ATTENTIONAL RESOURCING: THE ROLE OF COGNITIVE COMPLEXITY AND COGNITIVE LOAD

INTEGRATING MULTIPLE- AND UNITARY-MODEL PREDICTIONS INTO A TWO-LEVEL HIERARCHICAL FRAMEWORK

NATALIE MAY VIRGINIA MORRISON

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF WESTERN SYDNEY

SEPTEMBER 2010
DEDICATION

To Ben – My husband, My Rock and My Guide. Everyday and everyplace you are always by my side.

To Mum – My best friend. I just can’t put into words the help you have given me throughout my candidature and more importantly through all the years leading up to it and no doubt beyond it.

To Dad – For always being proud of me and never failing to challenge me to do better in all aspects of my life.

To Jonte – For making me feel like the centre of your universe with a simple nuzzle or a gentle snore.

To Ross and Sharon – For never letting a day go by without making sure my world is the best place to be.

To Linda, Greg and Jasmin – For providing that time-out every week that made getting up and working harder that little bit easier to manage.
ACKNOWLEDGEMENTS

Denis Burnham – My primary supervisor, for getting me through every tough moment and for providing hours of care, thought, and knowledge both for me and for this thesis.

Mark Wiggins – For inspiring me and encouraging me to take on my PhD and for keeping me going when perhaps no one else could have.

Roger Dean – For stepping in when times were tough and giving me the confidence to continue, and, for then providing the most insightful feedback as the thesis, and its write up, continued to develop.

MARCS Laboratories – For providing first rate support staff, equipment and resources for without this thesis would not be finished.

Chris Davis and Jeesun Kim – For getting me off the ground and running in the early stages of my candidature.

Tracy Mills and the Research Office Team – For always making sure that everything that needed doing throughout my candidature was done, in advance, with minimal stress.

To the Students (and Ex-Students) at MARCS who helped me through the big and small problems, and who sympathised and partied at all the right times. Thank you Ben Morrison, Jemma Harris, Nathan Perry, Bronson Harry, Shaun Halovic, Kirk Olsen, Dogu Erdener, Sonia Wilkie, Nicole Lees, Christa Lam-Cassatari, Mark Antoniou and Nan Xu and Josephine Terry.
STATEMENT OF AUTHENTICATION

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the test. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

______________________________
Natalie May Virginia Morrison

______________________________
Date this day, 31st of March 2010
TABLE OF CONTENTS

Title Page i
Dedications ii
Acknowledgements iii
Statement of Authentication iv
Table of Contents v
List of Tables xii
List of Figures xiv
Abstract xxi
Setting the Scene xxii

1. An Introduction to Auditory and Visual Attention 24
   1.1 Chapter Overview 25
   1.2 Conceptualising the Nature of Attention 25
   1.3 Attention: As a ‘Selective Filter’ or as a ‘Cognitive Resource’ 28
       1.3.1 Unitary-Resource Models 36
       1.3.2 Multiple-Resource Models 39

2. The Great Divide: A Review of the Literature Contributing to the Resource Divide 45
   2.1 Chapter Overview 46
   2.2 Introducing the Great Divide: One Phenomenon, Two Contradictory Models 46
   2.3 Explaining the Divide: Evaluating the Methods of Past Studies 49
       2.3.1 Experimental Methods of Multiple-Resource Studies 49
2.3.2 Experimental Methods of Unitary-Resource Studies
2.3.3 Summary: Methods Used in Studies Supporting the Classic Resource Models
2.3.4 Hybrid Attentional Models
2.4 Introducing Cognitive Load
2.5 Thesis Rationale and Research Questions
3. Experimental Design – Developing a Model-Neutral Method of Examining Attentional Resources and Cognitive Load
   3.1 Chapter Overview
   3.2 Introducing the Design Methods to be Used in this Series of Studies
      3.2.1 Independent Variables and Contrasts of Interest
      3.2.2 A Task Framework to Examine Multi-Modal Attention
      3.2.3 Manipulation of Cognitive Load
         3.2.3.1 Objective Measures of Cognitive Load
         3.2.3.2 Subjective Measures of Cognitive Load
      3.4 Summarising the Method Framework to be Used in this Series of Studies
4. Study One: Manipulating Cognitive Load with Masking
   4.1 Rationale
   4.2 Experimental Design
   4.3 Method
      4.3.1 Participants
      4.3.2 Materials and Apparatus
      4.3.3 Stimuli and Pilot Studies
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.3.1</td>
<td>Alphabetic Letter Presentation and Duration</td>
<td>97</td>
</tr>
<tr>
<td>4.3.3.2</td>
<td>Selection of Target and Distractor Letters</td>
<td>99</td>
</tr>
<tr>
<td>4.3.3.3</td>
<td>Developing the Auditory and Visual Letter String Sequences</td>
<td>100</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Procedure</td>
<td>104</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>105</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Data Reduction</td>
<td>105</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Data Screening</td>
<td>107</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Data Analysis Overview</td>
<td>108</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Analyses</td>
<td>108</td>
</tr>
<tr>
<td>4.4.4.1</td>
<td>Performance: Modality Combination and Cognitive Load</td>
<td>108</td>
</tr>
<tr>
<td>4.4.4.2</td>
<td>Complexity, Difficulty and Confidence Ratings</td>
<td>111</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion</td>
<td>114</td>
</tr>
<tr>
<td>5</td>
<td>Study Two: Manipulating Cognitive Load with Time</td>
<td>118</td>
</tr>
<tr>
<td>5.1</td>
<td>Rationale</td>
<td>119</td>
</tr>
<tr>
<td>5.2</td>
<td>Experimental Design</td>
<td>119</td>
</tr>
<tr>
<td>5.3</td>
<td>Method</td>
<td>120</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Participants</td>
<td>120</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Materials and Apparatus</td>
<td>120</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Stimuli</td>
<td>121</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Procedure</td>
<td>122</td>
</tr>
<tr>
<td>5.4</td>
<td>Results</td>
<td>124</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Data Reduction</td>
<td>124</td>
</tr>
</tbody>
</table>
5.4.2 Data Screening 124
5.4.3 Data Analysis Overview 125
5.4.4 Analyses 125
   5.4.4.1 Performance: Modality Combination and Cognitive Load 125
   5.4.4.2 Task Complexity and Task Difficulty: The Differences and Similarities 128
   5.4.4.3 Task Confidence: Establishing Appropriate RT Windows 131

5.4 Discussion 132

   6.1 Rationale 139
   6.2 Experimental Design 142
   6.3 Method 143
      6.3.1 Participants 143
      6.3.2 Materials and Apparatus 144
      6.3.3 Stimuli 145
      6.3.4 Procedure 145
   6.4 Results 148
      6.4.1 Data Reduction 148
      6.4.2 Data Screening 148
      6.4.3 Data Analysis Overview 149
      6.4.4 Analyses 150
         6.4.4.1 Performance: Modality Combination and Cognitive Load 150
6.4.4.2 Effects of Germane Cognitive Load, Motivation and Exposure, on Performance

6.4.4.3 Interactions of Extraneous and Germane Cognitive

6.4.4.4 Predicting Performance from Complexity, Difficulty and Confidence

6.4.4.5 Comparison of Subjective Measures and Objective Accuracy

6.5 Discussion

7. Study Four: A Hierarchical Model of Attentional Resourcing

  7.1 Rationale

  7.2 Experimental Design

  7.3 Method

    7.3.1 Participants

    7.3.2 Materials and Apparatus

    7.3.3 Stimuli

    7.3.4 Procedure

  7.4 Results

    7.4.1 Data Reduction

    7.4.2 Data Screening

    7.4.3 Data Analysis Overview

    7.4.4 Analyses

      7.4.4.1 Modality Combination and Cognitive Load: Performance Across Five Levels of Load

      7.4.4.2 Single- vs. Dual-Task Performance
7.4.4.3 Dual- vs. Tri-Task Performance 188
7.4.4.4 Complexity, Difficulty and Confidence 189

Ratings 189

7.5 Discussion 198

8. General Discussion 202

8.1 Chapter Overview 203
8.2 Overview of Experimental Results 204
8.3 Interpreting the Thesis Findings 208

8.3.1 Auditory and Visual Resource Modelling 208
8.3.2 Comparing Cognitive Load and Cognitive Complexity 211

8.4 Future Research Directions 213
8.5 Implications 215
8.6 Conclusions 218

9. References 220

Appendix A Listing and Pronunciation of all 26 Alphabetic Letters
Appendix B Ratings Questionnaire: Studies One, Two and Three
Appendix C Pilot Study One: Alphabetic Letter Presentation Duration
Appendix D Auditory Recording of 25 Letters Compressed to 150ms
Appendix E Pilot Study Two: Target-Distractor Similarity Ratings
Appendix F Table of Planned Contrasts: Studies One and Two
Appendix G Morrison, Burnham & Carlile (2010)
Appendix H Morrison, Burnham, Wiggins & Dean (2010)
<table>
<thead>
<tr>
<th>Appendix I</th>
<th>Table of Individualised RT Windows: Study Three</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix J</td>
<td>Table of Planned Contrasts: Study Three</td>
</tr>
<tr>
<td>Appendix K</td>
<td>Ratings Questionnaire: Study Four</td>
</tr>
<tr>
<td>Appendix L</td>
<td>Table of Individualised RT Windows: Study Four</td>
</tr>
<tr>
<td>Appendix M</td>
<td>Table of Planned Contrasts: Study Four</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

Table 1.1  Contrast the Unitary- and Multiple-Resource Models: Their Structures, Resource Allocation Networks and their Performance Predictions 44

Table 2.1  An Overview of the Research Studies Evaluated in this Thesis, the Measures Used, and the Model Supported 64

Table 2.2  Cognitive Load Measurement Categorisation and Example Measures for Each Category 71

Table 3.1  Experimental Methods, Stimuli and Measures as Used in those Studies Reviewed in Chapter 2 with a Summary of the Resource Model Subsequently Supported by the Study 80

Table 3.2  Validating Cognitive Load Manipulation using Subjective Questions 88

Table 6.1  Varimax Rotated Component Loadings for the Three Subjectively Rated Questions in the Reflections Questionnaire 160

Table 6.2  Means and Intercorrelations between Accuracy Percent Scores, Ratings of Complexity, Difficulty and Confidence for AA, VV and AV trials 161

Table 6.3  Standard Multiple Regression of Trial Complexity, Trial Difficulty and Trial Confidence as Predictors of Accuracy Performance 162

Table 6.4  Participant Ratings of Target Identification Confidence and Constitution of Actual Performance on These Trials 164
Table 7.1 Outline of Modality Combination x Cognitive Load Trials of Interest

Table 7.2 Participant Ratings of Target Identification Confidence and Constitution of Actual Performance on These Trials
## List of Figures

| Figure 1.1 | Example of speech stimuli presented binaurally to participants via headphones and the shadowed speech sentences produced by the participants in response to speech stimuli presented simultaneously. | 29 |
| Figure 1.2 | Unitary—Resource model performance predictions | 38 |
| Figure 1.3 | Multiple—Resource model performance predictions | 41 |
| Figure 2.1 | Depiction of how all three forms of cognitive load (extraneous, intrinsic and germane) are additive and influence the amount of residual cognitive resources available dependent on the relative size of the load for each. | 69 |
| Figure 2.2 | The proposal. (a) The proposed continuum ranging from strict multiple resources through weaker multiple- (e.g., separate-but-linked account) and unitary-support, through to strong unitary-resource support. This proposed continuum functions as a function of task complexity and perceived cognitive load. (b) A proposal of a hierarchical model which spans the entire continuum. Both the continuum and the hierarchical model span task characteristics from simple to complex, and low to high perceived cognitive load (or mental effort). | 74 |
Figure 4.1 The equipment set up for Study One. (a) dorsal view (b) lateral view.

