ford et al. (1997) established that the expanded working memory hypothesis remained robust across a range of information sources, including diagram and narration, text and narration, and text and tables (see section 3.3.5). There was little reason to believe that in Experiment 2 dual-modality principles would not be equally effective when applied to musical notation and explanatory statements. The current experiment has taken a key step in specifically extending theoretical predictions to the instruction of music. In general terms it adds further evidence to the tenet of working memory as the central consideration of instructional design.

In summary, Experiment 2 provided support and extended to music two core theoretical principles. First, integrated or contiguous presentation of information lowers cognitive load and builds the necessary connections to promote schema development (split-attention effect). Second, employing two sensory modes rather than one expands working memory capacity and increases the likelihood of mentally integrating instructional elements (modality effect). The implications of these findings carry some ecological consequence to a field in which the separation of musical notation and textual explanation remains commonplace. Text and reference books prefer the visual clarity of a layout that allows musical notation to be more easily reviewed and auditorially translated without distractions. Admittedly, music by nature raises presentational challenges when designing integrated formats; impediments to the mental streaming of notation might hinder understanding. Nonetheless, where a series of elements need to be combined for intelligibility, thoughtful integration of an instructional sequence will help allocate maximum working memory resources to the required learning. Likewise, well-structured auditory explanation incrementally delivered at the visual point of need should be preferred to a haphazard approach that does not consider the necessary simultaneity, size, sequence or association of
instructional elements.

The significance of these findings for music will become apparent in the next experiment where for dual-modal materials the auditory input of speech is replaced by recorded music. The results from Experiment 2 give reason to believe that potentially high cognitive loads in music, which otherwise exhaust processing capacity, might be offset by the cognitive efficiencies of simultaneous and dual-modal deliveries.
Chapter 7 Experiment 3

7.1 Introduction

The purpose of Experiment 3 was to compare uni- and dual-modal instructional formats. There were four conditions differentiated by modality (uni and dual) and presentation order (simultaneous and successive). In this experiment, dual-modal materials were structured as written explanatory text and auditory musical excerpts. Presenting explanatory text in written rather than spoken form during dual-modal presentation was a marked deviation from conventional CLT experiments. For example, previous dual-modal presentations usually consisted of diagrams and spoken text (Tindall-Ford et al., 1997); as indicated, in the current experiment the text was written rather than spoken, and combined with auditory music.

Uni-modal materials likewise consisted of explanatory text and musical excerpts, although all information was presented in the auditory mode. Again, this presentation contrasted with the normal uni-modal format used in previous CLT experiments, which usually consisted of two visual presentations such as a diagram and written text (Tindall-Ford et al., 1997). In the current experiment, the text is spoken rather than written, and is combined with auditory music.

For Experiment 2, the dual-modal format used a combination of spoken (auditory) text and musical notation (written) and it was assumed that this format conformed to the archetypal CLT design of spoken text and diagram. As hypothesised, the results of this experiment were in accord with other studies (Mayer & Anderson, 1991; Mayer & Simms, 1994; Mousavi et al., 1995; Tindall-Ford et al., 1997), and they confirmed the benefits of dual-modal presentation under these conventional conditions.

The use of either uni- or dual-modalities for text and auditory musical excerpts
is common to instructional practices in music. Even so, the working memory demands and therefore instructional effectiveness of combining these information sources in either modal arrangement were unknown. In relation to music, few if any empirical precedents have emerged and therefore addressed these issues. Nonetheless, considerable work (Frick, 1984; Jeung et al., 1997; Moreno & Mayer, 1999a; Mousavi et al., 1995; Tindall-Ford et al., 1997) has demonstrated across a range of non-musical materials that dual-modal formats expand working memory capacity and facilitate learning. In addition, research has provided evidence of dual-modality effects in relation to materials other than the more typical combination of auditory text and diagrams. For example, Tindall-Ford et al. (1997, Experiment 2) found that combining auditory text with tables was superior to written text and tables, and Mousavi et al. (1995, Experiment 5) found a similar dual-modal advantage even where both sources of information were comprised of verbal statements. Based on the findings for these diverse materials it was predicted that similar advantages would be evident for the arrangement of dual-modal materials designed for Experiment 3. In contrast, it was expected that working memory capacity would be limited when explanatory text and auditory musical excerpts shared the same auditory processing modality.

Experiment 3 materials were considered high in element interactivity. Previous studies (Chandler & Sweller, 1992; Mayer & Moreno, 1998) have established the instructional efficiency of either integrated or simultaneous formats when presenting materials that are high in element interactivity, both strategies avoiding the extraneous cognitive load associated with split-attention. Furthermore, Mousavi et al. (1995) established that dual-modal presentation was superior to uni-modal presentation irrespective of either simultaneous or successive delivery. These results suggest that in Experiment 3 successive presentation of uni-modal materials would incur the double
limitations of a split-attention design and the enforced processing of both information sources through the auditory mode. Accordingly, it was an important aim of this experiment to investigate the instructional implications of successively presenting either dual- (written-auditory) or uni- (auditory-auditory) modal music materials. Specifically, it was predicted that a dual-modal format would help offset the negative effects of a split-attention design.

In summary, the results of Experiment 2 provided support for the central claim that understanding is dependent upon the presentation of instructional materials in such a way as to maintain cognitive load within working memory limitations. Experiment 3 sought to extend these principles to conditions in which instructional factors were manipulated in ways germane to the teaching of music.
7.1.2 The Musical Aptitude Profile - Measuring Audiation Skills (Experiments 3-6).

Gordon (1993) defined musical aptitude and musical achievement respectively as “a measure of a student’s potential to learn” (p. 2), and a measure of what has been learned. He also argued that environmental influences combine with native potential up until about the age of nine to produce a developmental period of musical aptitude (Gordon, 1979)\(^{179}\); thereafter, musical aptitude remains stable, irrespective of further environmental influences. For instructional design in music (Experiments 3-6), recognising and understanding the important distinction between musical aptitude and musical achievement is a factor requiring careful consideration.

According to Gordon (1993), audiation is fundamental to musical aptitude and to musical achievement. Gordon (1995) defined audiation as,

the ability to hear and to comprehend music for which the sound is not physically present (as in recall), is no longer physically present (as in listening), or may never have been physically present (as in creativity and improvisation) (p. 8).

Furthermore, the process of audiation is not the same as musical perception; audiation involves the comprehension of musical stimuli and as a consequence follows the perception (hearing) of music. Given that music is streamed or perceived over time, audiation and perception can occur simultaneously (Gordon, 1993).

In order to objectively evaluate a student’s audiation skills, the primary construct of musical aptitude, Gordon (1995) developed the *Musical Aptitude Profile* (MAP). MAP is an objective battery of tests comprising three areas of musical audiation: tonal imagery, rhythm imagery and musical sensitivity\(^ {180}\). These tests have undergone

---

\(^{179}\) Both factors “contribute in unknown proportions to a child’s music aptitude” (Gordon, 1984, p. 1).

\(^{180}\) Gordon (1995) stated that the current terminology was coined before the verb *to audiate* was adopted in research (ca. 1976).
adaptations (e.g., Schleuter, 1983), evaluations (e.g., Gordon, 1986) and revisions (see Gordon, 1995, pp. 22-34) to establish and ensure their reliability and validity as an objective measure. Of importance within otherwise heterogeneous populations, the relations between previous musical experience, achievement and stabilized audiation skills do not warrant separate aptitude norms for these factors.

For Experiments 3 - 6 (Chapters 7 – 9), students were required to respond to auditory stimuli. In all these experiments, the independent variables under study were factors subject to various instructional interventions. Consequently, it was critical to the interpretation of results to exclude aptitudinal factors (stabilised audiation skills) that would otherwise bias collected data. To control for these factors, applicable areas of the MAP were administered prior to each experiment and, along with survey data, helped to compile an accurate profile of relevant participant characteristics. In turn, these characteristics were randomised and equally distributed across treatment groups. Gordon (1995) suggested that there is little if any relationship between academic aptitude and musical aptitude and that students need only cultural exposure to music for MAP to make a valid assessment of auditory abilities181.

Testing musical aptitude was but one important stage in strengthening the validity of the empirical work reported in the following chapters; identifying a student’s level of musical ability also strengthens the ecological validity of delivering materials that are conducive to the acquisition of knowledge and skills (Gordon, 1993). The following experiments explore within cognitive load theory the means by which appropriate instructional interventions facilitate the acquisition of such knowledge and skill within the domain of music.

181 Scores from these tests are standardised and given percentile ranks; they are based on a large representative sample of students (Gordon, 1995).
7.2 Method

7.2.1 Participants.

Eighty-five Year 7 students from an academically selective boys’ high school in Sydney participated in this experiment. As with previous experiments, students were selected as musically experienced in relation to their peers, ensuring current levels of understanding were sufficient to complete the problems once additional study was undertaken. All other participant characteristics and procedures for selection were as outlined in Experiment 1.

7.2.2 Materials and procedure.

Experiment 3 materials differed markedly from those described in the first two experiments. Two phases, consisting of five instructional examples, and similar (two sets of five problems) and transfer (two problems) test problems respectively, were presented without a break. The instruction phase consisted of four experimental treatments: simultaneous and successive dual-modal conditions (visual-auditory), and simultaneous and successive uni-modal conditions (auditory-auditory). All conditions demonstrated ways in which manipulating musical concepts produced melodic variation. In this experiment, melodic variation was defined as a new version of an original melody. The new melodic version changed one or more pitch features while preserving an identifiable aural relationship between the original melody and its variation.

The experiment was administered in four discrete sessions due to the different requirements of delivering auditory and visual materials. The instruction phase was delivered as either a Powerpoint presentation (dual-modal conditions), or by compact disc recording (uni-modal conditions). Powerpoint presentations were viewed within the
classroom on a large screen via a data projector. Test problems were presented on A4 sheets of paper and the associated auditory material was played on compact disc. The experimenter developed all instructional examples and test problems (including melodic variations). These materials built upon previous learning undertaken as part of the Year 7 curriculum in music.

The instruction phase explained the way in which five pitch concepts varied a melody. Each of the concepts—intervals, range, transposition, contour and modulation—was presented as a new instructional episode and without reference to previous concepts. Accordingly, all uni-modal materials were recorded as a continuous auditory sequence of learning episodes, whereas dual-modal episodes were viewed as new Powerpoint slides. The explanation of each concept was accompanied by two short comparative auditory excerpts (4 bars in length), which demonstrated how the original melody (first excerpt) was varied (second excerpt).

The melodic material (auditory excerpts) heard throughout the experiment was based on *Twinkle Twinkle Little Star*. A familiar nursery tune was deliberately chosen to ensure students already possessed a melodic schema for the original thematic material. In this way, working memory capacity was directed toward the varied pitch concepts rather than the need to concurrently build schemas for the melody itself. Estimating the cognitive load of materials was an essential consideration of this experimental design. To this end, it was assumed that increasing either the complexity or unfamiliarity of melodic material was synonymous with increasing intrinsic cognitive load.

In relation to the comparative auditory excerpts, students listened to the original melody, and then to the melodic variation after a brief pause. The variation illustrated the pitch concept briefly described in either of two corresponding auditory or written explanatory statements (Figure 20). Apart from variations based on the targeted pitch
concepts, the melodic materials otherwise remained the same (Appendix K.).

**Contour Change**

i. In the variation the pitch shape forms a rising contour. The original melody gradually rises and then falls, i.e. an arch-shaped contour.

ii. The first phrase is exactly the same in both the original melody and its variation.

**Large Range Change**

i. Repeated notes of the variation’s second phrase move to a lower pitch. These notes widen the overall range between the melody’s lowest to highest note.

ii. The first phrase is exactly the same in both the original melody and its variation.

*Figure 20.* Two examples (No 2 & No 4) of the explanatory statements from Powerpoint slides (dual-modal condition) used in Experiment 3.

Students were required to either read (dual modality) or listen (uni-modality) to explanatory statements and find the associated musical referents by identifying differences between the two comparative auditory excerpts. Holding the first excerpt in working memory while listening to the second excerpt, and then extracting the indicated variations, represented a task high in element interactivity. Each pair of comparative auditory excerpts, including a brief pause in between, lasted approximately 18 s. With respect to simultaneous conditions, the period of time allocated for each of the comparative auditory excerpts was also sufficient to either read or listen to the associated explanatory statements.

The two dual-modal conditions had 22 participants and the uni-modal simultaneous and successive conditions 21 and 20 participants respectively. The process
by which students were allocated to treatment groups departed sharply from the method of randomisation previously employed. All students were administered a melodic audiation test (Test T – Tonal Imagery, Part 1 Melody) from the Musical Aptitude Profile (7.2, Gordon, 1995), and ranked on their standardised scores. They were then assigned to one of four conditions by progressive allocation (1234…2341…3412 etc). Given the nature of the auditory materials, it was assumed that melodic audiation skills were an important factor in a student’s ability to learn the materials described above. They were not, however, a factor under analysis in this experiment. For that reason, distributing these skills across conditions minimised any interference with the intended experimental manipulation.

For the instruction phase, the explanations and auditory excerpts were identical for each condition, and were delivered by way of differing instructional formats. The two uni-modal conditions listened to the explanatory statements for each pitch concept either before (successive), or while (simultaneous) listening to the comparative auditory excerpts on compact disc. The dual-modal conditions read the explanatory statements for each pitch concept either before (successive), or while (simultaneous) listening to the comparative auditory excerpt as an audio file (via Powerpoint). Successive explanatory statements and auditory excerpts followed each other without a break.

The successive conditions were not permitted to review (hear or view) the explanatory statements once the auditory excerpts commenced. For the simultaneous conditions, the auditory excerpts commenced when students first either heard or viewed (uni- and dual-modalities respectively) the explanatory statements. This procedure guaranteed that both simultaneous and successive conditions had the same exposure to, and therefore equal time to study, each of the explanatory statements. As a result, the total instructional times for simultaneous conditions and for successive conditions were
unequal. That is, the simultaneous conditions were allocated only the time necessary to hear the auditory excerpts, whereas the successive conditions were given the elapsed time of the auditory excerpts and explanatory statements.

Following the completion of the first phase, students completed a 9-point subjective cognitive load rating scale, which was used as a basis to measure cognitive load and instructional efficiency during the instruction phase (as per the explanation for Experiment 1).

A test phase followed the instruction phase. Test materials were divided into similar and transfer problems and were identical for all four conditions. Similar problems were further divided into two sets of five problems each, and transfer problems were presented as two individual problems. In keeping with the instructional episodes, all problems consisted of comparative auditory excerpts that required students to identify one or more ways of varying a melody.

For the first set of five similar problems, students were also asked to provide any aspect that from one auditory excerpt to the next (comparative excerpts) remained the same. Identifying two elements increased problem difficulty, however it was expected, irrespective of instructional condition, that students would have little difficulty matching melodic features that remained the same. By design, the auditory excerpts were primarily comprised of common features, any of which could be readily confirmed while listening to the second excerpt. In contrast, identifying variations required extracting differences between auditory excerpts and then associating these with the earlier instructional explanations. The cognitive load associated with extracting differences was expected to be high and therefore differentiate between conditions.

For the second set of five similar problems, students had to identify two melodic variations that were based on previous instructional materials. Identifying two variations
(features that remained the same were no longer required) further raised problem difficulty and, as a direct indication of instructional understanding, was also predicted to differentiate between the four experimental conditions. The first set of similar test problems (played once) were administered without a break, and then followed after the experimenter’s brief explanation by the second set of similar problems (played once). Each set was completed within a short period of time (approximately 3 min 10 s).

The two transfer problems each consisted of a comparative auditory excerpt that included all five previously explained pitch variations. Identifying five variations raised problem difficulty even further, deviating considerably from the earlier instruction that delivered concept explanations and auditory excerpts one at a time. Aural analysis of this kind usually involves multiple overlapping solutions that are characteristic of ill-defined problems. Furthermore, variations that constitute one pitch concept are not (and cannot be) entirely independent of other musical concepts. Therefore, in order to have demonstrated a comprehensive and higher order understanding of transfer problems, students had to consider the interrelationships of multiple pitch concepts.