Figure 4.2 Depiction of the presentation of the auditory and visual stimuli: (a) Auditory/auditory (AA) condition; (b) Visual/visual (VV) condition, and; (c)+(d) Auditory/visual (AV) condition.

Figure 4.3 Template for letter presentation sequences for the left hand side (LHS) and right hand side (RHS) string pairings – top and bottom horizontal lines respectively. On both the LHS and RHS an auditory/visual prime will be presented at the 200ms mark (duration of 150ms) to ready participants for the trial commencement. Each letter (150ms in duration) is represented above by the rectangular boxes rising from the horizontal lines. As can be seen, no more than one letter is presented at the same time. The dashed vertical line indicates the point at which the first target may appear. The break (~) shown towards the end of both horizontal lines indicates that a period of silence was been omitted from this visual depiction of the temporal parameters of the trial.

Figure 4.4 Example experimental trials (a) AA trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load (b) VV trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load (c) AV trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load.

Figure 4.5 Mean participant RTs (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).
Figure 4.6 Mean participant accuracy percent scores (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).

Figure 4.7 Mean ratings of cognitive complexity (+SE) for each of the modality combination and cognitive load conditions.

Figure 4.8 Mean ratings of task difficulty (+SE) of all modality combination (AA, VV and AV) and perceived cognitive load (low, medium and high) conditions.

Figure 4.9 Mean ratings of task confidence (+SE) of all modality combination (AA, VV and AV) and perceived cognitive load (low, medium and high) conditions.

Figure 5.1 Template for letter presentation sequences under: (a) Low cognitive load (as shown for an AA trial). (b) Medium cognitive load (as shown for a VV trial). (c) High cognitive load (as shown for an AV trial).

Figure 5.2 Mean participant RTs (+SE) across modality combination and cognitive load conditions.

Figure 5.3 Mean participant accuracy percent scores (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).

Figure 5.4 Mean ratings of task complexity (+SE) across the three modality combination and three cognitive load levels.

Figure 5.5 Mean ratings of task difficulty (+SE) across the three modality combination and three cognitive load levels.
Figure 5.6 Mean ratings of confidence for the three modality combinations 
and three cognitive load conditions.

Figure 6.1 Mean participant accuracy percent scores across all modality 
combinations (AA, VV, and AV) and cognitive load levels (low, 
medium and high).

Figure 6.2 Mean participant accuracy percent scores (+SE) for pre- to post-
test sessions for the two motivation and exposure conditions.

Figure 6.3 Changes to modality combination by cognitive load interaction 
from pre- to post-test as a function of motivation and exposure.

Figure 6.4 Mean participant complexity ratings (+SE) for pre- to post-test 
sessions for the two motivation and exposure conditions.

Figure 6.5 Mean participant difficulty ratings (+SE) for pre- to post-test 
sessions for the two motivation and exposure conditions.

Figure 6.6 Mean participant confidence ratings (+SE) for pre- to post-test 
sessions for the two motivation and exposure conditions.

Figure 7.1 Depiction of the various attentional resource supply. (a) A LCL 
task resources by modality specific resources. (b) A HCL task 
requires modality-independent resources whereby (i) modality-
specific stores are ‘topped up’ (ii) modality-specific stores assume 
responsibility.
Figure 7.2 Template for letter presentation sequences in the tri-task trials. (a) A primary AV task together with a secondary visual task. Note that the primary visual task is on the RHS, while the secondary visual task is on the LHS and that letter for letter these do not onset at the same time. (a) A primary AV task together with a secondary auditory task. Note that the primary auditory task is on the LHS, while the secondary visual task is on the RHS and that letter for letter these do not onset at the same time.

Figure 7.3 Mean participant accuracy percent scores (+SE) for the three dual-task modality combination conditions across the five levels of cognitive load.

Figure 7.4 Mean accuracy performance (in %) comparisons between single- (hollow shapes) and dual- (solid shapes) task processing across the three levels of cognitive load.

Figure 7.5 Mean participant percent accuracy scores (+SE) in cross-modal (HCL) trials without (dark grey) and with (light grey) the addition of a secondary (LCL) task.

Figure 7.6 Mean ratings of task complexity (+SE) for each of the modality combinations. Single-task conditions are shown in light grey, dual-task conditions in darker grey and tri-task conditions in black.

Figure 7.7 Mean Ratings of task complexity (+SE) for the dual-task conditions across the 5 levels of cognitive load manipulated.

Figure 7.8 Mean ratings of task difficulty (+SE) for all modality combination trials. Tri-task conditions are shown in light grey, dual-task conditions in darker grey and single-task conditions in black.
Figure 7.9 Mean Ratings of task difficulty (+SE) for the dual-task conditions across the 5 levels of cognitive load manipulated.  

Figure 7.10 Mean ratings of task confidence (+SE) for all modality combination trials. Single-task trials are shown in light grey, dual-task trials in dark grey and tri-task trials in black.  

Figure 7.11 Mean ratings of trial confidence (+SE) for all the dual-task conditions across the 5 levels of cognitive load manipulated.
ABSTRACT

There are two primary attentional resource models that attempt to describe human attentional processing capabilities in multi-task situations. The unitary-resource model assumes a single store of modality-independent attentional resources which enable attentional processing for all modalities (Kahneman, 1972). It follows that performance predictions for within-modal dual-task processing (2 concurrent auditory (AA) or 2 concurrent visual (VV)) and for cross-modal dual-task processing (1 auditory concurrent with 1 visual (AV)) are similar – the two tasks compete for a share of the same resources with a consequent decrement in performance (in terms of accuracy and processing speed) irrespective of whether a cross- or within-modal task is being performed. In contrast, the multiple-resource model assumes multiple modality-specific attentional resource stores which are specific for processing in a single modality (Wickens, 1980), so it follows that this model would predict that cross-modal dual-task processing will be superior (in terms of accuracy and processing speed) to within-modal dual-task processing since in cross-modal processing the auditory and visual tasks will not compete for resources. Results from proponents of each class of model tend to support their models which could be seen as a theoretical impasse; however a review of the literature here indicates that the studies supporting each model tend to use particular methods that differ to those used in studies supporting the alternative. The proposition here is that these different methods lead to distinctly different levels of cognitive load being imposed on the individual. A model-continuum is proposed anchored at each end by these two models, with movement between these extremes being a function of the cognitive load of the task; lowest cognitive load being associated with strong multiple-resource support and high cognitive load with strong unitary-resource support. In a series of studies extraneous, intrinsic and germane cognitive load are manipulated to test the hypothesis that there will be a cross-modal advantage in conditions of low cognitive load and that as load increases the cross-modal
advantage will diminish. In all the studies a dual-task target identification paradigm was used with alphabetic letters presented in rapid serial visual presentation (RSVP) and rapid auditory presentation (RAP) sequences in within-modal (AA and VV) and cross-modal (AV) conditions with objective (accuracy and reaction time) and subjective (ratings of complexity, difficulty and confidence) measures taken. In Study One no changes to the cross-modal advantage were evident across load, however subjective data suggested the cognitive load manipulation (target-distractor similarity) was ineffective. In Study Two cognitive load was manipulated by stimulus presentation rate resulting in significant cross-modal advantages under low load conditions which diminished as load increased. In Study Three the effects of motivation and task exposure were evaluated with results demonstrating that increased motivation and exposure sustained the cross-modal advantage even in high load conditions. Moreover, rated difficulty was a superior predictor of performance to rated complexity, while rated confidence was a predictor specifically in VV within-modal processing. In Study Four a hierarchical-resource model was investigated by introducing additional levels of cognitive load, and contrasting single-, dual- and tri-task performance. Results suggested that in low load conditions modality-specific resources maintain performance while in high load conditions modality-independent resources replenish these reserves. The pattern of decrement in the cross-modal advantage as a function of cognitive load suggests the Interactive Hierarchical Resource Model is a more accurate representation of human resource system than either the multiple- or unitary-resource model alone. As support for a hierarchical model means that predictions of a categorically present or absent cross-modal advantage (as in the multiple and unitary-resource models respectively) are no longer appropriate, implications with respect to performance predictions and workplace designs are discussed with an appreciation of the role of the operator’s perceptions and the load requirements of task in the degree of any cross-modal advantage.
Air safety chiefs were warned yesterday that they were courting disaster after a report revealed that two airliners, with a total of 470 passengers on board, came within a split second of a devastating collision.

In one of the worst near misses seen at Heathrow airport, a Boeing 747 jumbo jet coming in to land missed an Airbus A321 preparing for take-off by little more than 100 feet. A report from the Air Accidents Investigation Branch (AAIB) said the potentially disastrous situation had developed at a time when a controller was supervising a trainee.

The AAIB inquiry discovered that the trainee, a 28-year-old woman, was initially in charge of the airliners' movements, but that her experienced colleague intervened when it became clear the situation was "tight". However, the 35-year-old male supervisor pursued the same course of action by calling on the A321 belonging to BMI formerly known as British Midland to take off immediately. He quickly realised, however, that it would not clear the runway in time so he ordered the British Airways Boeing 747, flying in from Japan, to pull up. The fuselage of the jumbo jet with 381 passengers on board came within about 112 feet of the tail fin of the airliner on the ground, which was bound for Brussels with 89 passengers.

[It was] pointed out that AAIB investigators had reported recently that an incident at Manchester airport last September had "all the hallmarks of a controller working under stress".

- Clement, 2001

"Air traffic controllers are always under stress, working in high load environments with many tasks to be undertaken at the same time. Glancing from screen to screen, talking with this pilot and that pilot. Mistakes will be made in such environments."

- AAIB Report, 2001
Chapter 1

An Introduction to Auditory and Visual Attention
1.1 **CHAPTER OVERVIEW**

The conceptualisation of attention as a cognitive resource is a contemporary approach in modelling the capabilities and limitations of human information processing. The conceptualisation of attention as a resource, first conceived by Kahneman (1973), is predated by a number of alternative structural models of attention (e.g., Broadbent’s (1958) Filter Theory and Treisman’s (1973) Attenuation Theory). This chapter will overview firstly the older *structural* approaches to modelling attention before moving into a discussion highlighting the *capacity, or resource*, approach – an approach which will form the foundations of this thesis. The chapter will conclude by introducing two seemingly contradictory resource approaches – unitary-resource and multiple-resource models – which are forerunners in the field at present. These two models result in distinctly different productions in multi-task situations and it is this that provides initial impetus for the experiments in this thesis.

1.2 **CONCEPTUALISING THE NATURE OF ATTENTION**

At any single moment in time, most, if not all, of an individual’s senses are confronted with a considerably large number of stimuli from the external and internal environments (Alain & Izenberg, 2003; Spence, 2001). However, processing every piece of available information is rarely achieved, as the human brain – as an information processor – may become overburdened with demands (Driver & Spence, 1998; Kahneman, 1973; Spence, 2001; Wickens, 2002) due to the complexity of the everyday environment (Gaver, 1993). In situations where more information is available than the system is able to process in a given time frame – cognitive capacity becomes ‘overloaded’ and is susceptible to errors. This coincides with a general decline in performance (Luck, Woodman & Vogel, 2000; Wickens, 2007). It is the role of the attentional system to filter incoming information and facilitate
optimum cognitive performance (Wickens, 2002) characterised by accurate and timely task execution.

To achieve this, the attentional system selects information for in-depth processing (Kahneman, 1973; Wickens, 2002) dependent upon: (1) the relative importance, or salience, of information (for instance, information signalling immediate potential danger (e.g., “incident ahead, proceed with caution”) will be prioritised over information indicating a future state of affairs (e.g., “special event, parking restrictions commencing in 10 days”)); and, (2) the amount and type of information to be processed (e.g., how many like tasks are completing at the same time) (Driver, & Spence, 1998). The latter are the most important – for should the environment provide many stimuli which need to be processed to achieve an optimal outcome, it is necessary to understand which of these will (or will not) receive additional processing. Determining which combinations of stimuli will be processed allows predictions of the individual’s optimal performance. The prediction of optimal performance is based on the knowledge of which, and how many, items may be processed simultaneously and this in turn depends on the structure of the attentional network of the individual (Alais, Morrone & Burr, 2006; Wickens, 2007).

Attention has been conceptualised in two ways – as a selective filter actively choosing some stimuli for extensive processing at the expense of others, and as a mental resource which can be allocated for task processing until it is expended. Both of these will be examined in 1.3. However, it is necessary to understand ‘attention’ more informally before models of attention can be appreciated. William James (1890) indicated that attention was necessary to prioritise the available stimuli for more in-depth examination and analysis, as processing of all stimuli would be overwhelming, and in many cases, unnecessary. James appeared to anticipate that attention, or more appropriately focalization, was a conscious
process which an individual engaged to maintain the efficient management of perceptual processing:

“Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, or consciousness are of its essence, It implied withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction”

- William James, 1890, pp. 403-404

William James’ conceptualisation of focalisation is analogous to the role of attention in both selection and resource models – to maintain optimal information processing by limiting what can be processed at any point thereby avoiding situations of cognitive overload where too many tasks are being undertaken concurrently (Luck et al., 2000). This is what is now referred to as voluntary or endogenous attention. Endogenous attention is under the individual’s own volition and can be investigated by asking an individual to actively process a given array of stimuli. For instance, sustained attention is engaged when reading a text to process all the text and encode it in memory in a meaningful way for later recall. In contrast exogenous attention is responsible for the ‘grabbing’ of attention (Spence, 2001), as in the startle response to a loud noise. This thesis focuses on the former – endogenous attention – as a crucial determinant of cognitive performance is the ability to maintain focus on a given activity and to avoid situations where voluntary attention is distracted by alternative stimuli.
Accordingly, the use of the word ‘attention’ in this thesis refers to endogenous attention, unless otherwise specified.

In summary, it appears that the role of the (endogenous) attentional system is to optimise an individual’s performance by actively limiting the stimuli which are processed during any given period. As such, attention can be conceptualised as both a source of efficiency, by filtering the information for processing, and as a limitation, in that important information may be overlooked if the system is operating near full capacity.

1.3 ATTENTION: AS A ‘SELECTIVE FILTER’ OR AS A ‘COGNITIVE RESOURCE’

The way in which the attentional system operates to limit the number of items processed at any given time is the question of interest for this thesis. Previous approaches have conceptualised attention as a filter which acted to select specific items from the environment for further, in-depth, processing. In such theories, for instance, filter theory (Broadbent, 1958) and attenuation theory (Treisman, 1973) - collectively referred to as structural selection (Parasuraman & Davies, 1984) or bottleneck approaches (Wickens, 2002), it was proposed that information processing begins as a parallel and unlimited process. During this unlimited processing stage, stimuli impinge on the individuals’ senses before specific ones are selected, by the attentional filter, for further processing which includes the involvement of working memory (Driver & Spence, 1998; Wickens, 2002). The parallel processing capabilities of the system cease once items reach the attentional filter, and instead serial processing ensues. This means that only one, or a few, of several coincident stimuli can be processed at any given time and it is this feature which discriminates the selection perspective from later alternatives – including resource model standpoints.

The limitations of such ‘selection’ views of attention became apparent in the 1970s and 1980s as researchers determined that individuals were able to process large numbers of
contemporaneous discrete items, sometimes in fine detail. For example, Cherry (1956) showed that participants were able to shadow (or repeat) a primary speech signal and were also able to report features of a secondary speech signal (i.e., language switching in the speech signal) even though they had been instructed that the secondary speech signal was irrelevant and should be ignored. Similarly, Moray (1959) demonstrated that participants were able to detect speaker gender changes, and switches between forward and backward speech of a secondary signal while shadowing a primary speech stimulus, while Treisman (1960) demonstrated that participants may often scramble the content of a primary and secondary speech stimuli and report a passage with amalgamated elements of both the primary and secondary stimuli (see Figure 1.1). Together, such findings demonstrate that the content of a supposedly unattended channel is focussed upon, albeit to a significantly lower extent than the attended channel.

![Diagram of a person with headphones showing speech stimuli and shadowed responses.]

**Figure 1.1.** Example of speech stimuli presented binaurally to participants via headphones and the shadowed speech sentences produced by the participants in response to speech stimuli presented simultaneously.
To account for these experimental outcomes, seemingly inconsistent with limitations of selection and filter theories (specifically, that more than one task could be completed at the same time), Moray (1967) suggested that attention might be viewed as a limited capacity cognitive resource in which attention is a limited cognitive resource or mental energy that can be allocated where necessary to enable the processing of the available stimuli (Kahneman, 1973). More specifically, each and every task requires a level of attentional resources to enable task completion and the system allocates the available attentional resources on a needs basis until the resources available within the system are depleted. This conceptualisation includes two important features which all resource models contend: that attention is both limited and allocatable.

Conceiving attention as a limited resource means that a number of tasks can be conceivably processed at the same time so long as the total attentional requirements of all the tasks does not exceed the attentional resources available within the system. Once all of the available resources are expended no further perceptual and cognitive processing is possible without performance decrements in one or more of the tasks (Kahneman, 1973). Many such performance decrements have been demonstrated in the literature, including: decreased driving performance in simulated driving tasks (Matthews, Sparkes, & Bygrave, 1996); decreased interaction in group interaction studies (Kellogg, 1988); and decreased performance in writing and arithmetic tasks (Lemaire, Abdi, & Fayol, 1996); compared to the same tasks performed under single task conditions. These results indicate that errors were not the result of individual incompetence (as performance in single-task conditions was high), but rather, a consequence of performing them simultaneously. Such outcomes have been used to support the claim that attentional resources are finite and can be used to facilitate optimum performance when tasks demands are low, and under situations of underload (Wickens, 2002).
Although the notion of limited resources is well established, the allocatable nature of these resources is less so. While proponents of resource models ascribe to the view that attention resources can be allocated during task completion, it remains uncertain how this distributional system is structured, and therefore, when, in any given dual-task scenario, the resources may be exhausted. Presently, no standardised resource model is recognised as representing the structure of the attentional resource network in human information processing.

To determine how the system is structured and therefore how resources are allocated, a review of situations in which the system is placed into overload is necessary. According to Wickens (2002), the structure of the attentional system can be inferred from the performance limitations evident during perceptual and cognitive processing. It is possible to determine the nature and the number of attentional resource stores of the system by monitoring performance on multiple tasks as the demands on the system are manipulated by the number of concurrent tasks being completed, or, the complexity or difficulty of the task(s). For this reason, dual-task experimental paradigms are regularly used in attentional resourcing research (Brünken, Plass & Leutner, 2003; Kahneman & Treisman, 1984; Wickens, 2002).

**Dual Task Experimental Designs:** The dual-task method involves a participant engaging in two simultaneous tasks. In some situations attentional capacity and limitations are inferred from the number and the nature of secondary tasks in which an individual can engage, while maintaining performance on a primary task equivalent to that primary task being undertaken in isolation (Brünken et al., 2003). This method enables researchers to evaluate the attentional demands of a single primary ‘selective’ stimulus/task by measuring whether any residual resources remain available to the system to process secondary, often ‘irrelevant’ stimulus/task.
The *Inattentional Blindness* (IB) paradigm (Sinnett, Costa & Soto-Faraco, 2006) is one such experimental approach which evaluates residual auditory and visual attentional resources. In these designs participants are required to engage in a primary task while a secondary task is presented simultaneously. However, this secondary task is explained to participants as being irrelevant and to concentrate only on the primary task. Such a selective attention design, therefore, estimates resource capabilities by evaluating primary task performance and then determining the knowledge the individual has about the apparent ‘irrelevant’ task. The assumption is that, in cases where the primary task requires fewer resources than available in the system, these residual resources will be used to scan any other available stimuli. In contrast, when system resources are expended, no processing of the irrelevant task will be achieved. The limitation of such residual resource measures is that, in most real world tasks, there is no easily identifiable primary task – instead individuals are faced with many individual tasks all demanding attentional resources. For this reason, the divided attention dual-task paradigm was introduced.