The first and second transfer problems were similar in all respects but for the melodic materials on which variations were based. The second transfer problem introduced a new melodic excerpt to replace the previously heard *Twinkle Twinkle Little Star*. Care was taken to retain the componential structure of the new melody and to avoid altering the intrinsic difficulty of the earlier melodic material (see Appendix L.). Naturally, without the schemas of a familiar melody, it was assumed that greater working memory resources would be devoted to both processing the comparative auditory excerpts and to finding relations between them. In this way, a problem continuum was achieved, spanning near to far transfer of instructional materials. For both transfer problems each pair of comparative auditory excerpts was played four
times, the first three of which were followed by a 30 s pause, and the final playing by a 45 s pause.

For all problems, once the time limit for one problem had expired the next problem was presented without further explanation. Students did not have access to instructional materials during the test phase. All groups were tested separately and the timing was dictated by the delivery of auditory materials. For similar problems, one mark was awarded for each correctly identified variation, giving 15 marks in total (5 marks and 10 marks for the first and second sets respectively). One mark was also awarded in the first set of similar problems for a correctly identified feature that remained the same, resulting in a possible mark of 5. Five pitch concepts were identifiable for each transfer problem, and for each of these, students also had to explain how and where the pitch concept was varied. Therefore, three marks were awarded for each pitch concept, giving a total of 15 marks (3x5) for each problem. The number of correct solutions was recorded. All questions were objectively marked and no marks were allocated for an incorrect answer.
7.3 Results

The variables under analysis were the number of correct similar and transfer test problems completed. For all problems, solutions were based upon identified variations between auditory excerpts. On the basis of these identified variations, hypothesised differences between conditions were tested. Additionally, an independent analysis of features that remained the same between comparative auditory excerpts was performed for the first set of similar problems.

The two sets of similar problems were of one type, increasing from one to two the number of pitch concepts that were varied. Similar problems were therefore reported as a single test total of identified variations. The first and second transfer problems were reported separately due to the change of melodic material from the first to second problems. Means and standard deviations for both features that remained the same (first set of similar problems) and identified variations are displayed in Tables 7.1 and 7.2 respectively.

Table 7.1 First Set of Similar Test Problems for Experiment 3:
Features that Remained the Same

<table>
<thead>
<tr>
<th>Features that Remained the Same</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phase</td>
<td>Max Score</td>
<td>Dual-Modality</td>
<td>Dual-Modality</td>
<td>Uni-Modality</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Simultaneous</td>
<td>Successive</td>
<td>Simultaneous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>First Similar Test Problems</td>
<td>5</td>
<td>2.23</td>
<td>1.66</td>
<td>1.95</td>
<td>1.53</td>
</tr>
</tbody>
</table>

Note. n = dual-modality simultaneous 22; dual-modality successive 22; uni-modality simultaneous 21; uni-modality successive 20
An analysis with sets of orthogonal planned contrasts was used in all phases to test predicted differences between uni- and dual-modal presentations. Two contrast tests examined the hypothesised differences for mode between the two simultaneous conditions and between the two successive conditions respectively. A third contrast test examined differences between the two combined simultaneous and the two combined successive presentations.

For the first set of similar test problems, there were no significant differences between conditions for features that remained the same. As hypothesised, an equivalent number of mean solutions were recorded between the two successive conditions, \( t(81) = 0.82 \), between the two simultaneous conditions, \( t(81) = 0.12 \), and between a combination of simultaneous and a combination of successive presentations, \( t(81) = 0.31 \). All conditions irrespective of format were capable of identifying similarities between the two auditory excerpts.

In contrast, a significant difference for mode was found in relation to identified variations for similar test problems (first and second sets combined). The uni-modal
successive condition identified a significantly higher number of mean variations than the dual-modal successive condition, \( t(81) = 2.28, d = 0.72 \). However, there were no significant differences for mode between the two simultaneous conditions, \( t(81) = 0.76 \), nor between a combination of simultaneous and a combination of successive presentations, \( t(81) = 1.35 \). This series of results ran somewhat contrary to the expectation that dual-modal and simultaneous formats would facilitate superior understanding.

For the first transfer test problem there were no significant differences for mode between the two successive conditions, \( t(81) = 0.51 \), between the two simultaneous conditions, \( t(81) = 0.12 \), nor between a combination of simultaneous and a combination of successive presentations, \( t(81) = 0.44 \). The largest non-significant difference in the mean number of identified variations occurred between the successive conditions, which preserved the direction of results found for combined similar problems. Moreover, the earlier significant difference for mode between successive conditions was evident for the second transfer test problem, again favouring uni-modal presentation, \( t(81) = 2.03, d = 0.63 \). As with previous problems, there were no significant differences for mode between the two simultaneous conditions, \( t(81) = 0.08 \), nor between a combination of simultaneous and a combination of successive presentations, \( t(81) = 0.66 \).

Means and standard deviations for cognitive load (subjective ratings) and efficiency measures are displayed in Table 7.3. The efficiency measure combined the standardised subjective cognitive load rating for the instruction phase with the standardised results from the total test problems.
Table 7.3 Cognitive Load and Efficiency Measures for Experiment 3

<table>
<thead>
<tr>
<th>Phase</th>
<th>Cognitive Load</th>
<th>Efficiency Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>Dual-Modality Simultaneous</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>Dual-Modality Successive</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Uni-Modality Simultaneous</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>Uni-Modality Successive</td>
<td>5.60</td>
</tr>
</tbody>
</table>

Note. n = dual-modality simultaneous 22; dual-modality successive 22; uni-modality simultaneous 21; uni-modality successive 20

An analysis of cognitive load indicated a marginally significant difference for mode between the two successive conditions, $t(81) = 1.97, d = 0.64$ (one-tailed), and a significant difference between the two simultaneous conditions, $t(81) = 3.97, d = 1.16$, both of which favoured dual-modal presentation. There was no significant difference between a combination of simultaneous and a combination of successive presentations, $t(81) = 0.32$. These data supported the hypothesis that uni-modal conditions would experience higher levels of cognitive load due to processing both sources of information through the same processing modality within working memory. In particular, the effect size for mode between simultaneous conditions was in the large range ($> 0.8$).

When standardised cognitive load ratings and performance scores (variations) were used to calculate instructional efficiency for all test problems, a significant difference for mode between the simultaneous conditions favoured dual-modal presentation, $t(81) = 2.50, d = 0.68$. There were no significant differences for mode between successive conditions, $t(81) = 0.09$, nor between a combination of simultaneous and a combination of successive presentations, $t(81) = 0.63$. These findings are consistent with the notion that simultaneously processing instructional elements through two modes is cognitively more efficient than processing the same
information through a single auditory mode. It would also appear that in this experiment other factors played an important role in determining the interaction between cognitive load and performance. These issues are briefly pursued in the discussion section below.

Figure 21 plots on a Cartesian plane efficiency measures for all test problems, which reflect in relation to the diagonal axis (hypothetical efficiency baseline) a moderate contrast between the relatively efficient simultaneous dual-modal format and, in particular, its uni-modal counterpart. Successive uni-modal and simultaneous dual-modal presentations registered neutral efficiency due to the self-countering advantages of test performance and mental load respectively.

![Diagram showing efficiency measures for all test problems, Experiment 3](image)

**Figure 21.** Relative condition efficiency for all Test Problems, Experiment 3 (see Paas and van Merriënboer, 1993).
To further highlight the difference between modalities for all test problems, Figure 22 plots instructional efficiency on an $x, y$ plane. As can be readily seen, there was a marked difference for instructional efficiency between simultaneous conditions.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure22.png}
\caption{Relative condition efficiency for all Test Problems, Experiment 3.}
\end{figure}
7.4 Discussion

Experiment 3 used CLT to test the relative advantages of uni- and dual-modal instructional formats in music. Specifically, these materials explored the predicted advantage of dual-modal presentation: the expanding of working memory capacity and the consequent lowering of cognitive load. To this end, it was hypothesised that dual-modal formats would be superior to uni-modal formats irrespective of whether materials were either simultaneously or successively presented.

For the contrasts between simultaneous formats there was evidence that dual-modal presentation was superior to its uni-modal counterpart. Support for this hypothesis came from the results of the subjective rating scale and efficiency measures, although it was predicted to arise primarily from a superior capacity to complete test problems. Despite both simultaneous conditions solving an equivalent number of problems, the students studying uni-modal materials experienced significantly higher levels of cognitive load than the students studying dual-modal materials. Moreover, it was found that studying dual-modal materials resulted in relatively higher levels of instructional efficiency, indicating that students possessed an increased capacity to process instructional materials. Together, these data are consistent with the theoretical principles that generated these experimental conditions.

For the contrasts between successive formats the predicted dual-modal advantages were unsupported. Students from the uni-modal condition demonstrated superior understanding by identifying significantly more similar and second transfer test problem variations than their dual-modal counterparts. Unexpectedly, this superior performance occurred even though the uni-modal condition recorded significantly higher cognitive load ratings. As a result, lower levels of cognitive load were not
associated with any instructional advantages, a marked contrast to the difference outlined above for simultaneous conditions. It is of interest that both uni- and dual-modal conditions experienced temporal split-attention, and therefore successive presentation alone cannot explain these differences for understanding.

In relation to successive formats the combination of higher cognitive load and superior test performance is difficult to reconcile with the hypotheses generated by CLT. These predications were based on the instructional priority of maintaining intrinsic and extraneous cognitive loads within working memory capacity. Consider however the possibility that the total cognitive load of materials was lower than expected, and as a consequence all conditions were able to readily process instructional materials. In such a case, any remaining mental capacity can be redirected toward the active construction of schemas, an aspect of mental effort referred to as germane cognitive load (Sweller et al., 1998). Viewed from this perspective, the uni-modal (successive) condition’s higher subjective ratings may, at least in part, represent an increased capacity to construct the necessary schemas for deeper understanding. Although such an interpretation is plausible for successive conditions, it is inconsistent with the data for simultaneous conditions, suggesting further investigation is required to accurately identify the nature of differential cognitive load under these experimental conditions.

In addition to the investigation of modality, contrast tests explored the difference between the two simultaneous and two successive formats, irrespective of modal presentation. Testing between these formats investigated whether the decreased instructional time of simultaneous formats occurred at the potential expense of imposing higher cognitive load. Conversely, successive formats were given the potential benefits of increased learning time, but at the expense of splitting attention and therefore raising
extraneous cognitive load. Nonetheless, testing found no significant differences for the mean number of test problem solutions for presentation order. It is notable, however, that non-significant means favoured a combination of simultaneous conditions in all but the first transfer problem, and that equivalent performance was achieved despite the successive conditions receiving almost double the period of instruction time.

In total, the data initially suggest that instructional benefits of modality are subject to the way in which the two sources of information are presented: dual modality favoured simultaneous presentation and uni-modality favoured successive presentation. However, when dual-modal materials are presented simultaneously the potential benefits appear, from the current evidence, to outweigh other simultaneous and successive formats. For example, instructional efficiency data (see Figures 21 and 22) indicated that students studying these materials enjoyed the advantage of relatively strong test performance while incurring a relatively lower load on mental processing. This difference for instructional efficiency is particularly marked between the two simultaneous conditions. Additionally, if compared to successive presentation, the dual-modal simultaneous condition achieved equivalent understanding within a substantially reduced period of instructional time.

The relative advantages indicated for simultaneous dual-modal presentation are in accord with CLT and suggest that this format offers a cognitively and temporally efficient method of learning. Equally, it is recognised that in future experiments of this nature that the conclusions drawn in relation to this condition would be strengthened if a capacity to solve a significantly greater number of test problems were also indicated. In relation to the successive conditions the interactive nature of these results warrants further attention. For instance, there is evidence from this and later experiments that splitting attention between dual-modal materials adversely affects understanding,
whereas splitting attention between uni-modal (auditory) materials produces better than expected results. Both these findings require further examination to accurately establish the conditions under which dual- and uni-modal formats provide an instructional advantage.
Chapter 8 Experiment 4

8.1 Introduction

The three previous experiments investigated ways in which combining two mutually referring information sources promoted understanding of music materials. Accordingly, instructional formats combined verbal explanations with either auditory musical excerpts or visual musical notation, and these were presented in either uni- or dual-modal formats. The primary aim of Experiment 4 was to explore ways in which combining all three information sources—text, musical notation and auditory excerpts—into a unified instructional format facilitated understanding. Fundamental to this investigation was whether the addition of auditory excerpts as a third information source imposed excessive levels of cognitive load on working memory.

To achieve this aim, three conditions were differentiated by modality and/or presentation order: The two dual-modal conditions (immediate auditory and delayed auditory) were structured as written explanatory text, musical notation and auditory excerpts (three sources). The uni-modal condition (no auditory) was structured as written explanatory text and musical notation (two sources). For all conditions, the auditory and/or visual instructional elements were presented simultaneously; the written materials of explanatory text and musical notation were presented in a visually integrated format.

Irrespective of how many information sources are combined, understanding remains bound by the complexity of materials (intrinsic cognitive load) and whether this complexity exceeds the limited processing capacity of working memory. For this reason it is unsurprising to find few CLT studies in which instructional formats combined more

---

182 From this point the term “auditory excerpt” implies musical excerpt.
than two sources of information. A priori, increasing the amount of information (quantity and/or sources) increases the possibility of overloading working memory.

Despite these possible limitations, there were reasons to believe that under certain conditions (delayed auditory) the nature of music materials would facilitate the processing of three information sources and avoid the negative impact of excessive cognitive load. First, the musical excerpts (sound) and their associated musical notation (symbol) were visual-auditory equivalents, each reinforcing the other in relation to explanatory text. Although equivalent materials can result in the negative effects of redundancy (Kalyuga et al., 1999; Yeung et al., 1997), the young music students in this experiment (Year 7) possessed low levels of musical expertise (conceptual and perceptual). Consequently, the combination of auditory and graphic musical representations was predicted to facilitate the necessary connections between instructional elements. In particular, it was considered easier to initially locate some musical elements on a static visual source, a strategy that also was employed for Experiment 2. Even so, these equivalent musical sources were not expected to expend the same processing capacity required when relating instructional elements between two mutually referring but independent sources, e.g., explanatory text and musical notation.

Second, many students find it difficult to form an accurate mental aural realisation from musical notation alone, i.e., hearing without sound. It was therefore expected that providing auditory excerpts would support the identification and integration of musical elements. Similarly, it was assumed that musical notation, as a static visual source, helped students identify what otherwise might be perceived as abstract auditory concepts. For novice music students, it was believed that relating instructional elements between musical examples and explanatory text was better facilitated by the complementary sources of auditory excerpts and musical notation than
by either source in isolation.

For Experiment 4 the to-be-learned concepts were expected to raise intrinsic cognitive load, and although the inclusion of both musical notation and auditory excerpts was expected to promote understanding, the requirement to simultaneously (immediate auditory condition) process all three elements (including explanatory text) was predicted to overload working memory. Simply attending to all elements at the same time, let alone finding the relations between them, was thought to be extremely complex.

As Pollock et al. (2002) remind us, the necessary assimilation of interacting elements will fail when working memory limitations are exceeded during complex learning. Under these circumstances it may be necessary to present these interacting elements in an isolated non-interacting form, even where, as a consequence, overall understanding is potentially compromised. Therefore, it was hypothesized that a simultaneous format in which the introduction of the auditory excerpts was delayed would moderate the increased cognitive load associated with otherwise initially processing three simultaneously presented sources of information.

At first, a format in which the introduction of auditory information was delayed appeared to contradict the well-established principles of either physically or temporally integrating mutually referring instructional elements (Moreno & Mayer, 1999a; Sweller et al., 1990; Ward & Sweller, 1990). In fact, all three conditions in this experiment used integrated designs, however the delayed auditory format sought to distribute the mental load by integrating the text and musical notation before introducing the auditory excerpts. In other words, if schemas associated with the text and musical notation are constructed first, it may be easier to later construct the essential higher-order schemas incorporating auditory music. Moreover, because the text and musical notation
remained while the auditory excerpt was heard, this format was expected to either minimize or exclude physical (visual-visual) and temporal (visual-auditory) split-attention. Therefore, only the placement (earlier or later) of the auditory excerpt differentiated between the two dual-modal conditions. Also, as outlined above, the auditory excerpts provided equivalent information to the musical notation rather than introducing new instructional elements per se. As a result, the total mental load for the delayed auditory condition was predicted to remain within processing capacity.