Divided attention tasks, as the name suggests, require individuals to engage in more than one task at a single moment in time. In contrast to the selective attention paradigm, neither task is identified as more critical than the other (e.g., simulated driving tasks where individuals must monitor all possible hazards, whilst maintaining appropriate speed, and possibly engaging in in-car conversation) (Wickens, 2002). The *Attentional Blink* (AB) paradigm (Arnell & Jolicœur, 1999) is one such method of evaluating divided attention, where participants are required to monitor the presentation of two target stimuli presented in quick succession (e.g., monitoring a string of alphabetic letters for the target letter “X” presented on two occasions). The offset of the two X’s would be manipulated from immediate (e.g., 90ms) to delayed (e.g., 600ms). The resource capacity of the system is, therefore, ascertained by establishing whether the two X’s are both identified or whether the
second one is overlooked by the participant (an attentional blink). Similar to the IB paradigm, the AB approach may present difficulties with ecological validity as using such a method limits performance to a dichotomous response (i.e., dual target identification versus single target identification). This means that individuals can either achieve dual-target identification or they cannot, but no record of declining ability is available. In terms of cognitive processing, it is clear that performance does not suddenly extinguish – but rather it tends to diminish over time.

To overcome this limitation, many applied domains concentrate on measures of accuracy and Reaction Time (RT) to evaluate performance changes. These measures allow the determination of the point (or condition) at which performance begins to deteriorate, rather than the point at which the task cannot be completed. For instance, using the AB paradigm, a researcher may determine only when an individual is no longer able to detect the two targets. However, using accuracy and RT measures allows the determination of the point at which two-target detection first begins to fail (e.g., when RTs begin to lengthen). Being able to predict not only failed performance, but also sub-optimal performance, is necessary in many applied domains. For instance, in the case of a near miss between aircraft at Heathrow Airport (see pp. xxiii), the supervisor’s performance had begun to show signs of deterioration (prolonged time frames to issue instructions to pilots, or changing directives) long before the near miss between the Boeing 747 and Airbus A321 occurred.

The value of the dual-task (RT and accuracy) experimental design is most evident when reviewing multi-modal stimuli and multi-modal processing. Until the 1990s, much attention research was focussed on single-modality processing – predominantly within the visual system (Spence & Driver, 1996). In this case, the dual task design enabled researchers to assess visual processing capabilities as additional stimuli were introduced and a linear summation could be used to predict the impact of adding further stimuli. However, the results
of a number of studies (Moray, 1959; Webster & Thompson, 1954; Wickens, 1980) in the 15-20 years prior – not specifically within the attentional resourcing domain – began to suggest that such a simple summation may not adequately explain human information processing, particularly when the multiple stimuli confront different modalities.

Cherry (1956) reported that participants could report features of a secondary speech stimulus while maintaining high performance on a primary speech shadowing task. However, those secondary speech features reported by participants were particularly rudimentary (e.g., changes in the gender of the voice, changes between forward-speech and backward-speech). In contrast, Parkes and Coleman (1990) demonstrated that participants who were performing a primary speech shadowing task, not unlike that of Cherry, could also complete quite difficult visual tasks (e.g., speed maintenance and lane keeping in simulated driving experiments). These results suggest that simultaneously processing two auditory stimuli (within modalities) (as in the Cherry experimental design) may not be analogous to simultaneously processing an auditory and a visual stimulus (cross modalities) as was the case in the Parkes and Coleman study. Accordingly, a strict summation of the attentional requirements of each task may not be appropriate – particularly when evaluating multi-modal stimuli.

Research prior to the late 1990s was concerned with single modality perception yet rarely is human information processing so categorical (Driver & Spence, 1998). More recently, multi-modal attentional research has grown as will be evident in Chapter Two. This thesis is primarily concerned with this more recent divergence into multi-modal attentional resourcing.

*Studying Auditory and Visual Multi-Modal Attention:* Multi-modal interactions and the attentional resourcing in such dual-task scenarios have been evaluated in a number of contexts to date: auditory and visual (Alais et al., 2006; Spence & Driver, 1996; Wickens,
2007); visual and tactile (Kennett, Eimer, Spence & Driver, 2001; Heller, Calcaterra, Green & Brown, 1999); auditory and tactile (Caclin, Soto-Faraco, Kingstone & Spence, 2002; Hötting, Rösler & Röder, 2002); and, auditory, tactile and visual (Eimer, 2001; Spence, Nicholls, Gillespie & Driver, 1998). The present studies focus on auditory and visual stimuli for a number of reasons.

Firstly, there is a large body of research across different laboratories, using different experimental designs, and across a number of different domains that have evaluated these two modalities. The result is a good deal of comparative data. Secondly, a number of studies have reported interactions between the auditory and visual modalities. For instance, in the McGurk Effect, McGurk and Macdonald (1976) demonstrated that when participants were presented simultaneously with the auditory vocalisation /ba/ and the visual lip movements of /ga/ they reported the perception of /da/ - a fusion of the auditory /ba/ and the visual /ga/. Moreover, this illusion is involuntary – participants continue to report the auditory-visual stimulus as /da/ even when they are aware of the auditory and visual components. The interdependence of auditory visual speech may come as little surprise due to the inherent auditory-visual nature of speech. However, the interaction between visual and auditory modalities also occurs in a number of non-speech illusions. One of these is the sound induced illusory flash effect (Shams, Kamitani, & Shimojo, 2002), or Shams Illusion, in which a single flash of light may be perceived as two flashes when accompanied by two successive auditory tones.

Despite the attention the multi-modal resourcing has received recently (Spence & Driver, 1996; Wickens, 2007) there remains no consensus in the field regarding the resourcing of single and multiple tasks. Indeed there are two seemingly incongruent frameworks advocated by difference researchers within the field.

The Multiple-Resource and Unitary-Resource Models are two attempts to describe the way in which humans’ limited attentional resources are allocated during perceptual
processing. A model of attentional resourcing is essential so that predictions of human
behaviour and performance can be made and an understanding of human processing
limitations can be determined. However, the predictions of the Unitary- and Multiple-
Resources models are significantly different, if not contradictory. In the introductory chapters
the course of such contradictions will be explored and in the following chapters studies will
be conducted which attempt to elucidate the conditions under which these contradictions
might occur.

While the Unitary-Resource Model is able to predict overall multi-task performance
based on the number and difficulty of concurrent tasks to be completed, the Multiple-
Resource Model instead predicts that overall multi-task performance is based on the number,
difficulty and the type of concurrent tasks to be performed. Each model and its dual-task
processing predictions will be discussed below.

1.3.1 Unitary-Resource Models

Unitary resource models have existed under a number of pseudonyms across many
years and researchers and have included: supramodal, amodal, uni-modal and single-capacity
resource models. Despite frequent relabelling, these models all conceptualise the attentional
resource structure as consisting of a single store, reservoir or pool (Knowles, 1963) of
resources which can be shared for all perceptual and cognitive activities (Kahneman, 1973).
A single resource limitation model essentially recognises that all tasks that an individual
undertakes will draw from the same attention resource pool, irrespective of the similarity of
these tasks. That is, a carpenter simultaneously listening to his/her iPod (auditory task),
reviewing plans for the table that he/she is crafting (visual task) and running his/her hand
along the recently sanded table top checking for imperfections (tactile task) are all tasks that
would share the single reservoir of attentional resources. Therefore, by summing the
requirements of the each of the individual tasks, it is possible to derive the combined
cognitive demands placed on the system. So long as this level of demand remains lower than
the attention resource capacity of the system, performance will remain high (limited only by
the ability of the operator). However, if the demand is greater than the system capacity -
performance in both domains, auditory and visual, will decline. Therefore, in the case of a
single reservoir system, a single resource limitation will be evident. That is, once the system
is overloaded, individuals should be unable to engage in any additional tasks – irrespective of
the similarity to concurrent tasks.

Support for unitary-resource models is evident in a number of studies during the past
half century (Arnell, 2006; Arnell & Jenkins, 2004; Arnell & Jolicœur, 1995, Arnell &
Farah, Wong, Monheit & Morrow, 1989; Shulman & Hsieh, 1995; Vohn, Fimm, Weber,
Schnitker, Thron, Spijkers et al., 2007; Ward, 1994). For example, Arnell and Jenkins
required participants to identify target letters using within-modal (two auditory tasks or 2
visual tasks) and cross-modal (an auditory and a visual task) stimuli and found that similar
errors were made irrespective of the type of trial. The authors concluded that the inability to
perform any form of secondary task must result from a single resource limitation - an
outcome consistent with a unitary-resource account of attentional resourcing. Similar within-
modal verses cross-modal outcomes have been evident across a number of disciplines and
using a number of experimental paradigms (Arnell; Bonnel & Hafter; Eimer & van Velzen;
Shulman & Hsieh; Ward).

To appreciate processing limitations in a unitary-resource structure, and to later
compare these to the processing limitations in a multiple-resource structure (see Chapter 2), a
simple illustration is offered in Figure 1.2. According to the unitary-resource structure, the
summation of individual demands into a system will determine the performance potential of
Performance on all tasks should remain high as total system demands are less than total system capacity (demands <10 units).

Total system demands are greater than total system capacity (demands >10 units) and as such performance on both the auditory and visual tasks will diminish.

Total system demands are greater than total system capacity (demands >10 units) and as such performance on both the auditory and visual tasks will diminish.

Performance on all tasks should remain high as total system demands are less than total system capacity (demands <10 units).

Figure 1.2. Unitary—Resource model performance predictions
the individual. Where the total demands placed on the system remain less than the cognitive capacity of the system, performance on all tasks should remain high (see Figure 1.2a). However, should the difficulty of any of the tasks (see Figure 1.2b-d) increase above system capacity, the individual will be unable to perform at an optimum level and performance will decrease on all of the tasks.

While a great deal of research supports the unitary-resource model, there remains doubt as to whether a strict summation of individual tasks requirements is sufficiently sensitive to predict human performance. These doubts were first expressed in studies by Allport, Antonis and Reynolds (1972), Shaffer (1975), and Wickens (1976). For instance, while numerous authors had demonstrated performance decrements when participants tried to perform two within-modal tasks (Kahneman, 1972), Allport et al. demonstrated that participants could sight-read music scores and engage in an auditory shadowing task (a cross-modal task) equally well as if they were performing each task in isolation. The unitary-resource account would account for such an outcome by positing that each task required minimal attentional resources and therefore the system demands were less than system capacity. However, more recent studies have suggested an alternative account as to why such a cross-modal outcome, or cross-modal advantage (Rodway, 2005), may have occurred.

1.3.2 Multiple-Resource Models

In response to Allport et al. (1972), Shaffer (1975), Wickens (1976) and others (Parkes & Coleman, 1990; Tannen, Nelson, Bolia, Warm & Dember, 2004; Webster & Thompson, 1954 to name a few), alternative models of attentional allocation were offered to explain why cross-modal performance appeared to be superior to within-modal performance. Specifically, these models addressed why individuals appeared more capable of performing a visual task (e.g., colour change judgements) together with an auditory task (e.g., monitoring a
musical piece for the presence of a three note string) than performing two auditory (e.g.,
monitoring a musical piece and listening for the introduction of static) or two visual (e.g.,
colour change judgements, and reading) tasks.

The multiple-resource account postulates the presence of a number of separate
resource pools which are modality specific (Wickens, 2007). Tasks requiring processing by
different modalities will draw from discrete attentional resource pools and, therefore, should
not interfere with processing in another modality as they are not competing for resources
from that same store. Figure 1.3 highlights the different attentional resource allocations for
each of these three tasks under both the unitary- and multiple-resource models.

Support for a multiple-resource model is evident in numerous studies, across a
number of domains, and using a number of experimental methods over the past 30 years. For
example, many of Wickens’ recent papers (Horrey & Wickens, 2006; Wickens, 2002;
Wickens, 2007; Wickens, 2008; Wickens & Liu, 1988) have demonstrated a significant
cross-modal advantage in simulated driving tasks where participants are required to carry out
a conversation with another person via a number of different methods (e.g., hand-free, hand-
held, windscreen mounted). Wickens’ work is often considered to be of an applied nature and
experimental control has been argued to be questionable using such high fidelity designs
(Strayer, Drews & Johnston, 2003). However, more recently similar results have been
reported in tightly-controlled laboratory studies (see Burr & Alais, 2006). Combined the
results from both applied and basic research domains provide reason to consider the multiple-
resource approach is both valid and reliable.

Returning to the earlier illustration introduced in Figure 1.2 the same scenarios will
now be considered in terms of multiple-resource predictions and illustrated in Figure 1.3.
Given the maximum possible processing of 10 cognitive demand units at a given time in the
multiple resource model this full system capacity is divided up between perceptual modalities
Performance on all tasks should remain high as no modality-specific capacity limits have been exceeded, and no reservoir is in a situation of overload.

Despite the total system demands exceeding system capacity (demands >10 units), performance on the auditory tasks should remain high as no auditory capacity limits have been reached (demands < 5 units), however, performance on the visual task will degrade as the visual modality is placed into a system of overload (demands >5 units).

Despite the total system demands exceeding system capacity (demands >10 units), performance will diminish only for the auditory tasks as only the auditory modality demands exceed auditory capacity (auditory > 5 units; visual <5 units).

Despite system demands being less than total system capacity (<10 units), performance will still diminish on the auditory tasks as the auditory modality demands exceed auditory capacity (auditory demand > 5 units). Performance on the visual task should be maintained as visual demands are less than the visual modality capacity.

Figure 1.3. Multiple—Resource model performance predictions
and so in this example the auditory modality is limited to 5 cognitive demand units and the visual modality to a separate 5 demand units. As depicted in Figure 1.3 performance on all tasks should remain high if the total demands placed on each sub-system remains less than the cognitive capacity of each sub-system (i.e., 5 demand units) (see Figure 1.3a). However, should the difficulty of either of the auditory tasks increase, performance will decrease for the auditory tasks but performance will be sustained on the visual task (see Figure 1.3b-d).

In summary, multiple-resource models allow for a separate pool of attentional resources for each modality. Each resource will only support processing within the relevant modality – no sharing of resources is possible; once modality-specific resources are depleted, no additional task of the same kind can be attempted without performance deficits occurring. Consistent with the unitary-resource model, there is a great deal of research support for the multiple-resources model (Alais & Burr, 2003; Alais et al., 2006, Burr & Alais, 2006; Duncan, Martens & Ward, 1997; Grady, Van Meter, Maisog, Pietrini, Krasusaki & Rauschecker, 1997; Horrey & Wickens, 2006; Mondor & Amirault, 1998; Potter, Chun, Banks & Muckenhoupt, 1998; Soto-Faraco, Morein-Zamir & Kingstone, 2005; Spence & Driver, 1994; Woodruff, Benson, Bandettini, Kwong, Howard, Telavage et al., 1996). Despite this evidence, there remains a question concerning the independence of each modality-specific resource. For instance, when approaching road-works or an accident, evidence suggests that drivers tend to turn down the volume of the radio, pause during a conversation with another person, or slow down their speed to maximise concentration (Takayama & Nass, 2008). Such outcomes are not consistent with the predictions of multiple-resource models, for these models, if considered in isolation of other perceptual and cognitive processes which may independently draw on other attentional resources the premise behind conceptions of multiple-resource advocates is that near perfect performance on cross-modal tasks should be maintained.
Given that the unitary- and multiple-resource models have distinctly different performance predictions (see Table 1.1 for a summary of the two models), it is difficult to understand why both are regularly supported in the research. Yet there remains evidence for both models. In Chapter 2, research studies supporting each model will be reviewed, with particular emphasis on the experimental designs, methods and stimuli in an attempt to understand the source of different predictions by and continued support of the two models.
Table 1.1

*Contrasting the Unitary- and Multiple-Resource Models: Their Structures, Resource Allocation Networks and their Performance Predictions*

<table>
<thead>
<tr>
<th>Model Features of Interest</th>
<th>Unitary-Resource Model</th>
<th>Multiple-Resource Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number and Type of Resource Stores</td>
<td>A single modality resource reservoir from which all processing modalities draw from during stimuli processing.</td>
<td>Multiple resource reservoirs which function independently of one another. Each processing modality thus draws from its own unique store of resources.</td>
</tr>
<tr>
<td>Pictorial Representation of Model Structure</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Within-Modality Performance Predictions</td>
<td>When two auditory or two visual tasks are performed concurrently competition for attentional resources <em>will be</em> evident as these within-modal tasks will be maintained by a modality-independent resource reserve.</td>
<td>When two auditory or two visual tasks are performed concurrently competition for attentional resources <em>will be</em> evident as these within-modal tasks will be maintained by a modality-specific resource reserve.</td>
</tr>
<tr>
<td>Cross-Modality Performance Predictions</td>
<td>When an auditory and a visual task are performed concurrently competition for resources <em>will be</em> equitable to within-modality performance as the two tasks are resourced by the same modality independent resource reserve.</td>
<td>When an auditory and a visual task are performed concurrently competition for resources <em>will not be</em> evident as each task is maintained by a discrete modality-specific resource.</td>
</tr>
</tbody>
</table>
THE GREAT DIVIDE

A review of the Literature Contributing to the Resource Divide: Why are Two Conflicting Models Supported So Commonly in the Literature?
2.1 CHAPTER OVERVIEW

The resource debate – or the great divide – has been growing for over 20 years. Different researchers continue to find experimental results conforming to both the unitary- and multiple-resource models and rarely questioning such inconsistent outcomes. Where attention has been directed towards explaining the difference, variations in experimental designs have normally been identified as points of difference (Potter et al., 1998). Differing experimental designs have been implicated in previous research as the basis of a divide between early and late models of attentional processing (Lavie, 1995). In fact, more recent research has revealed that rather than one or the other of these models being discredited, both models can explain aspects human attentional processes. However, they represent different aspects that were targeted as a result of differences in experimental design. Potentially, the apparent divide between the evidence supporting multiple- and unitary-resource approaches is a product of a similar difference in experimental methodologies.

In this chapter research representing both classic multiple- and unitary-resource models will be discussed, leading to more contemporary hybrid models (e.g., separate-but-linked and hierarchical). These models will be reviewed and categorised according to the features of their experimental designs as a means of explaining differences associated with the outcomes.

2.2 INTRODUCING THE GREAT DIVIDE: ONE PHENOMENON, TWO CONTRADICTORY MODELS

The multiple- and unitary-resource models described in Chapter 1 make cross-modal predictions that are not only contradictory but, if the models are taken in their purest forms, are incompatible with one another. Nevertheless, a number of published studies across a number of disciplines have found support for unitary-resource accounts of attention, and a
comparable number of studies have also demonstrated results which support a multiple resources account (Sinnett et al., 2006).

These studies are of interest for three reasons. Firstly, the literature supporting the multiple- or the unitary-resource accounts is substantial. In the past 10 years, over 1,000 articles have been published which comment on, and contribute to, the attention model debate (Alais et al., 2006; Arnell, 2006; Burr & Alais, 2006; Sinnett et al., 2006; Wickens, 2007). Secondly, the research articles belong to no single research domain but, instead, range in scope from psychology, to neurology, to computational modelling and through to philosophy (Lloyd, Merat, McGlone & Spence, 2006). Finally, no set of results has provided decisive support for one model over the other.

These three factors suggest that accepting one of these resource models to the exclusion of the other may be both premature and unjustified. Rather, in recent articles, researchers have begun to recognise that results which support one model do not logically imply that the alternative model has been falsified (Burr & Alais, 2006). Many authors are now less concerned with disproving the alternative model, and more concerned with identifying factors that may contribute to these inconsistent outcomes (Alais et al., 2006; Burr & Alais, 2006; Driver & Spence, 2004; Eimer & van Velzen, 2002; Schmitt, Postma & de Haan, 2000; Schwartz, Robert-Ribes & Escudier, 1998; Soto-Faraco et al., 2005; Wickens, 2002).

Many other cognitive psychology fields have been faced with similar theoretical divides (Lavie, 1995). Indeed, even within the attentional modelling field (Kahneman & Treisman, 1984; Lavie, 1995) the present unitary- versus multiple-resource divide is reminiscent of the older ‘early versus late’ selection dispute, which was concerned with whether the attentional selection of stimuli occurred during the preliminary stages of
perception (after analysis of basic physical features), or at a post perceptual stage in order to facilitate a response (Lavie, 1995).

The similarities between the present ‘resource divide’ and the older, early versus late selection dispute extends beyond the role of experimental designs and paradigm shifts (Kahneman & Treisman, 1984). Lavie (1995) and Kahneman and Treisman observed the gradual shift of experimental designs from more traditional ‘situation-based’ tasks to more contemporary ‘laboratory-based’ tasks within the attention research field over the preceding 10-15 years. Specifically, they commented that the early methods were relatively more ‘complex’, ‘effortful’ and ‘difficult’ for participants and were characterised by experimental designs using conversational auditory stimuli (Cherry’s (1953) Cocktail Party experiments) and/or decision making (Webster & Thompson’s (1954) Aviation tracking experiments). In contrast, they noted that the more recent approaches instead incorporated greater experimental control at the expense of situational reality (e.g., changing colour patches and altering tone pitch) requiring less ‘effort’ and ‘concentration’ by participants as they engaged in ‘simpler’ or ‘low complexity’ tasks (Lavie, 1995).