The no auditory condition (integrated text and musical notation) acted as a virtual control; this condition utilised only two written sources of information (unimodal) and the arrangement of these sources conformed to a previously successful instructional format (Experiment 2). It was hypothesised that these materials would be sufficient for students to understand the instructional procedures when given a written test, whereas their understanding of the same procedures would be relatively diminished when measured by an aural test. Music is essentially an auditory medium and therefore a comprehensive understanding of music requires responding to written, and in particular, aural materials. Accordingly, this experiment sought to establish a way in which multiple musical elements could be effectively delivered without the negative effects of extraneous and excessive cognitive loads.
8.2 Method

8.2.1 Participants.

Fifty-five Year 7 students from an academically selective boys’ high school in Sydney participated in this experiment. As with previous experiments, students were selected as musically experienced in relation to their peers, ensuring current levels of understanding were sufficient to complete the problems once additional study was undertaken. All other participant characteristics and procedures for selection were as outlined in Experiment 1.

8.2.2 Materials and procedure.

Experiment 4 materials were similar to those used in the first two experiments, explaining the function of upper and lower time signature numbers in relation to a bar of rhythm. Experimentally, it was easier to control instructional factors using the algorithmic solution procedures of these earlier materials (see Experiment 2). There were two phases: The first phase consisted of six instructional examples, which was then followed by the second phase consisting of written problems (two sets of twenty problems each) and auditory problems (two sets of six problems each). All problems were presented without a break. The instruction phase had three experimental treatments: immediate auditory (visual-visual-auditory), delayed auditory (visual-visual-auditory), and no auditory (visual-visual). In contrast to earlier experiments, some conditions were a combination of three (audio-visual, dual-modal) rather than two (visual only, uni modal) information sources, including explanatory statements, musical notation and auditory excerpts.

The experiment was administered in three discrete sessions due to the differential use of auditory excerpts. The instruction phase was delivered as a
Powerpoint presentation for all conditions. Written test problems were presented on A4 sheets of paper. Aural test problems were answered on A4 sheets of paper and the auditory excerpt for each question was delivered on compact disc. The experimenter developed all instructional examples and test problems, and the materials built upon previous learning undertaken as part of the Year 7 curriculum in music.

The instruction phase consisted of six examples (slides 1-6) that explained simple and compound time signatures and the values of notation in relation to a given unit of beat. The first three examples explained the functions of upper (slides 1 and 2) and lower (slide 3) numbers in relation to a time signature. The fourth to sixth examples (slides 4-6) followed an identical procedure, explaining the same functions in relation to a time signature (Figure 23).

a. Listen as the rhythm is played without the quaver beat and then with the quaver beat (lower notes).

b. Like a fraction, the lower number of a time signature indicates the type of note value by which all other notes will be measured.

= eighth note or quaver.

Figures 23. An example of the integrated written materials (all conditions) used in Experiment 4 (text size is reduced in relation to original Powerpoint projection).

c. Therefore, in this example, the quaver is used to measure the length of all other notes.

Care was taken to make each explanatory statement brief and intelligible, allowing students to simultaneously refer to the musical notation. For that reason, all conditions received the written text and musical notation as an instructionally efficient visually integrated format (as for Experiment 2). Additionally, the length of musical examples was restricted to either one or two bars in order to efficiently illustrate the
instructional procedures relating to time signatures. Each of the three examples for both \(\text{\textbullet} \) and \(\text{\textbullet} \) time signatures progressively contained more explanatory text. Learning sequences ranged from 17 s to 62 s (including auditory excerpts), and in total the instruction phase lasted for 3 min 36 s. No reference to earlier examples was allowed once a Powerpoint slide was advanced.

In contrast to Experiment 3, single rather than comparative auditory excerpts were used to illustrate explanatory text. Even so, element interactivity was considered high, arising from the manipulation of multiple explanatory elements and their relations to associated musical examples. What is more, the two dual-modal conditions (immediate auditory and delayed auditory) required students to relate elements from three sources and between two modes. At issue in this experiment was the number of elements that could be mentally integrated at one time, and whether the cognitive load generated by the equivalent elements of auditory excerpts and musical notation would remain within working memory capacity. For these reasons, the efficient no auditory (uni modality) format, incorporating two information sources, provided a direct comparison with formats that combined three information sources (dual-modality), including the addition of auditory excerpts.

The two dual-modal conditions (immediate auditory and delayed auditory) each had 18 participants and the uni-modal condition (no auditory) 19 participants. Students were allocated to treatment groups according to the process outlined in Experiment 3. All students were administered rhythmic audiation tests (Test R – Rhythmic Imagery, Part 1 Tempo, Part 2 Meter) from the Musical Aptitude Profile (Gordon, 1995) and ranked on their standardised scores. They were then assigned to one of three conditions by progressive allocation (1234…2341…3412 etc). It was again assumed that audiation skills helped predict a student’s potential to learn the instructional material. Audiation
skills were not a variable under analysis in this experiment and consequently distributing these skills across conditions minimised any interference with the intended experimental manipulation.

For the instruction phase, three conditions received identical explanations and musical notation by way of differing formats. All students were required to read the explanatory statements and relate them to the adjacent integrated notation. The no auditory condition received only written materials (explanatory statements and musical notation), whereas the immediate auditory condition also listened to the associated auditory excerpts. For the delayed auditory condition, the explanation and musical notation were presented first and then remained while the associated auditory excerpt was heard. All conditions presented each instructional example on a separate Powerpoint slide, and for the dual-modal conditions (immediate auditory and delayed auditory) auditory excerpts were incorporated as audio files.

Instructional times were identical for all three experimental treatments; these times were dictated by the delivery of auditory excerpts and the total elapsed time allocated to the delayed auditory condition. As a result, the no auditory and immediate auditory conditions initially received all materials (auditory and/or written) followed by additional time to study the written materials. Therefore, for the immediate auditory and delayed auditory conditions only the temporal location of the auditory excerpts varied in relation to the written information sources. For both these conditions, the written information remained permanently visible. The auditory component for the immediate auditory condition was presented as soon as the written material appeared, followed by a period in which only the written material was presented. For the delayed auditory condition, only the written materials were initially presented, followed by the introduction of the auditory component. For this condition the delivery of integrated
explanatory statements—musical notation and auditory excerpts followed each other without a break.

Following the completion of the first phase, students completed a 9-point subjective cognitive load rating scale, which was used as a basis to measure cognitive load and instructional efficiency during the learning phase (as per the explanation for Experiment 1).

A test phase followed the instruction phase. Test materials were identical for all three conditions, containing two sets of written problems (manuscript paper), and two sets of auditory problems (paper and compact disc). A $\frac{3}{4}$ time signature was used for both sets of written problems and the first set of auditory problems. Students were asked to either insert (written problems) or select (aural problems) one (first set) or two (second written set) note value/s to complete a bar of rhythm (Appendix M.). The second set of auditory problems were similarly structured (select one note value), however all examples were based on irregular metres, and as a result required the transfer of understanding from the familiar $\frac{3}{4}$ time signature to new metrical contexts.

The written tests each contained 20 problems (10 per page) and students had 1 min in which to complete as many of these problems as possible. There were more problems in each test set than could be completed in the allocated time. Therefore, as a product of understanding, differences between conditions reflected the capacity of students to find solutions in the shortest possible time. For written problems, one mark was awarded for each note value that correctly completed a bar, giving possible totals of 20 and 40 marks for the first and second sets respectively.

There were two sets of six auditory problems and each problem required the selection of one note value, which in combination with the given rhythm completed the bar according to the given time signature. These auditory problems were analogous to
the written problems, except the manipulated elements were delivered as auditory information. To select their answer, students circled a note value from a series of decreasing note values that spanned a semibreve to a semiquaver (Appendix M). This use of a multiple-choice response was believed to be an efficient way of identifying the correct answer. The length of auditory excerpts dictated the length of the auditory problems, each taking approximately 30 s. Once the time limit for the previous problem had expired, the next problem was presented without further explanation. One mark was awarded for each identified note value that correctly completed a bar of rhythm. A possible total of 6 marks was allocated for each of the auditory test problem sets (12 marks in total).

Each new test problem set was introduced with a brief explanation by the experimenter once the time limit for the previous set had expired. Students did not have access to instructional materials during the test phase. All groups were tested separately and the number of correct solutions was recorded. All questions were objectively marked and no marks were allocated for an incorrect answer.
8.3 Results

An analysis with sets of orthogonal planned contrasts was used in all phases to test predicted differences between instructional formats. It was hypothesized that the delayed auditory condition would demonstrate superior understanding and process a combination of three information sources more efficiently than the same three sources for its immediate auditory counterpart. It was assumed that the immediate auditory format would initially oblige students to engage with all three information sources. The cognitive load of processing all three sources at the same time was expected to overload working memory and interfere with understanding for both written and auditory problems.

A second contrast test compared a combination of the dual-modal conditions (immediate auditory and delayed auditory) with the uni-modal (no auditory) condition. This comparison investigated parallel differences between conditions with either two or three information sources, and therefore differences either with or without auditory excerpts.

It was predicted for written test problems that the conditions with either two or three information sources would perform at an equivalent level. Although dual-modal conditions (immediate auditory and delayed auditory) received additional auditory excerpts, the nature of written problems was in keeping with the understanding that could be extracted by the uni-modal condition (no auditory) from the two written sources alone (text and musical notation). Additionally, these two integrated sources could be processed in a cognitively efficient way. Aural problems however were predicted to favour a combination of the two dual-modal conditions (three sources, including auditory excerpts). In contrast, it was expected that the uni-modal condition
would find it difficult to appropriately manipulate the necessary auditory elements due to the absence of auditory excerpts during instruction.

The variables under analysis were number of correct written and aural test problems completed. Means and standard deviations are displayed in Table 8.1.

Table 8.1 Test Problems for Experiment 4

<table>
<thead>
<tr>
<th>Phase</th>
<th>Max Score</th>
<th>Uni-Modality M</th>
<th>Uni-Modality SD</th>
<th>Dual-Modality Immediate M</th>
<th>Dual-Modality Immediate SD</th>
<th>Dual-Modality Delayed M</th>
<th>Dual-Modality Delayed SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written Problems 1</td>
<td>20</td>
<td>3.00</td>
<td>2.62</td>
<td>2.06</td>
<td>2.04</td>
<td>4.17</td>
<td>2.43</td>
</tr>
<tr>
<td>Written Problems 2</td>
<td>40</td>
<td>2.32</td>
<td>2.14</td>
<td>2.61</td>
<td>2.25</td>
<td>3.06</td>
<td>1.70</td>
</tr>
<tr>
<td>Total Written Problems</td>
<td>60</td>
<td>5.32</td>
<td>4.41</td>
<td>4.67</td>
<td>4.10</td>
<td>7.22</td>
<td>3.96</td>
</tr>
<tr>
<td>Aural Problems 1</td>
<td>6</td>
<td>2.16</td>
<td>1.57</td>
<td>2.67</td>
<td>1.85</td>
<td>3.06</td>
<td>1.80</td>
</tr>
<tr>
<td>Aural Problems 2</td>
<td>6</td>
<td>2.16</td>
<td>1.57</td>
<td>2.50</td>
<td>2.00</td>
<td>3.00</td>
<td>1.72</td>
</tr>
<tr>
<td>Total Aural Problems</td>
<td>12</td>
<td>4.32</td>
<td>2.43</td>
<td>5.17</td>
<td>3.43</td>
<td>6.10</td>
<td>3.23</td>
</tr>
</tbody>
</table>

Note. n = simultaneous uni-modality19, simultaneous dual-modality18, successive dual-modality18.

For the first set of written problems the delayed auditory condition solved a significantly greater number of test problems than the immediate auditory condition, \( t(52) = 2.66, d = 0.94 \). There were no significant differences for the second set of written problems, \( t(52) = 0.65 \), however the difference was in the predicted direction, the delayed auditory condition recording the highest non-significant mean. These data were in accord with the hypothesis that despite processing three information sources, total cognitive load for the delayed auditory condition would remain within working memory capacity.

For the same written problems, no significant differences were found between a
combination of the dual-modal conditions (immediate auditory and delayed auditory) and the uni-modal simultaneous condition (no auditory) for the first set, \( t(52) = 0.16 \), or second set, \( t(52) = 0.89 \), of written problems. This result was consistent with the hypothesis that two- or three- source formats had sufficient information to perform at equivalent levels for written test problems.

For auditory problems there were no significant differences between dual-modal conditions (immediate auditory versus delayed auditory) for either the first, \( t(52) = 0.67 \), or second, \( t(52) = 0.85 \) sets. An observation of means indicated the results were in the predicted direction, although it was expected that the delayed auditory condition would significantly outperform its immediate auditory counterpart. In comparison to written tests, it is conceivable that the small number of aural problems (6 each set) and the use of multiple-choice responses contributed to the narrower range of means between conditions.

In line with the previous contrast, there were no significant differences between a combination of the dual-modal conditions (immediate auditory and delayed auditory) and the uni-modal condition (no auditory) for either the first, \( t(52) = 1.43 \), or second, \( t(52) = 1.18 \), sets of auditory problems. The no auditory condition recorded the lowest non-significant mean for both sets of aural problems, however the data did not indicate the expected differences between conditions.

Means and standard deviations for cognitive load and efficiency measures are displayed in Table 8.2. The efficiency measures combined the standardised subjective cognitive load ratings for the instruction phase with the standardised results for written and aural test problems.
Table 8.2 *Cognitive Load and Efficiency Measures for Experiment 4*

<table>
<thead>
<tr>
<th>Phase</th>
<th>Uni-Modality</th>
<th>Dual-Modality</th>
<th>Dual-Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td>3.26</td>
<td>1.76</td>
<td>3.11</td>
</tr>
<tr>
<td>Efficiency Written Problems</td>
<td>-0.21</td>
<td>1.31</td>
<td>-0.25</td>
</tr>
<tr>
<td>Efficiency Aural Problems</td>
<td>-0.34</td>
<td>1.21</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

*Note. n = simultaneous uni-modality19, simultaneous dual-modality18, successive dual-modality18.*

There were no significant differences between conditions for cognitive load, either between the immediate auditory condition and delayed auditory condition, $t(52) = 1.29$, or between a combination dual-modal conditions (immediate auditory and delayed auditory) and the uni-modal condition (no auditory), $t(52) = 1.10$. The delayed auditory condition recorded the lowest non-significant mean cognitive load.

Standardised cognitive load ratings and performance scores (number of correct solutions) were used to calculate instructional efficiency for written problems. There was a marginally significant difference that favoured the delayed auditory condition over the immediate auditory condition, $t(52) = 1.89$, $d = 0.68$ (one-tailed). This result was consistent with expectations that suggested the delayed auditory condition would benefit from the relative advantages of superior performance and lower levels of cognitive load. As predicted, there was no significant difference for written problems between a combination of the dual-modal conditions (immediate auditory and delayed auditory) and the uni-modal condition (no auditory), $t(52) = 0.99$.

For aural problems, there were no significant differences for efficiency measures between either the dual-modal conditions (immediate auditory versus delayed auditory),
\( t(52) = 1.62 \), or between a combination of the dual-modal conditions and the uni-modal condition, \( t(52) = 1.34 \). However, means were in the predicted direction and when written and auditory problems were combined (total test problems) there was a marginally significant difference that favoured the delayed auditory condition over the immediate auditory condition, \( t(52) = 1.73, \ d = 0.62 \) (one-tailed). Figures 24 and 25 plot, on Cartesian planes, efficiency measures for written problems and total test problems (written + auditory), which reflect in relation to the diagonal axis (hypothetical efficiency baseline) the relative efficiency of the delayed auditory format.

**Experiment 4**

![Efficiency Measures Written Problems](image)

*Figure 24.* Relative condition efficiency for Written Problems, Experiment 4 (see Paas and van Merrienboer (1994b)).
Experiment 4

Figure 25. Relative condition efficiency for Total Test Problems, Experiment 4 (see Paas and van Merriënboer, 1993).
8.4 Discussion

The results for this experiment provide some support for the instructional formats and hypotheses generated by CLT. In particular, these data are consistent with the notion that the number and complexity of interacting elements must be presented in such a way as to minimise the load on working memory. Accordingly, differences between conditions and the order of test problem means were in the predicted direction. Even so, cognitive load ratings suggest that students found the very short instructional sequence easier than expected. Without sufficient cognitive load arising from element interactivity, understanding is likely to occur irrespective of instructional format. Accordingly, only small differences between groups were obtained. Therefore, in future experiments increasing the intrinsic cognitive load associated with instruction could be expected to increase predicted differences between these conditions.

A primary aim of this experiment was to directly compare the two dual-modal formats, both of which were required to process three sources of information. Students studying the delayed auditory format demonstrated superior understanding by solving a significantly higher number of written problems and recording the highest mean score across all test problem sections. When performance and cognitive load ratings were viewed together, this same condition was found to be marginally more instructionally efficient than its immediate auditory counterpart. Similarly, Figures 24 and 25 illustrate that the delayed auditory condition maintained high relative efficiency across all problems.

These data for the two dual-modal formats were consistent with the prediction that processing three interacting sources of information would overload working memory unless presented in a way that minimised cognitive load. To this end, the
delayed auditory format enabled students to establish a schema/s for the written materials (presented first), which was then used to facilitate the integration of all three information sources once the auditory excerpt was presented. Any extraneous cognitive load generated when switching between written and auditory elements was either avoided or minimised by retaining the explanatory text and musical notation while students listened to the auditory excerpts. In addition, musical notation and auditory excerpts provided both mutually reinforcing elements and the perceptual means by which the material could be understood from both visual and auditory perspectives.

In contrast, the immediate auditory format initially delivered all three information sources together, and it is assumed the resultant intrinsic cognitive load exceeded working memory capacity. To achieve understanding, high element interactivity material must be simultaneously processed in working memory (Sweller et al., 1998); however, if task demands (intrinsic load) exceed cognitive capacity, cognitive overload occurs (Mayer & Moreno, 2003). It is assumed that for the immediate auditory condition’s (three information sources) novice learners, the total number of elements to be simultaneously processed exceeded limited mental resources. Moreover, under these conditions cognitive overload interferes with the process of developing the very schemas that (for experienced learners) initially restrict within working memory the number of individual elements required for understanding.

Pollock et al. (2002) referred to this situation as a paradox: For complex learning, establishing schemas may require the mental integration of multiple elements, the number of which in turn interfere with the establishment of these same schemas. Under these conditions they suggest an isolated-interacting elements technique where at first elements are learnt in isolation before, in a second phase, studying the interactions between them. Pollock et al. (2002) stated that by initially avoiding the interactions
between elements the intrinsic load of complex tasks can be artificially reduced. As suggested above, the delayed auditory format similarly reduced the intrinsic load by establishing at least a partial schema/s for the written materials before integrating the auditory excerpt.

For the no auditory condition (two information sources), the implications that arise from this experiment are less conclusive. The data for written problems were in accord with predictions that both dual- and uni- modal formats would facilitate equivalent understanding. It was assumed that a cognitively efficient integrated design of written text and musical notation, with or without auditory excerpts, would provide sufficient information to solve the associated written problems. It should be remembered that similar integrated written materials (uni-modal) used in Experiment 2 established the effectiveness of this format. Despite the proven efficiency of integrated designs, the absence of auditory excerpts during instruction was expected to interfere with understanding when students studying these (no auditory) materials were presented with aural problems. It is feasible that students would experience difficulty in responding to auditory elements if during instruction they were exposed only to written materials. Therefore, in this experiment it was important to establish whether auditory excerpts as a third information source would inhibit or enhance understanding.

The expected differences between formats either with or without auditory excerpts however were not indicated, although means for aural problems were in the predicted direction. Consequently, the relative instructional advantages of integrated written materials either with or without auditory excerpts remain uncertain. Nonetheless, these data do show in relation to the instructionally effective integrated format that delayed auditory presentation facilitated equivalent understanding. Moreover, non-significant differences between means favoured the delayed auditory
condition for all test problem sections and for mean cognitive load ratings. It can be concluded from these data that three information sources do not negatively affect understanding providing the text, musical notation and auditory excerpts are delivered in a way that maintains cognitive load within processing capacity.

Music requires working within visual and auditory environments, neither of which contributes to understanding in isolation. As a result, a key aim of this experiment was to investigate ways in which a unified approach to formatting three essential elements for the instruction of music promotes understanding. The evidence from these data suggests that the modifications applied to the delayed auditory format might provide a practical approach to, and theoretical rationale for, effectively combining three musical sources of information. Furthermore, a format that integrates all three elements essential for musical understanding is desirable for the added efficiency (time saved) that a holistic approach achieves.

In assessing the value of these results, it should also be remembered that all formats were founded upon the theoretically based instructional practices of integration and simultaneity, factors that avoid the negative effects of physical and temporal split-attention. As a consequence, any instructional advantages, especially those indicated for the delayed auditory format in this experiment, are viewed both as reinforcing the CLT principles upon which they were based, and as highlighting broader differences with other less efficient instructional designs.
Chapter 9 Experiments 5 and 6

9.1 Overview

The purpose of Experiments 5 and 6 was to compare dual-modal formats that combined written materials (musical notation and/or explanatory text) with auditory musical excerpts\textsuperscript{183}. To extend the investigation of dual modality, these experiments maintained identical conditions and groups of participants across two independent sets of materials. These materials increased, from Experiments 5 to 6, the level of element interactivity and therefore increased the level of intrinsic cognitive load. By using, in conjunction with explanatory text, both very short (Experiment 5) and extended (Experiment 6) musical examples, these investigations represented two valid ways of presenting dual-modal (multimedia) instructional formats in music.

\textsuperscript{183} From this point the term “auditory excerpt” implies musical excerpt.
Experiment 5

9.2 Introduction

There were four conditions differentiated by musical notation (with and without) and presentation order (simultaneous and successive). All conditions used written explanatory text and auditory excerpts (dual-modal): For two of these conditions musical notation was presented either with (simultaneous) or following (successive) the explanatory text. Therefore in line with Experiment 4 the current study also investigated whether a third information source facilitated or interfered with learning, although in this experiment musical notation rather than auditory excerpts acted as the extra source of information. In addition, the two conditions without musical notation were closely related to the dual-modal conditions of Experiment 3 (text and auditory excerpts), which also tested two-source formats under either simultaneous or successive conditions.

For Experiment 5, successive presentation conformed to a conventional temporal split-attention design, which required students to hold explanatory text in working memory while referring to the associated musical notation and/or auditory excerpts that followed. There was no attempt to modify formats in order to minimise temporal split-attention, and for that reason extraneous cognitive load was expected to interfere with understanding (Chandler & Sweller, 1992). By comparison, information for the simultaneous conditions was spatially (with notation) and temporally integrated, and therefore predicted to facilitate the necessary mental connections required between information sources (Chandler & Sweller, 1991; Mayer & Simms, 1994).

In addition, Experiment 5 explored the extent to which element interactivity informed the instructional design of music materials, especially where these materials were characterised by basic interactions between explanatory text and auditory excerpts
(common to all conditions). Previous studies (Jeung et al., 1997; Tindall-Ford et al., 1997) have established that effects generated by the presentation of instruction can vary according to different levels of element interactivity. For example, Tindall-Ford et al. (1997) found that “when instructions are low in element interactivity, the instructional format makes little appreciable difference to learning” (p. 279). On the other hand, understanding is promoted when the way in which information is presented lowers the extraneous cognitive load of high element interactivity (complex) materials (Sweller et al., 1998).

In order to investigate the differential effects of element interactivity, the materials of Experiment 5, using a large series of individually presented concepts, demonstrated ways of producing musical variety; each concept was associated with a singular musical feature to be memorised. In past studies (e.g., Tindall-Ford et al., 1997), low element interactivity materials have been structured as discrete lists of independent items (elements), where only minimal processing was required to learn each element: For instance, a term and its associated symbol constitute a single element in which only one mental relation is required for learning. Without the need to assimilate and simultaneously process multiple elements, intrinsic cognitive load remains low (Sweller, 1993).

In contrast to these previous studies, the materials of this experiment also included comparative auditory excerpts (as for Experiment 3), which required more than one relation to be made between information sources. Although individual textual, notational and/or auditory elements were very basic, information had to be referred between the elements within these episodes. Consequently, it was expected that processing (understanding) comparative sources of information would generate sufficient element interactivity to favour the instructional efficiency of simultaneous
presentation. In addition, none of the musical concepts was difficult to understand in isolation, although in total, it was difficult to learn a large series of individual concepts within the allocated time.

The inclusion of musical notation, in combination with its auditory musical equivalent and explanatory text (three sources), incorporated the common materials from which music students are often required to learn. Furthermore, it was assumed that for complex materials musical notation provides additional support when students are required to find relations between explanatory text and associated musical sources. Nonetheless, Experiment 5 materials were considered relatively basic and as a result it was expected that they would be learned with or without the addition of this third information source.

Under these circumstances, the inclusion of musical notation might be considered redundant or even questionable. However, testing whether these materials differentiated between conditions on the basis of musical notation was important for two reasons: First, as stated above, it is extremely common to find these materials combined for music instruction, and therefore it is important to establish whether such a practice interferes with learning. Second, it was expected that for these basic materials processing the written-auditory equivalents of musical notation and auditory excerpts would generate only minimal additional cognitive load. A priori, it was assumed that mentally integrating equivalent information sources would consume fewer mental resources than when students were required to integrate mutually referring but different information sources (e.g., text and notation). In short, it was of interest, in relation to basic comparative music materials, whether the addition of musical notation would overload working memory.
9.3 Method

9.3.1 Participants.

Seventy-seven Year 7 students from an academically selective boys’ high school in Sydney participated in this experiment. As with previous experiments, students were selected as musically experienced in relation to their peers, ensuring current levels of understanding were sufficient to complete the problems once additional study was undertaken. All other participant characteristics and procedures for selection were as outlined in Experiment 1.

9.3.2 Materials and procedure.

The materials for Experiment 5 were different to the procedural explanations of time signatures used in Experiments 1, 2 and 4. There were, however, similarities between these materials and those of Experiment 3, both explaining the concept of musical variation. Experiment 5 consisted of two phases: The first phase incorporated fifteen instructional examples, and the second phase incorporated three sets of test problems (twenty, ten and five examples respectively); each phase was presented without a break. The instruction phase was differentiated by four experimental treatments: simultaneous with notation condition; successive with notation condition; simultaneous without notation condition; and the successive without notation condition. Therefore, dual-modal materials varied between two (visual-auditory) and three (visual-visual-auditory) information sources by way of written explanations, and, written musical notation and/or auditory excerpts.

The experiment was administered in four discrete sessions due to the different requirements of delivering auditory and visual materials. The instruction phase was delivered as a Powerpoint presentation for all conditions. Test problems were presented
on A4 sheets of paper and the associated auditory material was played on compact disc. In contrast to previous experiments, all test problems were based upon aural materials. The experimenter developed all instructional examples and test problems, and the materials built upon previous learning undertaken as part of the Year 7 curriculum in music.

For Experiment 5 all conditions demonstrated the way in which fifteen concepts of music provided musical variation. As for Experiment 3, the concept of variation was defined as a new version of an original musical excerpt that preserved an identifiable aural relationship between the original melody and its variation. In contrast to Experiment 3, variation was not restricted to melodic features, but encompassed rhythmic, pitch and structural concepts (Appendix N.). The explanations for each concept appeared on a new series of PowerPoint slides and they were presented without reference to previous concepts.

The explanation of each new concept was accompanied by two short comparative auditory excerpts (1-2 bars in length), which demonstrated how an original excerpt (first slide) was varied (second slide). Each concept was accompanied by a sound representing the beat and a ‘ting’ that indicated the first beat of each bar. The audible beat and metre indications gave students auditory reference points for the variation of rhythmic elements such as metre, and were thought necessary for the experience level of students in this experiment. The musical material differed for each concept, however all examples were simply structured and based upon a C major tonality. The length and complexity of each individual musical excerpt was believed to fall within working memory capacity.

For each instructional concept students were asked to remember the given term or phrase that described the demonstrated variation. Consistent with the basic
instructional materials, the explanatory statements that accompanied each excerpt were very short and typically consisted of only 1-3 words. The first statement established the concept and the second statement in relation to this concept described either specific (Figure 26) or general changes (Figure 27). Each pair of comparative auditory excerpts, including a brief pause in between, lasted approximately 13 s. For simultaneous conditions, explanatory statements could be read within the duration of each auditory excerpt.

*First Powerpoint slide*

8. Constant metre

![Figure 26. An example of the integrated written materials used in Experiment 5 indicating a specific variation.](image)

*Second Powerpoint slide*

Mixed metre

![Figure 27. An example of the integrated written materials used in Experiment 5 indicating a general change.](image)

1. Rhythmic pattern….
Students were required to read the explanatory statements and identify the indicated difference between the comparative auditory excerpts. Because explanatory statements, musical notation and/or auditory excerpts were very short, these differences were easily recognized. Therefore, element interactivity was relatively low insofar as any musical learning activity that requires reading and/or listening to music can be characterised in this way. Even so, an element count confirmed that students would still need to consider a small number of basic interactions for each comparative episode (estimated at five interactions; Appendix O.).

Each of the two parts comprising the comparative materials was, at least on the surface, not dissimilar to previously employed materials described by low element counts (Tindall-Ford et al., 1997). It was of primary interest, however, whether basic comparative episodes in music generated the predicted cognitive load despite the otherwise basic understanding (element interactions) required for learning. Although the relations between elements were minimal, learning difficulty was increased for all conditions by changing the slides in quick succession, each taking only the time necessary to read and hear the explanatory statement and auditory excerpt respectively. Furthermore, students were exposed within a short period of time to fifteen of these concepts. Consequently, concepts were difficult to remember in total, but as outlined above, relatively easy to learn in isolation.

The simultaneous without notation condition had 20 participants and all other conditions 19 participants. Students were allocated to treatment groups according to the process outlined in Experiments 3 and 4. All students were administered audiation tests (Test T – Tonal Imagery, Part 1 Melody; Tests R – Rhythm Imagery, Part 1 Tempo; Part 2 Meter) from the Musical Aptitude Profile (Gordon, 1995), and ranked on their standardised scores. They were then assigned to one of four conditions by progressive
allocation (1234…2341…3412 etc). The fifteen concepts used in this experiment covered a range of audiation skills, which were a factor in a student’s ability to learn the materials described above. Therefore a composite score of audiation tests was used to rank students according to the indicative auditory skills required to complete these materials. As with Experiments 3 and 4, distributing these skills across conditions minimised any interference with the intended experimental manipulation.

For the instruction phase the explanatory statements and auditory excerpts were identical for all conditions and delivered by way of differing instructional formats. The very brief explanatory statements for each concept were read either before (successive), or while (simultaneous) listening to the comparative auditory excerpts. For the two conditions with three information sources the explanatory text was placed either just above (simultaneous)—fully integrated—or preceding (successive) the musical notation. All auditory excerpts were delivered as audio files attached to either the associated explanatory Powerpoint slide (simultaneous conditions) or a blank Powerpoint slide either with or without musical notation (successive conditions). Successive explanatory statements and the musical notation and/or auditory excerpts followed each other without a break.

The successive conditions were not permitted to review the explanatory statements once the musical notation and/or auditory excerpts commenced. As a result, explanatory statements had to be held in working memory until associated written and/or auditory referents were studied (temporal split-attention format). For the simultaneous conditions, the musical notation and/or auditory excerpts commenced when students first viewed the explanatory statements. All conditions were given equal instructional time. Accordingly, simultaneous conditions studied each explanatory statement, musical notation and/or auditory excerpt for the cumulative time taken by
successive conditions to study the same materials.

Following the completion of the first phase, students completed a 9-point subjective cognitive load rating scale, which was used as a basis to measure cognitive load and instructional efficiency during the instruction phase (as per the explanation for Experiment 1).

A test phase followed the instruction phase. Three sets of problems were similarly structured to the earlier instructional episodes, including the use of comparative auditory excerpts. Each set gradually increased test difficulty by requiring students to identify one, two and then three concepts that varied between comparative auditory excerpts. There were 20, 10 and 5 questions in sets 1-3 respectively (Appendix P.), and all earlier instructed concepts were randomly represented at least once. The questions were heard either once or twice for the first two sets and third set respectively. The length of auditory excerpts dictated the length of the auditory problems, each taking, with a short break in between, approximately 15 s to complete.

For all problems, once the time limit for one problem had expired the next problem was presented without further explanation. Students did not have access to instructional materials during the test phase. All groups were tested separately and the timing was dictated by the delivery of auditory materials. For the three sets of problems, one mark was awarded for each correctly identified variation, giving 20 (20 x one correct answer), 20 (10 x two correct answers) and 15 (5 x three correct answers) marks in total respectively. The number of correct solutions was recorded. All questions were objectively marked and no marks were allocated for an incorrect answer.
9.4 Results

A 2 (order: simultaneous or successive) x 2 (notation: with or without) analysis of variance (ANOVA) was calculated to test differences between the four conditions. It was hypothesised that element interactivity associated with integrating explanatory text and auditory excerpts would be sufficient to favour the simultaneous presentation of materials. For successive presentation, the temporal split-attention between mutually referring information sources was expected, despite the processing of relatively basic materials, to raise extraneous cognitive load. Because each comparative instructional episode was very short and contained few elements, it was also predicted that students would understand these materials with or without the addition of musical notation.