More recent changes in experimental paradigms within the attentional resourcing literature appear comparable to the changes documented by Kahneman and Treisman (1984) and Lavie (1995) above in relation to the early versus late dispute. For instance, within attentional resourcing literature, the early experimental designs can be described as ‘complex’ in nature; participants were required to complete comprehension, understanding and decision making tasks with meaningful stimuli, thereby investing significant mental effort (Kahneman, 1972). It was during this time that unitary-resource model support dominated the literature (Driver & Spence, 1998). However, as greater levels of experimental control were implemented, methods involving simple stimuli and non-complex tasks were introduced (e.g., visual colour change and auditory pitch discrimination tasks) – tasks that could generally be
described as requiring little mental effort. Coincident with this change, multiple-resource accounts of attentional allocation came to the fore (Wickens, 2007). The remainder of this chapter provides a review of evidence for multiple- and unitary-resources with the expectation that similar insights to those gained from the early versus late selection debate may be obtained.

2.3 EXPLAINING THE DIVIDE: EVALUATING THE METHODS OF PAST STUDIES

Given the large number of published articles contributing to the multiple- versus unitary-resource debate, the present review is selective with due regard for different paradigms (e.g., behavioural, neurophysiological, electrophysiological and clinical domains). For the purposes of this review, a subset of 34 papers (those most regularly cited in the literature in the past 5 years) was examined representing advocates of both multiple- and unitary-resourcing propositions. The review will highlight ‘strong advocates’ and others ‘weak advocates’ of each of these models. Strong advocates are those that provide support for one model against the other, whereas weak advocates are those that maintain their position of support for a model despite results that do not refute the predictions of the alternative model (e.g., Burr & Alais, 2006). Following a review of the studies supporting the multiple- then the unitary-resource models, there is a review of hybrid models e.g., the separate-but-linked model (Spence & Driver, 1994) and the hierarchical model (Posner, 1980).

2.3.1 Experimental Methods of Multiple-Resource Studies

Prior to the formalisation of the multiple-resource model, which posits that individual resource capacities exist for each processing modality (and other dimensions also, like, stages, responses and codes) (Wickens, 1980), various studies had demonstrated that cross-modal rather than within-modal stimulus presentation facilitated greater performance (Glass,
While the Wickens model was originally conceived as a workload model, it has been incorporated into accounts of attentional resourcing outside of the applied domain. The model was developed by Wickens to explain dual-task processing costs in predominantly simulated driving tasks where distractions like mobile (cell) phones or other passenger conversations were engaged simultaneously. As a workload model, it was able to predict relative performance, dependent upon the type of stimuli presented (same modality versus different modality), the stages of processing (perception, cognition and response), the required responses (manual spatial or vocal verbal) and the codes (spatial or linguistic). Essentially, the model proposed that maximum performance was possible when two tasks required processing in different areas of the model. For instance, visual driving, paired with a motor driving movement response was performed accurately in the company of an auditory modality and verbal response conversation.

Potter et al. (1998) were among the first within the basic human information processing field to agree with the predictions of Wickens’ multiple-resource model (1980) – albeit at the time not identifying specifically with the model. Potter et al’s (1998) experimental approach was considerably different to that used in many of the studies conducted by Wickens’ and his colleagues (Horrey & Wickens, 2006; Liu & Wickens, 1992; Wickens, 1976; Yeh & Wickens, 1988). For instance, while Wickens predominantly investigated attention within simulated driving experimental parameters, and other simulated environments (e.g., cockpit and air traffic control), Potter et al. used more laboratory-based experimental procedures already being applied within attentional resourcing research – specifically the AB paradigm. Therefore, while Wickens was concerned with gross performance measures (e.g., lane keeping, speed, obstruction detection, obedience to signposting, stopping distance), the latter was concerned with more specific performance measures (e.g., reaction times to target presentation, accuracy of target detection, recall of
target features). The use of converging methods to produce the same results is taken in
behavioural science as strong evidence for the veracity of the results and in principle,
provides strong support for the multiple-resource model. In this regard, an examination of the
similarities and differences in these two methodological approaches will be informative.

Potter et al. (1998) asked participants to detect an alphabetic letter within a stream of
auditorily or visually presented numerals (and visa versa), where participants sequentially
monitored either two auditory or two visual or one auditory and one visual alpha-numerical
streams. Participants were asked to complete a pattern recognition task by remembering the
odd symbol out in the alpha-numerical stream (i.e., a letter among numbers, or a number
among letters) which was presented at a relatively slow presentation rate (Arnell & Larson,
2002). This design is said to require little mental effort\(^1\) by the participant as the discrepant
symbol ‘pop[s] out’ for them (Arnell & Jolicœur, 1999). Moreover, with significant time
between the presentation of each symbol within the stream, participants have ample time to
detect and encode (for later recall) the target symbol. Therefore, in finding support for a
multiple-resource account Potter et al. appeared to employ a relatively undemanding task in
which participants were required to invest relatively little mental effort.

In contrast, the tasks used by Liu and Wickens (1992) and Yeh and Wickens (1988)
appeared to require considerable effort, as participants engaged in simulated driving tasks –
tasks which appear to be somewhat more taxing or more difficult than the symbol detection
task (Potter et al., 1998). However, the driving tasks used in many of Wickens’ experiments
can be considered somewhat automatised (that is, they could be performed without apparent
consideration, thought or consciousness), thus requiring reduced mental effort; especially
given that Wickens’ participants appeared to be experienced drivers. Therefore, Potter et al.

\(^1\) Mental effort is a term regularly used in the literature to describe the cognitive exertion experienced by an
individual as they engage in any type of task (Wickens, 2002). Mental effort is a subjective account of the
demands a particular task imposes on an individual’s limited cognitive resources. For the purposes of this thesis
the term will be used to express the cognitive experience the average participant would observe/report.
and Wickens’ (Liu & Wickens, 1992; Wickens, 1976; Yeh & Wickens, 1988) tasks may be more similar than at first glance as while different in appearance they both require low levels of mental effort.

This claim – that it is the relative mental effort engendered by a task, rather than the features of the task itself which is responsible for the differing experimental outcomes contributing to the resource divide – is further supported by a study by Horrey and Wickens (2006) who conducted a meta-analysis of 23 studies of the effects of mobile telephone use on driving ability. They found no interference between auditory and visual dual-task processing in studies where measures of lane-keeping or tracking performance were reported. However, cross-modal performance costs were evident in measures of critical road hazard avoidance and detection, differences that may have resulted from different resources being tapped by the different dependent variables (e.g., lane-keeping requires peripheral vision and hazard response requires focal vision). Over and above this it would be argued that lane-keeping and tracking are tasks that drivers always perform and are therefore more likely to be automated. In contrast, hazard avoidance occurs far less regularly when driving and more often than not is varied in form (e.g., cars changing lanes, pedestrians, cars braking suddenly) and therefore are less likely to be automated.

The proposition being offered thus far is that studies supporting multiple-resources usually use tasks that require little effort. This proposition is supported by a series of studies recently conducted by Burr and Alais (2006) who required participants to localise blobs of light presented visually and clicks of sound. In dual-task conditions, no significant deficits in cross-modal detection accuracy were found compared with significant within-modal accuracy deficits, evidence that Burr and Alais took to be consistent with a multiple-resourcing account of attention, and not with a unitary-resource account. Similarly, in Potter et al. (1998), the stimuli employed a detection task requiring participants to determine whether the target
origin was to the right or left of the probe origin. This is a straightforward localisation task presumably automatised through regular use in everyday life. Their results also showed less cross-than within-modal performance deficits, supporting the notion that multiple-resource studies employing simple and/or automatic tasks result in cross-modal benefits.

Further support for this notion can be derived from Duncan et al. (1997) in which participants were required to detect a target word (high pitched ‘nap’ or ‘nab’; low pitched ‘cot’ or ‘cod’) from a non-target syllable (always /guh/) auditory series, or to detect a target word (vertical presentation (one above the other) ‘cot’ or ‘cod’; horizontal (side-by-side) presentation ‘nap’ or ‘nab’) from a non-target ‘XXX’ visual series. Although the participants were required to monitor these four different target possibilities, the targets tended to ‘pop out’ as (1) the target stimuli were real words, while the distractors were non-words and (2) the targets were presented in a different pitch (either higher or lower) than the surrounding distractor stimuli. As such, while at first glance this task appears to require significantly more effort than those previously discussed, the ‘pop out’ feature significantly reduces the overall difficulty of the task. Therefore, Duncan et al. concluded that support for multiple-resources fits well within the argument presented in this thesis: that effortless tasks are associated with multiple-resource support.

These and other behavioural studies (Alais et al., 2003; Glass, 1980; Mondor & Amirault, 1998; Schmitt et al., 2000; Soto-Faraco et al., 2005; Spence, Ranson & Driver, 2000) support the proposition in this thesis that support for the multiple-resource models tends to be evident when non-taxing or effortless tasks are used. To determine the generality of this argument, further studies outside the behavioural realm should be considered to establish whether this model-method association can be sustained outside of the behavioural experiments so far reported. For instance, in recent years, neurophysiological studies have also begun to be conducted in the resource field. The addition of neurophysiological evidence
has not disambiguated the divide within the attentional resource field. Rather, it has variously corroborated the predictions of both multiple- and unitary-resource models.

In terms of support for the multiple-resource model found in such neuroimaging studies (using such methods as fMRI, PET, ERP), a number of studies have concluded that cross-modal processing involves some distinct areas of the brain not involved in either visual or auditory within-modal processing. For example, in an fMRI study Woodruff et al. (1996) asked participants to perform the relatively non-taxing and effortless task of detecting a target number (eight(8)) from an array of distractor numbers (one through to seven (1-7) and nine (9)) across both auditory and visual modalities. Importantly, the choice of this experimental design was such that task-switching was of no concern when interpreting the findings (Potter et al., 1998). Woodruff et al. demonstrated in this study significant autonomy between the auditory and visual modalities by evaluating the areas of the brain involved in auditory-only, visual-only and auditory-visual dual task processing. The authors concluded that there was support for the multiple-resource model. Similarly, a number of additional neurophysiological studies have supported the claim of independence between the auditory and visual attentional systems (Grady et al., 1997; Jancke, Mirzazade & Shah, 1999; Luck, Chelazzi, Hillyard & Desimono, 1997; Motter, 1993; Sommers, Dale, Seiffert & Tootell, 1999) and all these have used experimental methods which could be categorised as ‘easy’, ‘simple’, ‘effortless’ and ‘non-taxing’ whereby the individual appraises the situation as requiring few resources and therefore well within their coping resources (Matthews, 2001).

In summary, a significant proportion of studies reviewed here demonstrate quite strong support for multiple-resources with little evidence of cross-modal interference. One exception was reported in a meta-analysis (Horrey & Wickens, 2006) which indicated that some level of cross-modal interference may appear when more effortful or taxing dependent variables were used (e.g., hazard avoidance compared to lane-keeping). The research
evaluated here, irrespective of whether the results could be viewed as weak or strong, suggests that relatively simple or non-complex experimental methods may be especially associated with results that support modality-specific attentional resource pools. However, despite the wealth of research advocating a multiple-resource position, there exists an equally persuasive argument for the unitary-resource position to which we now turn, with a particularly emphasis on differentiating the methods used in studies that support the unitary-resource vs. the multiple-resource models.