The variables under analysis were the number of correct similar test problems completed. Means and standard deviations are displayed in Table 9.1.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Max Score</th>
<th>Condition</th>
<th>Simultaneous</th>
<th>Successive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>+ Notation M</td>
<td>+ Notation M</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Experiment 5 Problems</td>
<td>55</td>
<td></td>
<td>11.18</td>
<td>5.74</td>
</tr>
</tbody>
</table>

Note. n = simultaneous + notation 19; simultaneous - notation 20, successive + notation 19; successive - notation 19.

For total problems (sets 1-3) there was a significant main effect for presentation order that favoured simultaneous formats, $F(3, 73) = 6.31$, $MS_e = 153.91$, $\eta^2_p = 0.08$ (partial Eta squared is calculated as $\eta^2_p = \frac{SS_{effect}}{SS_{effect} + SS_{error}}$). This difference was in accord with the hypothesis that simultaneous presentation facilitates the integration of instructional elements.
There was no significant main effect for musical notation $F(3, 73) = 0.58$, $MS_e = 14.05$, nor any significant interaction, $F(3, 73) = 1.81$, $MS_e = 44.09$. These results were consistent with the prediction that students would understand these basic concepts with or without musical notation. Nonetheless, Table 9.1 indicates that the two conditions with additional musical notation (simultaneous and successive) solved the highest and lowest mean number of problems respectively. This difference provides some evidence that a third information source might have increased the load on working memory when materials were presented in a successive or split-attention format.

Means and standard deviations for cognitive load and efficiency measures are displayed in Table 9.2. The efficiency measure combined the standardised subjective cognitive load rating for the instruction phase with the standardised results for total test problems.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Simultaneous</th>
<th>Successive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Notation</td>
<td>- Notation</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>4.74</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>5.68</td>
<td>1.97</td>
</tr>
<tr>
<td>Exp.5</td>
<td>Efficiency</td>
<td></td>
</tr>
<tr>
<td>Total Problems</td>
<td>0.43</td>
<td>1.23</td>
</tr>
</tbody>
</table>

Note. $n =$ simultaneous + notation 19; simultaneous - notation 20, successive + notation 19; successive - notation 19.

For cognitive load ratings, there were no significant main effects for either presentation order, $F(3, 73) = 0.59$, $MS_e = 2.05$, or musical notation, $F(3, 73) = 0.10$, $MS_e = 0.03$. Likewise, there was no significant interaction between presentation order and musical notation, $F(3, 73) = 2.15$, $MS_e = 7.42$. In keeping with the data for test problems, both the successive and simultaneous conditions with added musical notation
recorded the highest and lowest non-significant mean cognitive loads respectively.

Standardised cognitive load ratings and performance scores were used to calculate instructional efficiency for total similar problems. There was a significant main effect for presentation order, $F(3, 73) = 3.81$, $MS_e = 5.10$, $p = .05$, $\eta^2_p = 0.05$, that favoured simultaneous presentation, and supported the previous differences and observations for test problem and cognitive load data. There was no significant main effect for musical notation, $F(3, 73) = 0.15$, $MS_e = 0.20$, nor any significant interaction, $F(3, 73) = 2.86$, $MS_e = 3.82$; however the consistent rank order difference between the conditions with added musical notation (Table 9.2) was maintained for this measure.

Figure 28 highlights on a Cartesian plane the relative inefficiency of successively presenting three information sources (with musical notation), especially in relation to its simultaneous counterpart. These data confirm that successive presentation negatively affected understanding, however it was unexpected that these basic comparative materials would be sensitive to the addition of musical notation when presented in a successive or split-attention format (Figure 29).
Experiment 5

Figure 28. Experiment 5 efficiency measures for total problems (see Paas and van Merriënboer, 1993).

Figure 29. Experiment 5 efficiency measures, highlighting the difference between conditions with musical notation.
9.5 Discussion

Cognitive load theory was used to generate the instructional formats under investigation in this experiment, which in turn produced results in accord with theoretical predictions. Primarily, this study investigated the physical and temporal integration of information across dual-modal materials that required the manipulation of basic elements within a series of independent learning episodes. Integral to these investigations was whether the addition of musical notation, as a third information source, interfered with understanding.

The major findings for Experiment 5 support the presentation of information in ways that avoid the extraneous cognitive load associated with holding and searching for mutually referring elements in working memory. Accordingly, test problem data confirm the superiority of simultaneous presentation in preference to the same information delivered in successive or split-attention formats. Moreover, this superiority of simultaneous presentation was maintained across either two or three information sources (with or without musical notation).

Despite the basic nature and relatively lower demands (element interactivity) of these materials, learning remained sensitive to the processing advantages of simultaneous (integrated) presentation. Accordingly, Experiment 5 data provide evidence that even basic comparative episodes increase the number of relations between elements such that cognitive load remains an important consideration of instructional design. When students possess lower levels of domain knowledge, the assimilation of even relatively few elements expends otherwise limited mental resources in the task of constructing the necessary domain schemas for learning. Under these circumstances, simultaneous conditions in this experiment were better able to utilise working memory
resources for this task.

The results for musical notation were consistent with predictions that each individual concept was intelligible with or without musical notation, and that any additional processing would remain within working memory capacity. As expected, the element interactions between text and musical sources were cognitively more demanding than those involving the visual-auditory equivalents of musical notation and auditory excerpts. Hence, evidence for differential performance arose for presentation order, but not for musical notation.
Experiment 6

9.6 Introduction

The purpose of Experiment 6 was to further investigate the dual-modal formats of the previous study (Experiment 5), where understanding required the mental integration of explanatory text, musical notation and/or auditory excerpts. For music students it is not unusual to learn by way of the simply structured episodes studied in Experiment 5, in which each brief episode was intelligible without reference to any other. Nevertheless, music materials are often presented in much longer sequential episodes than those used in Experiment 5; longer episodes provide more authentic, albeit more complex, learning contexts. Accordingly, the complexity of Experiment 6 materials was substantially increased to examine whether the processing of either two or three interrelating information sources would produce larger effects than those found for the previous experiment.

The four conditions were identically structured to those of Experiment 5, each format differentiated by presentation order (simultaneous and successive) and musical notation (with and without). Element interactivity was considerably higher in relation to Experiment 5 materials, and as a consequence the load on working memory was similarly increased (see comparative element counts, Appendices O and Q). Consistent with earlier predictions, integrated and simultaneous formats were expected to lower extraneous cognitive load and facilitate understanding in comparison to conventional split-attention designs.

To achieve greater complexity, the materials of Experiment 6 demonstrated

184 Although it is unlikely that in a classroom setting that so many episodes would be presented in one session.
ways in which a small series of concepts affected musical texture. In stark contrast to Experiment 5, the explanation and demonstration of each concept was longer and required the manipulation of multiple interdependent elements. Moreover, auditory excerpts incorporated two- rather than single- instrument examples to demonstrate musical concepts, and each of these were longer than the auditory excerpts previously employed. As a result, students had to reconcile explanatory statements with two different but interrelated auditory events. Additionally, students were required to refer information between comparative episodes, and between the elements within those episodes. Taken together, element interactivity was expected to be appreciably higher for these materials.

For Experiment 6, the complex nature of music materials raised a number of challenges in relation to the design and expectations of instructional formats. The results of Experiment 4 suggested that simultaneous delivery of three information sources might interfere with a novice student’s capacity to mentally integrate instructional elements. Moreover, the unitary block of explanatory text used in this experiment was relatively large and, for the simultaneous condition, directly adjacent to the musical notation; it was not, however, integrated as a series of sequential statements (see Experiment 4).

Despite these possible limitations, musical notation was considered essential for understanding complex materials, although only when presented as part of an instructionally efficient format. Musical notation was assumed to reduce search, and in turn free scarce working memory resources in order to mentally integrate the increased number of elements associated with these materials. For these reasons, the simultaneous with notation condition was expected to facilitate the connections between textual and auditory musical referents. In contrast, the temporal and physical split attention
associated with successively presenting three information sources was predicted to increase search and significantly raise cognitive load.
9.7 Method

9.7.1 Participants.

Participant details were identical to those reported for Experiment 5, students remaining in the same conditions from one experiment to the next.

9.7.2 Materials and procedure.

The materials for Experiment 6 were unrelated to the previous experiments and for the first time introduced musical examples based upon two instrumental parts; the contrasting tone colours of oboe and trumpet (both synthesized) were used to represent these parts. Although the materials for Experiments 5 and 6 were distinct, both served a coherent exploration of dual-modal music instruction. In keeping with this overall design, Experiment 6 also consisted of two phases: The first phase included four instructional examples, and the second phase included twelve test problems; each phase was presented without a break. Likewise, details relating to experimental treatment groups, the administration, development and delivery of instructional and test problem materials were as outlined for the previous experiment.

For Experiment 6 all conditions demonstrated ways in which the manipulation of the given musical concepts affected musical texture. Texture was defined as the way in which musical concepts change the auditory relationship between foreground and background instrumental parts. Foreground described the factors that made one or both instruments obvious to the listener. Accordingly, the application of one or more of these factors placed either one or both instruments in the musical foreground. The four musical concepts used to manipulate texture were rhythmic similarity, rhythmic activity, volume, and pitch. Consistent with Experiment 5, each new concept was presented both on a new series of Powerpoint slides and without reference to previous concepts.
Comparative auditory excerpts once more accompanied each concept and demonstrated how the texture of the first excerpt (first slide) was varied (second slide). However, in contrast to Experiment 5, the length and complexity of each excerpt was increased to three bars and two instrument parts respectively (Figure 30). Because of these factors, it was assumed that in relation to Experiment 5, cognitive load would be considerably higher when processing these materials (estimated at seven interactions; Appendix Q.). The musical examples differed for each concept and they were based upon a C major tonality. Audible beat and metre indications were deemed unnecessary for students to identify the manipulated concepts and they were therefore excluded.

For each textural concept students were asked to remember the type of change that described the demonstrated variation. The explanatory statements that accompanied each excerpt were much longer than those for Experiment 5 and ranged from 26-30 words. The first statement described in relation to the concept how both instruments could be either identical or similar, resulting in both being perceived in the textural foreground. The second statement described how the concept could differentiate between the two instruments and therefore promote only one instrument into the textual foreground (Figure 30, second slide). Each pair of comparative auditory excerpts, including a brief pause in between, lasted approximately 50 s. For simultaneous conditions, explanatory statements could be read within the performance duration of the comparative auditory excerpts.
2. The rhythmic activity or number of notes in each instrument may be very similar or almost identical. Similar numbers of notes place both instruments in the foreground as important parts.

In contrast, the activity or number of notes in each instrument may be mostly or completely different. The more active instrument places it in the foreground as the more important part.

Figure 30. Concept 2 of the integrated written materials used in Experiment 6.

In line with previous comparative auditory excerpts, students were required to identify the indicated difference between excerpts after reading the explanatory statements. Unlike the materials of Experiment 5, there were fewer concepts, but each was more complex to process in working memory. Accordingly, the difficulty of these materials was not in learning a large number of concepts in quick succession, but in manipulating the inherent number of interactive elements between a range of factors: two instrumental parts that comprised the auditory excerpts; longer explanatory statements and associated two-part auditory excerpts; four interrelated concepts that in either isolation or combination affected texture.
The treatment conditions were identical in number and student composition to those outlined for Experiment 5. Therefore, the earlier described procedures for assigning participants—including audiation tests—were applicable to both experiments. For the instruction phase, the explanatory statements and auditory excerpts were identical for all conditions and delivered by way of differing instructional formats. Following the completion of the first phase, students completed a 9-point subjective cognitive load rating scale, which was used as a basis to measure cognitive load and instructional efficiency during the instruction phase (as per the explanation for Experiment 1).

A test phase followed the instruction phase. For each question, test problems required students to answer two parts. The options One / Both were given in order to identify whether one or two instrument parts were promoted into the textural foreground. To complete the question, students had to also identify one concept (from the earlier instruction) that affected the textural relationship between the two instrument parts (Appendix R.). There were 12 questions in total, and all instructional elements were randomly represented at least once. Each problem was heard once and the length of comparative auditory excerpts dictated the length of the auditory problems, each playing taking, with a short break in between, approximately 14 s to complete.

Once the time limit for one problem had expired the next problem was presented without further explanation. Students did not have access to instructional materials during the test phase. All groups were tested separately and the timing was dictated by the delivery of auditory materials. For these problems, one mark was awarded for each correctly circled option and each correctly identified concept, giving 24 (12x2) marks in total. The number of correct solutions was recorded. All questions were objectively marked and no marks were allocated for an incorrect answer.
9.8 Results

A 2 (order: simultaneous or successive) x 2 (notation: with or without) analysis of variance (ANOVA) was calculated to test differences between the four conditions. The high element interactivity of instruction was expected to raise levels of intrinsic cognitive load, and therefore the incorporation of simultaneous presentation and musical notation were considered necessary to facilitate the processing of these complex materials. Conversely, successive presentation (split attention) of either two or three information sources was predicted to substantially raise extraneous cognitive load and, for already complex materials, exceed the capacity of working memory.

The variables under analysis were the number of correct test problems completed. Means and standard deviations are displayed in Table 9.3.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Max Score</th>
<th>Simultaneous</th>
<th></th>
<th></th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ Notation</td>
<td>- Notation</td>
<td></td>
<td>+ Notation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Experiment 6</td>
<td>24</td>
<td>13.71</td>
<td>3.44</td>
<td>15.43</td>
<td>4.75</td>
</tr>
<tr>
<td>Problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.58</td>
<td>2.76</td>
<td>11.74</td>
<td>4.51</td>
<td></td>
</tr>
</tbody>
</table>

*Note. n = simultaneous + notation 19; simultaneous - notation 20; successive + notation 19; successive - notation 19.*

For test problems there was a significant main effect for presentation order that favoured simultaneous presentation, $F(3, 73) = 14.27, MS_e = 223.71, \eta_p^2 = 0.16$, which was in accord with predictions and the results for Experiment 5. Strong evidence from both Experiments 5 and 6 supported the complementary hypotheses that simultaneous and successive presentations lower and raise extraneous cognitive load respectively, irrespective of whether students process either two or three sources of information.
Contrary to the prediction that musical notation would be essential to understand these complex instructional materials, there was no significant main effect for this third information source, $F(3, 73) = 2.53$, $MS_e = 39.69$. The successive with notation condition again recorded the lowest non-significant mean number of problems solved, although differences between this and other conditions were not quite as marked for this experiment. The interaction between presentation order and musical notation was not significant, $F(3, 73) = 0.10$, $MS_e = 1.49$, indicating that the superiority of simultaneous conditions was maintained across either two- or three- source formats.

Means and standard deviations for cognitive load and efficiency measures are displayed in Table 9.4. As reported for Experiment 5, the efficiency measures combined the standardised subjective cognitive load ratings for the instruction phase with the standardised results for total test problems.

<table>
<thead>
<tr>
<th>Table 9.4 Cognitive Load and Efficiency Measures for Experiment 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condition</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>M</td>
</tr>
<tr>
<td>Experiment 6 Cognitive Load</td>
</tr>
<tr>
<td>Exp. 6 Efficiency Problems</td>
</tr>
</tbody>
</table>

Note: $n = $ simultaneous + notation 19; simultaneous - notation 20; successive + notation 19; successive - notation 19.

Cognitive load ratings indicated there were no significant main effects for either presentation order, $F(3, 73) = 0.42$, $MS_e = 1.17$, or musical notation, $F(3, 73) = 0.77$, $MS_e = 2.13$. There was however, a significant interaction, $F(3, 73) = 4.70$, $MS_e = 13.10$, $\eta^2 = 0.06$. A simple effects test indicated that the successive with notation condition experienced higher levels of cognitive load than its without notation counterpart, $t(36) = 2.36$, $d = 0.77$ (Figure 31). This result was consistent with the prediction that for
complex materials splitting attention between three information sources (added musical notation) raises cognitive load and as a result reduces a student’s capacity to manipulate the relevant elements in working memory.

![Experiment 6](image)

**Figure 31.** Significant interaction for Experiment 6 mean cognitive load ratings.

Standardised cognitive load ratings and performance scores were used to calculate instructional efficiency. There were no significant main effects for either presentation order, $F(3, 73) = 3.17, MS_e = 4.02$, musical notation, $F(3, 73) = 2.12, MS_e = 2.69$, nor any significant interaction, $F(3, 73) = 1.34, MS_e = 1.70$. Figures 32 and 33 illustrate relative condition efficiency and the difference between mean efficiency respectively. Both graphs point to relatively poor efficiency where attention is split between three sources of information. Similar to test problems, both the successive and simultaneous conditions with added musical notation recorded the highest and lowest non-significant means respectively.
Figure 32. Relative condition efficiency for Experiment 6 (see Paas and van Merriënboer, 1993).

Figure 33. Experiment 6 mean instructional efficiency, highlighting the difference between successive conditions with and without notation.
9.9 Discussion

The results of Experiment 6 were in accord with those for the previous experiment, extending the findings to dual-modal materials of high element interactivity. Specifically, Experiment 6 reinforced the necessity of presenting mutually referring elements simultaneously in order to avoid the temporal split attention of successive formats. Moreover, the predicted difference between simultaneous and successive conditions (for test problems) increased from Experiments 5 to 6 ($\eta_p^2 = .08$ and .16 respectively), as the differential nature of materials likewise raised levels of instructional complexity. It is therefore assumed that processing simultaneously delivered materials better utilises available cognitive resources, resources that are rapidly exhausted by the costly search and extraneous cognitive load of split-source (successive) instructional designs.

Furthermore, there was evidence from this experiment that the addition of musical notation (as a third information source) places an even greater load on working memory when materials are presented in successive formats. For these complex materials, a significant interaction indicated that cognitive load was higher when successive formats included this third information source. Clearly, from the evidence provided by both Experiments 5 and 6, the addition of musical notation, irrespective of any potential processing advantages, does not ameliorate the negative effects of split attention.

It is uncertain as to why musical notation did not facilitate understanding of these complex materials, especially when simultaneously presented with text and auditory excerpts. In relation to these data, two possible explanations are considered: The musical notation was placed adjacent to blocks of explanatory text, the extent of
which, despite simultaneous delivery, conceivably required some search to identify mutually referring elements. Problem search either raises cognitive load (Sweller, 1988), or presumably disposes students to avoid the incorporation of these sources, a result that may explain why conditions with musical notation performed no better than those without. Presumably, an equivalent source such as musical notation is easily avoided if not physically structured in ways that compel its mental integration.

An alternative interpretation suggests that as complexity increases, so might the need for prior knowledge of visual and/or auditory musical excerpts. Whether excerpts are familiar or not, understanding can proceed only when all necessary elements can be mentally integrated in working memory. For this to occur, the appropriate number and size of elements is dependent on the nature of students’ current schemas and, as complexity increases, their degree of automation. Given these constraints, it is unlikely novice students possessed schemas for the complex and unknown musical excerpts of Experiment 6. Under these circumstances it is plausible that musical notation, in combination with other elements, was difficult to access and therefore to process in working memory; as a result it was of little instructional benefit.

Assuming that the current formats did not promote complete integration of three information sources (especially notation), two modifications are suggested for future presentation of these complex materials: First, restructuring the text into a series of sequential statements that are fully integrated–proximal to specific notational referents–may facilitate the mental correspondence of elements achieved in previous experiments (e.g., 2 and 4). Second, initially isolating elements for novice learners would establish the necessary schemas that enable later integration and understanding of complex materials (see isolated-interacting elements effect, Experiment 4, Pollock et al., 2002). In both cases, the primary goal of instruction is to assist in the construction of
schemas by maintaining processing loads within working memory capacity.

In summary, the results of both experiments add to a series of cognitive load and multimedia studies across a wide range of materials that have demonstrated the negative effects of physical and/or temporal split attention (Jeung et al., 1997; Mayer & Moreno, 1998, 2003; Mousavi et al., 1995), both of which were evident for the dual-modal materials of Experiments 5 and 6. In addition, there is evidence from these experiments that the extraneous cognitive load generated by successive presentation is further exacerbated when information is split between three rather than two information sources. It is assumed that in all forms, split-attention decreases the working memory capacity dedicated to processing instructional elements essential for the development of schemas.
Chapter 10: General Discussion

10.1 Overview of Research

The experiments reported in this thesis explored the application and extension of cognitive load theory (CLT) to instructional contexts of music. Accordingly, these experiments support both practical suggestions for the effective formatting of music instruction, and the generalizability of CLT to a continually widening range of materials and learning environments. Of primary importance, all six experiments lend support to the fundamental claim of CLT that working memory limitations are a central determinant of instructional design. It was proposed that formatting music materials without also considering these limitations would result in similar learning impairments to those found across other technical (e.g., Chandler & Sweller, 1991) and humanities areas (e.g., Yeung, 1999).

Materials for the instruction of music can assume many forms, which usually include various combinations of explanatory text, musical notation and auditory musical excerpts. Some of these formats rely exclusively on visual materials such as written text and musical notation\(^\text{185}\), or employ a more conventional dual-modal delivery of auditory text and musical notation. The broader application of visual and auditory modalities in these ways is of course familiar to both other disciplines and to previous CLT studies. Not surprisingly, auditory musical excerpts are frequently incorporated into music instruction, and as a result generate either an exclusively auditory format (spoken text and music), or a dual-modal format in which the explanations are delivered as visual text. Moreover, combining all three instructional forms (text, musical notation and

\(^{185}\) Given the perceptual nature of music, the exclusive use of visual materials is usually a temporary or intermediate step towards the inclusion of auditory materials.
auditory excerpts) produces a dual-modal format in which three rather than two sources of information need to be managed in a cognitively efficient way. Few if any previous studies have applied the principles of CLT to the exploration of these reconfigured dual-modal formats.

The nature of instructional formats and the modalities they employ have a major bearing on the efficiency of learning. According to CLT, working memory is comprised of at least partially separate and independent processors for visual and auditory materials. As a consequence, it has been shown that using two modes instead of one (visual) can expand processing capacity and facilitate the mental integration of otherwise disparate sources of information (see section 3.3.5). Without this mental integration of instructional elements, the acquisition of schemas, upon which all understanding depends, cannot occur. By extension, acquiring schemas is dependent upon the capacity demands or mental load of integrating instructional elements. This mental load is determined by the element interactivity that arises from both the intrinsic nature of materials (unalterable) and the extraneous processing generated by the way in which materials are formatted. Hence, formats that reduce or eliminate extraneous processing facilitate understanding by decreasing the load on the limited mental capacity available to accommodate the intrinsic difficulty of materials.

These CLT principles are assumed to be fundamental to the proposal that in order to facilitate understanding, cognitively efficient delivery of two- and three-source dual-modal music materials is essential. For this reason, it was hypothesised that physically integrated and simultaneously presented information sources (explanatory text, music notation and auditory excerpts) would lower extraneous cognitive load and increase available working memory capacity when processing dual-modal music materials. Notwithstanding the differences between the dual-modal formats in these
experiments (visual text) and the previously tested formats of auditory text and diagrams, it was nevertheless assumed that dual-modal processing would similarly increase working memory capacity in relation to uni-modal (auditory-auditory) presentation. Whether, in relation to dual-modal materials, the partial independence of auditory processors (music and text) offered advantages unrelated to previously employed uni-modal (visual-visual) formats was uncertain.

These theoretical principles were tested using a range of music materials that varied in their degree of element interactivity and to the extent to which they represented either defined or ill-defined musical problems. From a theoretical perspective, the design of all materials was guided by reference to and an understanding of cognitive processes, an essential component of any approach to instruction (Chandler, 2004).
10.2 Results

Experiment 1 provided some support for the benefits of physically integrating two mutually referring visual information sources (text and musical notation). Under these conditions, students were better able to understand a combination of explanatory text and its associated musical notation than when the same materials were presented as two discrete and adjacent sources. Although the advantages of integrated presentation were only evident for acquisition problems (learning phase) and not the test phase, it was assumed that these findings were a demonstration of the split-attention effect. On this basis, processing capacity for the split-source group was devoted to the extraneous search associated with mentally integrating both sources of information. Despite the obvious indications across experimental phases that students readily understood these materials\textsuperscript{186}, they remained sensitive during the acquisition phase to the adverse effects of split-source formatting.

Identical materials were used in Experiment 2, however the general level of difficulty was increased; a conventional dual-modal condition (auditory text and visual musical notation) was also included as a third experimental treatment. Students from both the integrated (visual text and musical notation) and dual-modal conditions demonstrated consistently superior understanding in both the acquisition and test phases when compared to the students who were required to study the same materials structured in a split-source format (visual text and musical notation). These findings replicated and considerably strengthened the evidence for a split-attention effect identified in Experiment 1. Moreover, it was assumed that the advantages associated with assigning information between visual and auditory modes were due to the modality

\textsuperscript{186} These indicators included cognitive load ratings and test scores.
effect, resulting in an effective increase of working memory capacity. These results provide a clear demonstration that CLT effects are readily transferable to the materials of music.

Experiment 3 also provided some evidence for a modality effect, but only when materials were presented simultaneously: The instructional efficiency measures, comprising relatively lower cognitive load and relatively higher test performance, favoured the dual-modal condition. In combination with evidence already reviewed (see section 3.3.5), these initial results suggest that there may be dual-modal benefits whether explanatory text is delivered in either the auditory (conventional) or visual modes. However, in contrast to previous studies, dual-modal effects did not extend to successive presentation. Under these conditions, uni-modal (auditory) presentation resulted in superior test performance, although in relation to instructional efficiency, this was somewhat offset by significantly higher levels of cognitive load.

When considering both performance and mental load factors, these data suggest that simultaneously presenting music and text between two modes may provide the most effective way to optimise processing capacity and therefore learning. As summarised for Experiment 2, these dual-modal benefits support the notion of an expanded working memory, however considerable research is needed if, for this configuration of dual-modal materials, the basis of these benefits is fully understood. Recall that although in Experiment 3 students in the simultaneous dual-modal condition recorded superior efficiency measures in comparison to their uni-modal counterparts, this advantage did not extend to superior test performance. Whether this points to the current materials
exerting insufficient levels of difficulty\textsuperscript{187}, or to other factors related to auditory only processing (see section 10.3), is yet to be fully determined.

Experiment 4 tested three instructional formats hypothesized to be efficient: two simultaneously presented dual-modal conditions (three information sources) and a third visually integrated uni-modal condition (two information sources). Given the use of integrated and simultaneously delivered instructional elements, any identified advantages between these conditions were assumed to be associated with a reduction in the intrinsic cognitive load of working memory rather than the extraneous cognitive load generated by instructional design. To this end, presenting information in a two-stage process—integrated text-notation followed by the addition of auditory excerpts—was superior to delivering all three sources at the same time (immediate auditory condition). That is, all components were simultaneously presented as part of an additive and overlapping design (delayed auditory condition), presumably reducing the extraneous cognitive load normally associated with split-source formats. Furthermore, there were no significant differences between the students studying the delayed auditory condition and the students studying the instructionally efficient two-source integrated format (no auditory condition). It is therefore possible that for the delayed auditory condition the additional processing of auditory excerpts remained within the capacity of working memory.

Experiment 4 demonstrated, by way of instructional design, the importance of delivering materials in such a way that levels of intrinsic cognitive load remain within a learner’s working memory capacity. This result is of some importance to a discipline in which three information sources containing streamed (real time) material are essential to

\textsuperscript{187} That is, despite the differences in cognitive load, the auditory conditions may have had sufficient available mental capacity to understand these materials.
learning. According to CLT, excessive levels of intrinsic cognitive load overwhelm working memory capacity and as a result interfere with understanding. Accordingly, this two-stage format was assumed to effectively spread the mental load between two instructionally efficient phases of presentation. In this way the necessary schemas for three integrated sources (dual-modal) were, at least in part, constructed progressively, but without the burden of holding referents from one stage (text and notation) to the next (text, notation and auditory music).

Experiment 5, under conditions of relatively low element interactivity, extended the previous findings by comparing four dual-modal formats: visual text and auditory music with or without musical notation, each format either simultaneously or successively delivered. The results strongly support the presentation of either three or two information sources in integrated and/or simultaneous formats respectively. Conversely, formats that physically and/or temporally split-attention between visual text and musical notation and/or auditory excerpts were assumed to raise extraneous cognitive load and interfere with understanding. Furthermore, when compared with two-source formats, the addition of musical notation (third source) did not result in significant performance and cognitive load differences; however the processing of these three sources was significantly more efficient when presented simultaneously. Notably, for both Experiments 4 and 5, processing efficiency was subject to the way in which instructional formatting accounted for intrinsic and extraneous cognitive loads respectively.

Experiment 6, incorporating the same dual-modal conditions, extended the findings of Experiment 5 to materials of high element interactivity. Moreover, as the complexity of materials was increased from Experiments 5 to 6, the impairment to learning caused by physical and/or temporal split-attention also increased. Contrary to
predictions, the addition of musical notation did not enhance understanding, although a marked difference (interaction) between levels of cognitive load indicated that simultaneously processing three sources of information was less demanding than the same information delivered in a successive format. As would be expected, increasing element interactivity would likewise increase the importance for instructional design of extraneous cognitive load associated with a split-source design. These data are entirely consistent with Experiment 5, and together indicate that for these instructional materials the more crucial concern is the cognitive load generated by the way in which information is presented rather than the number of information sources per se.

Split-attention effects were evident in all but one of the six experiments\(^{188}\), a consistent and significant thread running through this body of empirical work. With respect to these effects, all findings were in accord with the hypotheses generated by CLT for the three phases outlined earlier (Chapter 4). Of note, the negative effects of split-attention were readily apparent irrespective of modality and whether two or more information sources were physically and/or temporally separated. What is more, these effects persisted across a range of musical formats, including visual text with either or both visual notation and auditory music, and were also sensitive to varying levels of element interactivity.

To my knowledge the music materials for these experiments explored the split-attention effect using novel formats that manipulated in new ways the application of both modality and the number of information sources. These instructional formats, which are intrinsic to music, applied the theoretical conception of extraneous cognitive load to a complex of music stimuli separated both in space and over time. In

\(^{188}\) Split-attention was not relevant to Experiment 4; all conditions were presented in cognitively efficient formats.
combination with previous results the unavoidable conclusion from this current series of experiments is clear: Split-attention is a pervasive effect not only applicable across many other instructional contexts, but a significant factor in the instruction of music. Alternating between musical information sources may at times appear intuitively sound, however such a practice may impede the cross-source integration of mutually referring elements.

Although the dual-modality effect was only directly tested in two experiments (2 and 3), all but one of these studies employed dual-modal formats, a feature entirely in keeping with the nature of music materials. Two broad findings are reported and discussed below: First, under certain conditions dual-modal music instruction is associated with working memory advantages when compared to the same materials delivered in a uni-modal format. Although further work is needed, these cognitive advantages may hold whether explanatory material is delivered as either auditory or visual text. Second, the appropriate sequencing and presentation of dual-modal music materials is crucial to the instructional efficacy of these formats. Dual-modal advantages are quickly eroded by other instructional factors that increase intrinsic and/or extraneous cognitive loads.
10.3 Theoretical and Future Research Implications and Limitations of Findings

The findings summarised for split-attention, modality and the efficient management of cognitive load are fully consistent with CLT, which in turn gave rise to the hypotheses and results reported above. Collectively, they add to a series of reported effects that have been replicated across a diversity of learning materials (see sections 3.3.2, 3.3.5 and 3.3.4 respectively). Nonetheless, to my knowledge the six experiments reported above are the first to extend instructional effects to the materials of music, materials which by nature are removed from the more traditional technical areas that initially dominated CLT research.

Although these studies were based on sound theoretical principles, it was uncertain as to whether the hypothesised effects would apply to the novel ways in which modality and the idiosyncratic sources of musical information are combined. Clearly, generalizability is a primary test for the utility of any instructional theory: replicating results for increasingly varied materials strengthens the reliability and validity of a theory’s underlying principles. For this reason, the current studies can be taken as further evidence for the theoretical constructs that generated the instructional designs.