2.3.2 Experimental Methods of Unitary-Resource Studies

One of the earliest studies claiming support for a unitary-resource account of the attentional system was conducted by Ward (1994). Ward’s method required participants to monitor a computer screen for the presentation of an alphabetic letter target. Preceding this target was either a non-predictive auditory (a tone) or visual (colour brightening) cue, or no cue, all of which they were told to ignore. The experiments were designed to measure the extent to which the participants’ attention was captured by the stimulus cues they had been told to ignore. Ward’s results showed similar outcomes whether the cue was within the same or different modality to the target, and this lead Ward to conclude that support had been provided for the unitary-resource model. However, Farah et al. (1989) who also claimed support for the unitary-resource account suggested slightly different cross- compared to within-modal performance. Specifically, Farah et al. suggested slightly differing auditory and visual cuing effects where Ward had found none. Therefore, Farah et al. support a relatively weak unitary-resource position. Why might it be the case that two seemingly similar studies have produced somewhat discrepant results?

Common to both Ward (1994) and Farah et al. (1989) was the fact that the auditory cues were single tones and the visual cues involved brightening the colour of targets.
Importantly, these cues were non-predictive – providing no valid information about the type or positioning of the upcoming target. Therefore, the tones and colours used by both Ward and Farah et al. could almost be relabelled as ‘anti-cues’ in that performance will be hindered if participants divert attentional resources to the cue. The use of such non-predictive, distracting, cues would be quite taxing on participants just as ignoring (as opposed to focussing on) task features is considered to be a particularly difficult task (Kramer, Wickens & Donchin, 1985) requiring significant mental effort (Kahneman, 1972).

Despite the apparent similarity of the experimental designs used by Ward (1994) and Farah et al. (1989) there was a particular feature of the cues used in each study that potentially lead to the Ward study resulting in relatively stronger support of the unitary-resource model than the Farah et al. study. Ward used ‘mixed-cues’, that is, a series of differing auditory and visual cues which preceded the onset of the target, whereas Farah et al. (1989) used only a single auditory and single visual cue. It is possible that, by having a greater array of cues, Ward induced more attention-grabbing distractors. In contrast, participants in Farah et al. were faced with single cues and so, may have been able to become more quickly accustomed or habituated to the cues. It may be suggested that while overall the mental effort required of participants in both these studies lead to support being found for the unitary-resource model, however, the required effort in the Ward study was considerably higher than that in the Farah et al. study (due to the non-habitual nature of the cues in the former compared to the later). Therefore, while the methods of these two studies were extremely similar, the slightly different experimental results obtained by Ward and Farah et al. may be explained by the apparent differences in the cues used in the two studies and provides evidence that not only is there support for the unitary-resource model, but that there can be relatively different levels of the strength of this support which are mediated by the required effort of the task.
A similar juxtaposition across studies that are using essentially the same methods may be made for studies by Arnell and Jolicœur (1995), who found support for a unitary-resource account, and Potter et al. (1998), who provided support for a multiple-resource account of attention. Both studies used the Attentional Blink (AB) paradigm (see Section 1.2 for further details) to investigate the relationship between the auditory and visual modalities. They required participants to monitor dual alpha-numerical strings presented auditorily (via headphones) or visually (on a computer screen), and detect and remember, in order, targets (any of 17 alphabetic letters) presented randomly among a set of distractor numerals.

Arnell and Jolicœur (1995) found significant cross-modal ABs across all the reported experiments thereby providing unreserved support for the unitary-resource position. In contrast, Potter et al. (1998) generally found ABs in within-modal but no cross-modal conditions. However, Potter et al. did manage in one experiment to duplicate the results of Arnell and Jolicœur. Potter et al. explained these outcomes not by supposing that there was evidence for a unitary-resource pool per se – but instead, by suggesting that, because the auditory and visual tasks were different and were presented in different locations, the resulting cross-modal deficits were due to participants needing to alternate attention between spatially and functionally different tasks (task-set switching) rather than because they were sharing attentional resources (Arnell & Larson, 2002).

Two follow up studies (Arnell & Jolicœur, 1999; Arnell & Larson, 2002) in which the need for any task-switching was removed, continued to reveal a reduction in the detection of the second visual target (T2) by a preceding visual or auditory target (T1), an outcome suggestive of attentional competition compatible with a unitary-resource approach. These findings supported the earlier work of Arnell and Jolicœur (1995) and show that a cross-modal AB could be found in the absence of task-switching. A review of the methods used in these studies reveals an alternative to the task-switching hypothesis proposed by Potter et al.
– namely that the faster the presentation rate of visual and auditory stimuli the more likely it is to produce results consistent with the predictions of a unitary-resource model.

 Arnell and Jolicœur (1999) demonstrated that when using RSVP/RAP presentation rates of between 105 and 120ms (similar to that used in Arnell and Jolicœur, 1995) significant cross-modal ABs could be found, but that if the presentation rate was increased to 135ms (as used in the Potter et al., 1988 study) no AB could be detected. Further support of this notion is seen in a study by Shulman and Hsieh (1995) where cross-modal ABs were found using an intermediate presentation rate of 126ms, however, the AB was significantly less robust than those ABs demonstrated by Arnell and Jolicœur (1995; 1999). As such, it would appear that cross-modal ABs are absent under fast presentation rates, present under slow presentation rates, and although present only minimally so at moderate presentation rates. Since faster presentation rates are generally considered to represent a more difficult task compared to identical methods using considerably slower presentation rates (Lavie, 1995), presumably due to target identification decision making being more taxing when less time is available for making judgements (Woods, 1988), it would seem that the role of task difficulty or effort is again revealed as a factor that differentiates the two resource models.

 The implication that heightened task difficulty is part-way responsible for results favouring the unitary-resource position is further reinforced by consideration of neurophysiological studies. In a combined fMRI and behavioural study, Vohn et al. (2007) required participants to monitor low and high pitched tones and to press a key when they heard a change from high to low (or visa versa) on two successive notes. Similarly, in the visual conditions participants monitored circles and squares which changed dimension throughout the course of the experiment. As only differences between successive items were regarded as correct detections, the task therefore required participants to detect not only a change but the temporal characteristics of the change. This constant vigilance and the
involvement of working memory means that the task used by Vohn et al. was considerably difficult. Therefore, it comes as little surprise that performance accuracy was similar in within- and cross-modal conditions, and moreover, that processing in these conditions was undertaken in the same regions of the brain – outcomes reminiscent of the unitary-resource model proposition. Moreover, unlike the tasks used by Wickens (1984), participants were not provided with predictive cues which might facilitate the development of strategies or the activation of automatised responses. For these reasons it is suggested here that the neurophysiological results described by Vohn et al. do, in fact, coincide with the results of the discussed behavioural studies (Arnell, 2006; Arnell & Jenkins, 2004; Arnell & Jolicœur, 1995; Bonnel & Hafter, 1998; Farah et al., 1989; Ward, 1994) and support the suggestion that research that provides support for the unitary-resource model is characterised by experimental tasks that are considerably more difficult that those in studies that provide support the multiple-resource model.

In summary, both behavioural and neurophysiological studies the results from which support a unitary-resource framework have used quite different methods and measurement techniques. Nevertheless, despite the diverse methods these tasks appear to share an underlying dimension – the level of difficulty imposed by the task or the level of effort required of the participant.

2.3.3 Summary: Methods Used in Studies Supporting the Classic Resource Models

There appears good reason to suggest that task difficulty, or required effort, is an experimental design feature that differentiates those studies claiming support for multiple-resources or unitary-resources. Thus, it might be argued that the two models do not represent entirely opposing and unrelated positions. However, this is not an entirely adequate conclusion for two reasons.
Firstly, suggesting that task difficulty is responsible for different resource models does little to explain why different methods would produce such differing experimental outcomes. Could this imply that humans have two different and independent attentional systems which are activated dependent upon the difficulty of the task? Could it mean that the attentional system is ‘staged’ such that modality-specific and modality-independent networks function collaboratively depending upon the task requirements? Or it could mean that ‘easy’ methods are not so taxing as to test resource limitations? Or could it mean even that the ‘hard’ methods perhaps exhaust not only attentional resources but also drain other cognitive resources and that the exhaustion of these other non-attentional resources gives rise to the interference in the unitary-resource position? So, in addition to detailing how the studies differ, it is important to also provide an explanation as to why the task differences between the studies have yielded different results.

Secondly, given that some studies provide weak support for both the multiple- and unitary-resource positions, it appears that a continuum may exist between the two classic models based on the level of task difficulty. Accordingly, it may be viable to consider a model of attentional resourcing which incorporates elements of both the multiple- and the unitary-resource models. Such a hybrid model may be either represent: (1) one of the classic positions more strongly, incorporating features of the opposing model to elucidate findings e.g., the separate-but-linked account proposed by Spence and Driver (1996); or (2) an amalgamation of the two classic models so that, under some conditions, modality-specific resources are used whereas in others modality-independent or ‘general’ resources are required e.g., a hierarchical model like that proposed by Posner (1980). A further examination of such a multiple-resource → unitary-resource continuum will now be evaluated in light of the more recently proposed hybrid models of attention.
2.3.4 Hybrid Attentional Resource Models

Consistent with the multiple-resource model, Spence and Driver (1996) proposed an information resource network that consists of modality-specific attentional resource pools in which auditory processing is served by auditory-only resources and visual processing by visual-only resources, without any sharing of resource between modalities. The predictions of this proposal are consistent with the auditory-visual cross-modal advantage which characterises the classic multiple-resource model (Wickens, 1980). Spence and Driver’s proposition, however, incorporates \textit{links} between these modality-specific systems, which they propose provide “spatial synergy” (pp. 1026) between the modalities but no resource sharing.

Spence and Driver (1996) incorporated these links into an otherwise multiple-resource structure in response to their experimental results which suggested difficulty in decoupling auditory from visual processing when the auditory and visual stimuli were located in similar spatial locations. Spence and Driver found that the detection of an auditory target is faster when it is presented immediately following a visual cue in the same location – a ‘spatial cuing’. Further, they found that both auditory and visual attention is directed to the same spatial location, even when participants were made aware that there was no advantage by colocating auditory and visual attention. As such, these links can be seen to provide a coupling between the focus of attention in the two modalities, while in no way actually sharing resources during this focusing and processing.