The estimation and management of cognitive load factors formed the basis of all six experiments and it is from these perspectives that certain theoretical observations are made and recommendations for future research suggested. These observations also focus on issues that move beyond the theoretical implications of phenomena such as split attention and dual modality per se: For these effects, the results reported above strongly reinforce the inferences that have already been drawn from previous studies in relation to the mechanisms of human cognition (see Chapter 3). Of interest though, is the extent to which these results inform future instructional designs in music, and more broadly, the extent to which they may further our understanding of the cognitive
constructs that support these instructional designs. To this end, four broad issues are identified in the following discussion.

The first issue of interest concerns the limitations of dual modality as a cognitive mechanism (see Ginns, 2005, for a recent meta-analysis of 43 independent modality effects). Despite the potential dual-modal processing advantages offered by an expanded working memory, holding musical sequences in one mode while searching for their associated referents in another proved problematic across multiple experiments. Indeed, the temporal split attention that is commonly associated with the study of music materials may not simply be a key factor, but the key factor to consider when designing instructional formats for this discipline: such was the obvious impairment across these studies. Holding sufficient real-time auditory material as a means of facilitating cue abstraction (elements), let alone to facilitate the integration of individual elements, may quickly exhaust working memory resources, including any increase in capacity due to dual-modal processing. On the other hand, the temporal proximity of simultaneous presentation appears to facilitate the mental correspondence of musical elements, which, given the consistent instructional advantages, is assumed to appreciably reduce the operational loads on working memory.

Compared to equivalent uni-modal visual materials, previous research has demonstrated that dual-modal formats result in superior understanding, even where both are delivered as split-source formats (Mousavi et al., 1995, Experiments 3 and 4). Therefore, under certain conditions, dual modality may at least partially compensate for the concomitant extraneous cognitive load of split attention. However, temporal split attention associated with dual-modal multimedia materials was also found to impair understanding (see Mayer & Moreno, 2003, "Synchronizing", p. 50) when compared with contiguously presented alternatives. Together, these findings strongly recommend
that the working memory advantage associated with dual modality is not by itself an alternative to inefficient instructional design, but rather viewed as a compatible strategy that can further enhance the cognitive efficiencies of simultaneous presentation\textsuperscript{189}.

This proposition is fully supported by the findings of the experiments reported above, although it may contradict intuitive or traditional practices in music: Listening to music, explaining concepts, or pointing out score references, are often undertaken as ostensibly discrete activities, cross referencing where necessary to integrate relevant sources of information. Regardless of how either attractive or instructionally neat this might first appear, the current findings indicate that understanding will be superior when mutually referring information sources are mentally integrated as simultaneously presented materials. In order to control the total cognitive load imposed by the simultaneous processing of multiple elements, the sequencing, length and complexity of learning episodes are assumed to act as mediating factors. The manipulations of these intervening factors suggest useful areas for continuing investigation.

Of relevance to these issues is the recent work of Tabbers, et al. (2004) in which they found reverse modality effects when comparing visual-visual and audio-visual formats, results that appear to differ significantly from earlier research. They sought to test the modality effect within a classroom setting under self- rather than system- paced instructions\textsuperscript{190}. Under these conditions, where students have more time to refer elements between information sources, Tabbers et al. suggested that the usual advantages associated with bimodal simultaneity disappear; skipping between visual sources

\textsuperscript{189} See Experiment 4 Discussion: the dual-modal design used alphabetical markers to ensure the simultaneous delivery of mutually referring elements.

\textsuperscript{190} See Hypothesis 3 from Ginns’ meta-analysis (2002), in which effects were larger for system-paced presentation.
becomes easier because spoken text is linear by nature. It should be noted nevertheless, that the materials used by Tabbers et al. did not conform to the normal split-source materials required to obtain the modality effect. The modality effect, like the split-attention effect, should be obtainable only when using materials that are unintelligible in isolation and therefore require integration to be understood. Tabbers et al. used materials in which the visual information could be understood in isolation and in which the auditory information was redundant. The modality effect should not be obtainable under such circumstances.

A second area of interest arising from these studies concerns the theoretical implications associated with the auditory processing of both music and text. Even though uni-modal auditory processing was not a central consideration of this empirical work, the results of Experiment 3 raised certain issues relating to our understanding of working memory constructs. Despite the equivocal nature of these data, there is preliminary evidence that auditory presentation of both music and text placed increased processing loads on working memory. Nonetheless, both auditory conditions performed either equally or above their simultaneous and successive dual-modal counterparts respectively.

Although future research is required into the relative merits of auditory only musical formats, more recent studies do provide findings pertinent to the current results (see also section 1.4.4). For example, Baddeley (1990) refers to a modality effect in which presented items are more likely to be remembered if heard rather than read, an encoding effect (Cowan, Scott Saults, & Brown, 2004) largely attributable to superior recall of more recent items. What is more, this effect is inhibited by additional speech, but not by non-speech sounds (see Crowder & Morton, 1969), a result that holds only if sounds are perceived in a music- rather than speech- based context (Ayres, Jonides,
Reitman, Egan, & Howard, 1979). Across a number of studies, this modality effect has been associated with materials that range from serial recall to more complex sequential processes (Conway & Christiansen, 2005). Even though the basis of this effect is yet to be fully determined, possible explanations point to the temporal advantages of auditory delivery (Cowan et al., 2004) and/or the increased investment of attentional resources that are devoted to auditory processing (Foos & Goolkasian, 2005).

When taken together, these findings support the proposition that auditory processing of text carries inherent advantages over the same materials processed visually. Consider for instance, the working memory model (see section 1.4.3) indicates that although auditory speech has direct access to the phonological store via auditory perception, visual text has only indirect access via a conversion process in the articulatory loop (Jones, Macken, & Nicholls, 2004). Furthermore, it appears that although auditory musical excerpts may be processed and integrated with only minimal disruption to auditory speech-based materials (see Deutsch, 1975a), it is uncertain whether music is independent of the phonological loop (see section 2.2). At present, these data point to at least partially independent working memory mechanisms for auditory music and auditory text (see Berz, 1995); such mechanisms may at least partially avoid either the potential conflict and/or increased mental load that might otherwise arise from combining elements received through the same processing modality. These interpretations, although somewhat speculative, are compatible with the better than expected performance of the auditory conditions for Experiment 3. Additional research would help to establish within a CLT framework whether auditory only processing offers a viable instructional alternative for these music materials.

Notwithstanding the results for these auditory conditions, they do not explain the relatively poor performance of the successively delivered dual-modal condition. Even
though both successive (uni- and dual-modal) groups were associated with temporal split attention, the negative effects of holding and referring relevant information between sources was more evident when the text was at first read rather than heard. It would appear that mentally integrating successively presented dual-modal materials, including the reprocessing of visual text into an acoustic (phonological) code, is particularly susceptible to interference (e.g., see Norris, Baddeley, and Page (2004) and their findings for articulatory suppression\textsuperscript{191}). These issues deserve further attention in order to understand the underlying cognitive mechanisms of both dual- and uni-modal (auditory) processing.

A third issue generated by the findings of this empirical work corresponds to a recent trend in the extension and modification of cognitive load theory: the management of intrinsic cognitive load as a factor of authentic and complex learning tasks (van Merriënboer & Ayres, 2005; van Merriënboer & Sweller, 2005)\textsuperscript{192}. Somewhat unusually, Experiment 4 considered the total cognitive load of efficient instructional formats rather than the additional cognitive load generated by inefficient design. Because the extent to which elements interact is a product of both acquired schemas and the intrinsic difficulty of the materials, students may find, regardless of instructional formatting, some materials intractable to learn without further modification. To this end, the results for the overlapping design of Experiment 4 are assumed to provide indirect

\textsuperscript{191} Norris et al. (2004) studied the effects of irrelevant speech on serial recall and provided “additional support for the WM [working memory] view that the reason AS [articulatory suppression] eliminates the effect of IS [irrelevant sound] during the input of visual lists is because suppression prevents the phonological recoding of visual material” (p. 1104).

\textsuperscript{192} Van Merriënboer & Sweller (2005) also state, “With regard to media, CLT is no longer limited to straightforward instructional presentations or linear media but has found several applications in the design of multimedia systems” (p. 173); materials within this thesis have of course pursued this instructional direction.
evidence for the construct of intrinsic cognitive load and the necessity for learners to acquire schemas within the quite severe limitations of working memory.

As reported earlier, a similar rationale of lowering cognitive load by, in effect, distributing or fragmenting it across instructional stages was used to generate the isolated interacting elements effect (Pollock et al., 2002). As earlier outlined, this instructional technique at first precludes holistic understanding in order to construct the very schemas upon which this understanding depends. However, in contrast to the notion of part-to-whole understanding, Experiment 4 maintained holistic meaning across both stages, although perception was incrementally reinforced from both visual (notation) and auditory (excerpt) perspectives\(^{193}\). This modified approach was seen as an efficient way to employ limited mental resources and yet to readily achieve holistic understanding.

Music materials lend themselves to this efficient design due to the reinforcement achieved by dual representation (visual notation-auditory excerpts). For music, this format can be seen as an accompanying strategy to the isolated interacting elements technique, both procedures recognising that schemas can only be constructed when processing loads remain within working memory capacity. However, where the mental load of large informational episodes (e.g., longer auditory excerpts) exceeds working memory capacity, the isolated interacting elements technique may be the only viable means of mentally integrating the required elements. Further research would help to establish the instructional contexts in which these techniques should be adopted, ensuring for both techniques a more accurate match between musical formats and target

\(^{193}\) Recall that the two visual sources were presented whole before the third source (auditory excerpt) was superimposed. In addition, this auditory source constituted an equivalent or complementary source to information already presented (musical notation).
With respect to music, identifying theoretical principles underpinning the incorporation of three information sources is crucial. The integrated use of explanatory text, auditory excerpts and musical notation is not only instructionally common, but is also an efficient way of learning about music: All three are essential to any claim of musical literacy\(^{194}\). Nonetheless, whether the currently employed methods of combining these instructional sources are, from a cognitive load perspective, always effective is a question that requires ongoing research.

It is a limitation of the current studies that the performance measures, levels of cognitive load and/or the structuring of learning episodes may not have been sensitive enough to fully investigate differences between conditions with or without a third source of information (auditory excerpts or musical notation)\(^ {195}\). For some materials (e.g., Experiment 5), it is plausible that simple corresponding notational-auditory excerpts might represent superfluous or even redundant material, although for more complex materials (e.g., Experiment 6) it is difficult to believe that auditory excerpts alone would be sufficient to facilitate understanding. Further research is required to ensure that the potential interaction between complexity and equivalent music sources (notational-auditory) is accounted for in future experimental materials. Although the current results can only tentatively address these broader issues, they suggest that cognitive load factors play a major role in determining effective design and therefore CLT provides a useful basis for further exploring these issues within the domain of

\(^{194}\) We should also be mindful that there is considerable evidence for the instructional benefits of combining verbal coding and non verbal coding of new materials (Clark & Paivio, 1991).

\(^{195}\) Also, see recommendations that were made in the Discussion for Experiment 6 regarding future design recommendations in order to fully investigate this issue.
Within the current theoretical framework, a fourth and final issue arising from this empirical work reflects a crucial concern of effective instructional design: the accurate estimation and measurement of cognitive load. Variably, the current studies directly (subjective ratings) and indirectly (performance) support the central importance of this cognitive construct. For example, when cognitive load is closely matched with learner populations (e.g., Experiments 2), appropriate instructional formats maximise understanding. On the other hand, where cognitive load is too low and/or there is sufficient instructional time and practice (e.g., Experiment 1), learning proceeds irrespective of the efficiencies, or otherwise, of instructional formats. Likewise, where the number of interacting elements generates excessive cognitive load (Experiment 4), understanding is likely to be impaired, despite an otherwise efficient instructional design. Notwithstanding these findings, the expected pattern of relative lower cognitive load and superior performance results was not always evident. The factors that might lead to such an observation are discussed below.

To strengthen these observations, and to further generalize findings, it is recommended that future investigations incorporate different populations, including mainstream secondary students. It is also uncertain as to the accuracy and validity of cognitive load measurements associated with young students, a significant issue and current limitation given the nature of the student population used throughout the reported experiments. Additionally, although motivation—a component of mental effort and performance—has been recently identified as a cognitive load factor (Paas, Tuovinen, van Merriënboer, & Darabi, 2005), it was not considered in the current

---

196 The students in the current studies ranged from 11-13 years of age.
experiments. Therefore, further work is recommended in relation to motivation and its influence on the effectiveness of music formats. Somewhat related to this issue is the nature of the cognitive load that is measured: Higher ratings may be a reflection of either excessive load (intrinsic) or a reflection of invested mental effort (germane load). Whatever the issue, accurately measuring and comparing cognitive load has always proved a challenge (see section 3.2.4); be that as it may, it remains the central theoretical issue underpinning CLT (Paas et al., 2003b).

These experiments also required that subjective ratings accurately reflect cognitive load differences regardless of whether problems were either well defined or ill defined by nature. For example, it was relatively straightforward to adjust levels of element interactivity for Experiments 1, 2 and 4 due to the algorithmic processes that directed and engaged the learner through a systematic series of instructional steps. However, it was far more challenging to predict processing loads when dealing with either the melodic or textural variations of Experiments 3 and 6 respectively. In these cases, understanding was not necessarily directed by a series of predetermined steps, but by the connections students were able, or willing, to make between elements. This uncertainty when estimating cognitive load may impose limitations on the way findings can be transferred, or predictions made, from one set of materials to another.

Investigating the extent to which ill-defined problems ensure learner engagement and generate estimable levels of cognitive load present productive challenges for future research. Of related interest, multimedia environments are now providing greater flexibility for learners to interact with instructional formats.197 Nevertheless, recent studies have emphasised that when applying the principles of CLT within flexible

197 These interactions include self-paced learning, selecting elements, or receiving feedback.
instructional contexts that choice has to be moderated by meaningful feedback and sufficient guidance (Moreno & Mayer, 2005; Moreno & Valdez, 2005). Within the current studies music materials were limited by the extent to which learners could control the acquisition of schemas, an area that in the future would provide a useful extension of these findings.
10.4 Practical Implications of Research Findings

Generating effective instructional formats for music was a key objective of these empirical investigations: Any instructional theory is valuable in so far as it can produce workable and transferable instructional applications (Sweller & Chandler, 1991). For this reason, the following recommendations for the structuring and presentation of music materials are suggested by the findings of these experiments:

1. Two mutually referring visual information sources (e.g., explanatory text and musical notation) should be physically integrated in preference to presenting the same information in a split-source format. This physical integration of sources is only necessary where both are required for intelligibility.

2. Allocating two mutually referring information sources between two modes (e.g., auditory text and musical notation) is an effective alternative to physically integrating two sources of information, and a superior alternative to presenting the same information in a split-source format (e.g., visual text and musical notation).

3. Dual-modal information sources are processed more efficiently if presented simultaneously rather than in a format that requires the successive processing of each source. Where total cognitive load remains within working memory capacity, the principle of simultaneity applies to either a combination of two (either auditory or visual text, and either visual or auditory music) or three (e.g., visual text, visual music and auditory music) sources of mutually referring information.
4. Where total cognitive load of processing two or three information sources exceeds working memory capacity (intrinsic cognitive load), the materials should be presented in an incremental process that either avoids or minimises the extraneous cognitive load of split-attention. Depending on the complexity of the materials, this can be achieved in either a part-to-whole format or in a way that preserves holistic understanding. In this way the intrinsic load is distributed between two instructional episodes without otherwise violating the principles and cognitive efficiencies of integration and simultaneity.

5. Comparative materials should always be presented in ways that consider instructional efficiency and the negative effects of extraneous cognitive load. Experiment 5 demonstrated that when comparative excerpts consist of even the most basic combination of text (1-2 words) and musical excerpts, levels of cognitive load are sufficient to differentiate between instructional formats.

At present, there are no reasons to believe that these five principles would not remain viable for all combinations of music-related information, subject of course to the total cognitive load generated by the instructional materials and mediated by the experience (schemas) of learners\(^{198}\). The length of both explanatory statements and musical excerpts remains a practical limitation associated with all five recommendations. The effects reported in this thesis are based upon the delivery of information in manageable blocks that control the number of instructional elements and therefore the intrinsic cognitive load of integrating information. This observation is

\(^{198}\) These practical recommendations should also be viewed within a broader framework of effective instructional principles (see for example Astleitner, 2005).
especially pertinent to instructional activities such as score reading, where the visual and/or auditory tracking of large musical passages is often required. It is not of course suggested that such staple instructional practices are invalid; the breadth of many musical works requires extended observation. However, where understanding of musical concepts requires the integration of multiple elements, educators should be mindful of the cognitive load generated by the magnitude of the information source/s they ask students to comprehend.

The recommendations above are practical and could be readily transferred to a classroom environment. Now more than ever, new technologies provide the affordable means by which the materials of music can be appropriately structured and transmitted to either individuals or larger groups. In particular, computer-based instruction provides relatively quick access to many design features (Koroghlanian & Klein, 2004); the possibility of sophisticated and tailored teacher-designed instruction is now a reality. Of course, the widespread adoption of digital technology does not guarantee effective instruction. On the contrary, the allure of multiple functions, extraordinary sound fidelity and engaging activity can often mask, or worse counteract, the means by which technology should be used as an instructional tool. “In short there is nothing special about cutting-edge educational technology that automatically fosters learning; instead, effective instructional methods are needed regardless of the delivery medium that is used” (Moreno & Mayer, 2004, p. 171).

The implications of these suggestions also extend beyond teacher-designed

199 In fact, a preparatory instructional activity could involve students familiarising themselves with instructional exemplars, together or in isolation, in order to provide the knowledge differentiation and “contrasting cases” necessary for later integration and deeper understanding (Schwartz & Bransford, 1998).

200 For example, see Morrison and Anglin (2005) for a recent review of CLT strategies for e-learning.
instruction to the many commercial products that now flood the market. Many of these products structure their materials on aesthetic, desktop publishing or commercial principles rather than considering the more important cognitive imperatives that should guide instructional design. For example, in recent years interactive computer software has been increasingly incorporated into mainstream classrooms for purposes that range from musicological research to structured aural training. The potential benefits of self-paced and autonomous learning make these activities an attractive option. Nevertheless, there is currently a paucity of research to support the methodologies often preferred by software designers and, of even greater concern, there exists a considerable lag between published research and its acceptance by the mainstream education community.
10.5 Conclusion

Cognitive load theory provides the means to direct the effective design and presentation of music instruction. This conclusion is based upon the successful application of instructional designs across a range of familiar and novel settings. It should be emphasised that consistent explanations of cognitive load phenomena were found despite the incorporation of both modality and multiple information sources in ways quite distinctive to the instruction of music. In addition, all experiments were conducted with academically able students, further generalising the application of cognitive load principles to participant populations that now span primary school children (Bobis et al., 1993; Bobis, Sweller, & Cooper, 1994) through to elderly learners (Paas, Camp, & Rikers, 2001; van Gerven et al., 2002).

Our understanding of cognitive architecture and the universal principles that determine the ways in which we learn have developed enormously in recent decades. Based on this understanding, CLT posits a series of principles that were fundamental to developing the instructional designs within this thesis. First, working memory is inordinately small; only a few novel interacting elements can be processed at one time. Second, processing materials between two modes expands effective working memory capacity. Third, if mutually referring explanatory text and auditory music have to be presented as split-source materials, an auditory only format should be preferred over dual-modal delivery. Fourth, the extraneous processing required when physically and/or temporally splitting attention between two or three sources of information raises cognitive load and impairs understanding. Fifth, the development of schemas is the basis of all understanding, a process that can only occur when the total cognitive load of interacting elements remains within working memory capacity. These five principles are essential to the effective and cognitively efficient design of music instruction, a
conclusion supported within this thesis by the instructional formats that facilitated understanding.


Learning, Memory and Cognition, 15, 977-989.


superiority effect in serial recall: Separating input and output factors. *Journal of
Experimental Psychology: Learning, Memory and Cognition, 30*, 639-644.

multimedia educational environments: Effects of agent properties, picture features,
and redundancy. *Journal of Educational Psychology, 94*, 428-434.

Learning and Verbal Behavior, 8*, 658-664.


Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in

L. Nilsson & T. Archer (Eds.), *Perspectives on learning and memory* (pp. 19-37).

*Thinking in sound: The cognitive psychology of human audition* (pp. 113-145).

and Psychophysics, 5*, 365-373.


Psychology, 3*, 255-267.


Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook


Appendix A

Examples of perceptual recognition phenomena and Gestalt grouping laws

Sensory perception is at times described by the universal tendencies to organise or group patterns of stimuli (Gestalt grouping laws). The examples below illustrate the ways in which we perceive certain objects; in some cases the same object gives rise to two different perceptions.

(Taken from Leman, 1995, p. 62)

(Taken from Sloboda, 1985, p. 168)

(Taken from Leahey and Harris, 1997, p. 116).
Appendix B

Examples of mnemonic techniques

Visual Peg

Each item is linked to a ‘landmark’, which will anchor the associated memory in the mind. Where more than one item is to be memorised, they can be linked to a common landmark, e.g., ball, boot and grass can be linked to football.

Story Method

Each item to be memorised is linked to the other in a narrative that makes use of associations between items and an interesting (surreal, fantastic, colourful etc) progression of narrative ideas.

Journey Method

Each independent item is associated with a pre-selected landmark or location that is highly familiar and memorable to the learner. For a series of items that need to be remembered in order, the landmarks can form a natural progression of locations, such as walking through various rooms of a home. This method uses the principles underlying the peg and narrative methods.

(The above mnemonic ideas were taken from O’Brien, 2000)

First Letter Method

To remember a series of items, the first letter of each can be placed into a memorable phrase or sometimes even a word. For example, the notes of the treble and bass staves in music have been remembered for generations in the following ways (bold):

Treble Stave:  **Every Good Boy** Deserves Fruit (lines)  **FACE** (spaces)
Bass Stave:    **Good Boy** Deserves Fruit **Always** (lines)  **All Cows Eat Grass** (spaces)
Appendix C

The complete “washing clothes” passage from Bransford and Johnson (1973)

The passage below illustrates the power of a schema to direct, or misdirect, understanding of new information. Intelligibility of the instructions below relies on the reader having activated a schema for washing clothes.

The procedure is actually quite simple. First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily arise. A mistake can be expensive as well. At first the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to see any end for the necessity of this task in the immediate future, but then one can never tell. After the procedure is completed one arranges the materials into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, that is part of life.

(Taken from Bransford and Johnson, 1973, p. 722).
Appendix D

The Monk problem

The following provides a description of the classic Monk problem.

Once there was a monk who lived in a monastery at the foot of a mountain. Every year the monk made a pilgrimage to the top of the mountain to fast and to pray. He would start out on the mountain path at 6 a.m., climbing and resting as the spirit struck him, but making sure that he reached the shrine at exactly 6 p.m. that evening. He then prayed and fasted all night. At exactly 6 a.m. the next morning, he began to descend the mountain path, resting here and there along the way, but making sure that he reached his monastery again by 5 p.m. of that day.

That evening as he was hastening to a much needed dinner, he was stopped by the monastery’s visiting mathematician, who said to him, “Do you know, I suddenly realized a very curious thing. Every time you make your pilgrimage there is always some point on the mountain path, perhaps different on each trip, that you pass at the same time when you are climbing up as when you are climbing down.” “What!” snorted the monk, annoyed. “Why, that’s ridiculous! I walk at all manner of different paces up and down the path. It would be a great coincidence if I should pass any spot at the same time of day going up as coming down. The idea that such a coincidence might happen time after time surpasses belief!” The mathematician, who had a touch of friendliness in his soul, smiled sweetly and said, “Bless you, Brother, not only should you believe it, but if you will just think about it in the right way, it’s obvious.” He then locked himself in his cell, confident that he had spoiled the monk’s dinner and probably his night’s sleep as well.

The graph below provides an alternative representation of the Monk problem in which the solution is readily observable.

(Passage and graph taken from Hayes, 1978, pp. 177-178).
Appendix E

An example of functional fixity

The following examples of functional fixity demonstrate the tendency to view problems from a preconceived perspective rather than flexibly adapt our thinking to the requirements of a problem.

The goal is to place three small candles at eye level on a door. Among other objects on a nearby table are a few tacks and three small boxes about the size of matchboxes. In one condition the boxes were filled with candles, tacks, and matches. In another condition the boxes were empty. The solution requires tacking the boxes to the door so they can serve as platforms for the candles.

(Taken from Reed, 1992, p. 287).

It was more difficult for the group with the complete matchbox to envision its use in an unconventional way, whereas the group with the empty matchbox were more amenable to its flexible adaptation—as a platform—to solve the problem.

(Taken from Reed, 1992, p. 287).
Appendix F

The Water Jug problem

Given three jugs (A, B, C,) where A can hold 8 units, B 5 units and C 3 units, A is initially full, and B and C are empty. The subject’s task is to find a sequence of pourings that would divide the contents of the largest jug evenly between the largest and middle-sized jug. In pouring, water is transferred until the jug the subject is pouring from is empty or the jug being poured into is full. Water cannot be added or discarded during the course of solving the problem.

(Gilhooly, 1996, p. 29)
Appendix G

Means-ends analysis model used by Sweller (1988)

The following model was used to track the number of productions required by a learner to reach a goal.

(Taken from Sweller, 1988, p. 269)
Appendix H

Survey of musical experience

The experimenter guided students through the survey and ensured there was a consistent understanding and response from each group.

STUDENT SURVEY

History of Music Experience

A. Do you or have you played a musical instrument? Yes / No circle
   If Yes, name the:

   Instrument 1
   Indicate the year in which you commenced playing .......... and the year you ceased playing .......... circle
   Did/do you have tutoring or lessons? Yes / No circle
   If Yes, for how many years? 1 2 3 4 5 6 7+ circle
   Have you completed any external examinations for your instrument (e.g. AMEB, Trinity etc). Yes / No circle
   If Yes, give the last grade .......... Year ..........
   Have you played in a musical group/band etc (school or community). Yes / No circle
   If Yes, name the:
   Group: ........................................

   Circle the years in which you were a member:

B. Have you completed any external examinations in musicianship or theory (e.g. AMEB, Trinity etc).
   Yes / No circle
   If Yes, give the last grade .......... Year in which attempted ..........

C. In the previous two years of Primary school, did you have regular or occasional music lessons?
   circle one
   In these lessons, did you learn to read and understand the language of music? Yes / No Circle
   On arriving at this school, would you describe your understanding of music as:
   very strong strong moderately strong moderately weak weak very weak

(Scanned image is smaller than the original A4 sheet)
Appendix I

Music Response Sheet

An example of the subjective rating scale used to calculate cognitive load.

Music Response Sheet

Please answer the following question by ticking the box under the number that nearest reflects your response to the flash card shown below.

The Flash Card represents problems in the Problem Sheet you have just completed.

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Extremely easy to understand</td>
<td>Very easy to understand</td>
<td>Fairly easy to understand</td>
<td>Tending to be easy to understand</td>
<td>Neither easy nor difficult to understand</td>
<td>Tending to be hard to understand</td>
<td>Fairly hard to understand</td>
<td>Very hard to understand</td>
<td>Extremely hard to understand</td>
</tr>
</tbody>
</table>

(Scanned image is smaller than the original A4 sheet)
Appendix J

Examples of the Experiment 2 instructional format for the integrated condition

(Please refer to the instructional sheet for the time signatures and note values.)

Example 4 gives the value of each note when the lower number of a time signature indicates a half note or minim.

Example 4a: \[ \frac{3}{2} \quad \text{or} \quad \frac{4}{2} \]

Study Example 4b: Follow the sequence of letters, a to g below.

a. semiquaver = \[ \frac{1}{8} \] of a minim beat
b. quaver = \[ \frac{1}{4} \] of a minim beat
c. crotchet = \[ \frac{3}{8} \] of a minim beat
d. dotted crotchet = \[ \frac{5}{8} \] of a minim beat
e. minim = \[ \frac{1}{2} \] of a minim beat
f. dotted minim = \[ \frac{3}{2} \] minim beats
g. semibreve = \[ \frac{2}{2} \] minim beats

Turn page when instructed.

(Scanned image is smaller than the original the A4 sheet)
Study Example 2: Follow the sequence of letters, a to c below.

a. The upper number of a time signature indicates the total value of beats in each bar.

\[
\begin{align*}
2 \text{ beats} & \quad 1 \text{ beat} \quad \frac{1}{2} \text{ beat} \quad \frac{1}{2} \text{ beat} = 4 \text{ beats} \\
1 \text{ beat} & \quad \frac{1}{2} \text{ beat} \quad \frac{1}{2} \text{ beat} \quad 2 \text{ beats} = 4 \text{ beats}
\end{align*}
\]

b. Bars 1 and 2 give two possible note combinations that equal the value of 4 beats in each bar.

c. Like a fraction, the lower number of a time signature indicates the type of note value by which all other notes will be measured. For example:

\[
\text{Lower Number } 4 = \begin{cases} 
\text{quarter note} \\
\text{crotchet}
\end{cases}
\]

Therefore, in this example, the crotchet is used to measure the length of all other notes.

Turn page when instructed.
Appendix K

Examples of the melodic material from Experiment 3

The melodic material below is based on *Twinkle Twinkle Little Star*. Each version varied a targeted pitch concept.

Pitch concept – rising contour

Theme

![Melodic material](image1)

Variation

![Melodic material](image2)

Pitch concept – larger range

Theme

![Melodic material](image3)

Variation

![Melodic material](image4)
Examples of the melodic material used for transfer problems in Experiment 3

The example retained the same level of intrinsic difficulty as the earlier melodic material.

Transfer Problem 2

Theme

Variation
Appendix M

Examples of written and aural tests for Experiment 4

The written questions required the completion of each bar by the addition of either one or two notes.

**Written Tests**

1. 

2. 

3. 

4. 

**Aural Test**

In the following problems you will hear an incomplete bar of rhythm. After listening to the given rhythm, you are to complete the bar by adding one note. Circle your answer from the selection of note values beside the question number on the sheet provided. Each new question number will be announced.

**Practice Problem**

Circle one note that completes the given bar of rhythm

1. 

(Scanned image is smaller than the original A4 sheet)
Examples of musical variations for Experiment 5

Variations represented a range of concepts, each stave (for the simultaneous with notation condition) with accompanying text and audio was delivered on a separate Powerpoint slide.

Rhythm values …. 

Pitch direction …. 

May be opposite 

Shorter phrases …. 

Longer phrases
Appendix O

Estimate of interacting elements for Experiment 5

Elements that need to be considered when learning each concept (equal to two comparative Powerpoint slides).

Slide 1.
1. Record the name of the concept reference (e.g., rhythmic pattern, or, smooth shape)
2. Relate concept and descriptive reference to music.

Slide 2.
3. Record the name of the concept reference (e.g., may change, or, angular shape, respectively).
4. Relate concept and descriptive reference to music.
5. Compare slides 1 and 2.
Appendix P

Examples of the test questions for Experiment 5

Single Element - Legato to staccato

Theme
\[ \frac{4}{4} \]
\[ \begin{array}{cccccccc}
\text{C} & \text{E} & \text{G} & \text{C} & \text{G} & \text{C} & \text{G} & \text{C} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\end{array} \]

(legato)

Variation
\[ \frac{4}{4} \]
\[ \begin{array}{cccccccc}
\text{C} & \text{E} & \text{G} & \text{C} & \text{G} & \text{C} & \text{G} & \text{C} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\end{array} \]

(staccato)

Multiple Elements - Phrase length and accent

Theme
\[ \frac{4}{4} \]
\[ \begin{array}{cccccccccccc}
\text{C} & \text{E} & \text{G} & \text{C} & \text{G} & \text{C} & \text{G} & \text{C} & \text{G} & \text{C} & \text{G} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\end{array} \]

(staccato)
Appendix Q

Estimate of interacting elements for Experiment 6

Elements that need to be considered when learning each concept (equal to two comparative Powerpoint slides).

Slide 1.
1. Record the similarity/sameness of the concept for two instruments (e.g., rhythms, or, rhythmic activity).
2. Relate the concept similarity/sameness to a unified foreground for two instruments.
3. Relate the concept similarity/sameness for two instruments to the music (may involve more than one element for some or most students).

Slide 2.
4. Record the different emphasis of the concept between two instruments (e.g., rhythms, or, rhythmic activity).
5. Relate the different emphasis of the concept between two instruments to one or both instruments in the foreground.
6. Relate the different emphasis of the concept between two instruments to the music (may involve more than one element for some or most students).
7. Compare slides 1 and 2.
Appendix R

Examples of the test questions for Experiment 6

One instrument is promoted to the foreground by the use of greater rhythmic activity. These materials were delivered aurally and therefore only heard.

\[ \text{Instrument 1} \]

\[ \text{Instrument 2} \]