As with the other research in which support for this multiple-resource model was evident, the tasks used by Spence and Driver (1996) could be classified as relatively easy; i.e. participants were required to make elevation detection judgements of visual light emitting diode (LED) illumination and of auditory white noise bursts across a series of seven experiments in which cue validity was manipulated. The fact that links were demonstrated by Spence and Driver and not in previous works is most probably due to other experimenters
using headphones to present auditory stimuli, while Spence and Driver used audio
loudspeakers. Presenting auditory stimuli via headphones interferes with the distal origins of
the visual stimuli (e.g., on a computer screen to the front of the participant) and the
corresponding auditory stimuli (to the left and right hand sides of the participant’s head),
whereas loudspeakers allow ecologically valid presentation of auditory and visual stimulus
components (Arnell & Jenkins, 2004). Therefore, Spence and Driver’s separate-but-linked
account does not, in effect, present an argument with the predictions of multiple-resource
models – it simply provides a qualification of the model and therefore the methods of such
experiments should be examined alongside support for the classic multiple-resource model.

Support for the separate-but-linked hypothesis (Spence & Driver, 1996) is growing
(Mondor & Amirault, 1998; Schmitt et al., 2000; Sinnett et al., 2006; Spence et al., 2000).
Consistent with the studies reviewed here, support is evident in situations where little effort is
required of the participant. Nevertheless, not all researchers agree that the account can
explain all of the reported outcomes so decisively (Soto-Faraco et al., 2005).

Soto-Faraco et al. (2005) found that participants could decouple auditory and visual
attention to process auditory stimuli at a separate spatial location to simultaneously presented
visual stimuli. Such an outcome is not easily accounted for by the separate-but-linked
proposal, unless it is possible to voluntarily engage and disengage such links. However, this
is unlikely given that the links have been demonstrated not only in endogenous (voluntary)
atention experiments (Spence & Driver, 1996) but additionally in exogenous (involuntary)
atention research (Spence & Driver, 1997).

Given these difficulties some researchers have instead posited a unimodal-plus-
supramodal (Spence & Driver, 1996), staged (Arnell & Duncan, 2002), or hierarchical
(Posner, 1990) model of attentional resourcing. In fact, it has been suggested that it is
unlikely a unitary-resource or a multiple-resource (including the separate but linked) account
of resourcing could accurately depict the attentional networking of humans (Alais et al., 2006; Sinnett et al., 2006). Instead, some have proposed the feasibility of a hierarchical system which maximises performance by allocating resources from both modality-specific and modality-independent stores depending upon the task (Farah et al., 1989). Proponents of such a hierarchical account come from both multiple-resource and unitary-resource model advocates (Arnell & Duncan, 2002; Farah et al., 1989; Soto-Faraco & Spence, 2002; Spence & Driver, 1996).

Posner (1990) suggests that, among other elements that a hierarchical model would incorporate both modality-specific reservoirs and an overarching modality-independent or amodal reservoir which could direct attentional resources to the modality that was in need of support. However, a precise model of the relationship between these two processes has not been examined – the hierarchical model is little more than a proposal at present (Alais et al., 2006). Given the review thus far, there is reason to propose a relationship between such modality-specific and modality-independent stages that may, in some way, be connected to the difficulty of the tasks. Before any such proposal of a hierarchical network can be hypothesised it is necessary to evaluate why task difficulty has continually been implicated in studies supporting multiple- or unitary-resource models. Table 2.1 provides a summary of the studies reviewed so far including the type of measure (behavioural versus neurological) and the model they subsequently supported in the study.
### Table 2.1

*An Overview of the Research Studies Evaluated in this Thesis, the Measures Used, and the Model Supported*

<table>
<thead>
<tr>
<th>Research Study</th>
<th>Multiple-Resource Support</th>
<th>Unitary-Resource Support</th>
<th>Experimental Measures Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alais and Burr (2003)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Alais, Morrone and Burr (2006)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Arnell (2006)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Arnell and Jenkins (2004)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Arnell and Jolicœur (1995)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Arnell and Jolicœur (1999)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Arnell and Larson (2002)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Bonnel and Hafter (1998)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Duncan, Martens and Ward (1997)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Eimer &amp; van Velzen (2002)</td>
<td></td>
<td>✓</td>
<td>Neurological</td>
</tr>
<tr>
<td>Farah, Wong, Monheit and Morrow (1989)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Glass (1980)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Grady, Van Meter, Maisog, Pietrini, Krasuski and Rauschecker (1997)</td>
<td></td>
<td>✓</td>
<td>Neurological</td>
</tr>
<tr>
<td>Horrey and Wickens (2006)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Jancke, Mirzazade and Shah (1999)</td>
<td></td>
<td>✓</td>
<td>Neurological</td>
</tr>
<tr>
<td>Liu and Wickens (1992)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Luck, Chelazzi, Hillyard and Desimono (1997)</td>
<td></td>
<td>✓</td>
<td>Neurological</td>
</tr>
</tbody>
</table>
Table 2.1 (Cont.)

<table>
<thead>
<tr>
<th>Research Study</th>
<th>Multiple-Resource Support</th>
<th>Unitary-Resource Support</th>
<th>Experimental Measures Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mondor and Amirault (1998)</td>
<td>✓</td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Motter (1993)</td>
<td>✓</td>
<td></td>
<td>Neurological</td>
</tr>
<tr>
<td>Potter, Chun, Banks and Muckenhoupt (1998)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Schmitt, Postma and De Haan (2000)</td>
<td>✓</td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Shulman and Hsieh (1995)</td>
<td></td>
<td>✓</td>
<td>Neurological</td>
</tr>
<tr>
<td>Sommers, Dale, Seiffert and Tottell (1999)</td>
<td>✓</td>
<td></td>
<td>Neurological</td>
</tr>
<tr>
<td>Spence and Driver (1994)</td>
<td>✓</td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Spence, Ranson and Driver (2000)</td>
<td>✓</td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Ward (1994)</td>
<td></td>
<td>✓</td>
<td>Behavioural</td>
</tr>
<tr>
<td>Webster and Thompson (1954)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Wickens (1976)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Wickens (1984)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
<tr>
<td>Woodruff, Benson, Bandettini, Kwong, Howard, Telavage et al. (1996)</td>
<td>✓</td>
<td></td>
<td>Neurological</td>
</tr>
<tr>
<td>Yeh and Wickens (1988)</td>
<td>✓</td>
<td></td>
<td>Behavioural</td>
</tr>
</tbody>
</table>

* although conforming to the multiple-resources stance on attentional resource modality independence these studies recognise that some degree of interrelatedness is required - the proposed separate-but-linked model.
2.4 **INTRODUCING COGNITIVE LOAD**

When attempting to explain why a relationship appears to exist between task difficulty and the support for either the multiple- or unitary resource models, it appears reasonable to consider a model of the relationship between cognitive processing, performance and the factor(s) of task difficulty or perceived mental effort. For example, Cognitive Load Theory (CLT) (Sweller, 1988), posits several task completion aspects of importance to the present review on attention resourcing. In Chapter 1, this notion of ‘load’ was considered implicitly. However, the vagueness with which it has been approached in much of the literature reviewed here, that is, in many studies the role of load is not directly considered during experimental design, is part of the reason that work is not decisive.

Firstly, CLT indicates that all cognitive resources are limited, a proposition supported throughout the attention resourcing literature (Brünken et al., 2003). Secondly, it is proposed that every task in which an individual engages will require some portion of these cognitive resources (Wiebe, Roberts & Behrend, In Press). However, each task will require a unique share of the resources relative to the burden imposed by the external features of the task, mediated by the individual’s ability to undertake the task. For instance, a more skilled operator may find a particular task to be considerably less effortful than a novice undertaking the same task due to the development of additional techniques (e.g., task automation, decision heuristics). Thirdly, CLT proposes that as the relevant cognitive resources are expended, operator performance will, at some point, begin to diminish. Such predictions appear make the CLT model well suited to contribute to the involvement of task difficulty to the present multiple- vs. unitary-resource debate.

The CLT model, in its original form, conceived of ‘load’ as the subjective experience of an individual engaged in a task (Sweller, 1988), rather than a feature, entity or event of the task that can be objectively quantified. As such, Sweller suggests that it is not possible to
define a task as being ‘low’ or ‘high’ load, only to ask individuals retrospectively whether they experienced ‘low’ or ‘high’ load when completing the task. Thus, it could be argued that Sweller conceives load as perceived or subjective load (Weibe et al., In Press).

This distinction has important implications for experimental design as, according to Sweller, cognitive load would be an outcome of the experiment, rather than a manipulation used within the experiment. Nevertheless, it seems reasonable to suggest that perceived load would, in many ways, relate to the complexity of the task. While it must be recognised that individuals will report different perceived experiences of load for the same task, for any given individual, there should be a relationship between the task complexity and the load they report.

A similar relationship has been proposed by Woods (1988), who conceived of this ‘burden’ as the cognitive demands imposed on the individuals as they complete a task. Woods believed that the demand of any task could be calculated by evaluating the features of the task itself (which he referred to as the “world”), the abilities of the individual (or ‘agent’), and the individuals ‘representation’ of the task to arrive at a measure of cognitive complexity. Therefore, rather than being concerned with the experience of individuals and how that experience impacted their performance (c.f. Sweller, 1988), Woods was interested in measuring the parameters of the task, ascribing a cognitive complexity score, and then predicting performance.

Essentially, the review conducted earlier in this chapter was concerned with task features and how they appeared to share a relationship with the outcomes concerning the presence and size of any cross-modal advantage. As such, there is an assumption that the varying task complexity parameters reviewed shared a positive linear relationship with the level of cognitive load that is experienced by the individual.
While Sweller (1988) and Woods (1988) seem to be concerned with different ways of predicting task performance, these approaches are not incompatible methods of addressing how a task may impact upon an individual’s performance. The differentiation between Sweller’s perceived cognitive load and Woods’ cognitive-complexity concepts is rarely made for many seem to believe the two are so closely associated that difficulty (a measure of required effort) and complexity (a measure of task features) are interchangeable terms. For instance, Brünken et al. (2003) propose a three-part taxonomy to describe those task-operator features which affect ‘cognitive load’.

Firstly, *intrinsic cognitive load* is suggested to arise from the complexity of the materials or information units required to be processed in a task. Secondly, Brünken et al. (2003) describe *extraneous cognitive load* as a result of how the materials, information or stimuli are physically presented to the individual. These first two elements of the Brünken et al. taxonomy are not especially distinct from Woods’ (1988) description of how the ‘world’ (via its dynamism, interactivity, uncertainty and associated risks of the tasks being engaged in) impacts the cognitive complexity of a task and could, for this reason, be regarded as ‘task-related cognitive load’ components. Moreover, they very much afford to Wood’s position that a task can be analysed to determine the impact that it will have on an individual’s performance. The only difference being is that Brünken et al. contend that these measures can not be assessed directly to determine performance. Rather, the two features impose on the capacity of the individual and create the experience of load. Therefore, performance can be determined as a product of the amount of experienced load.

The third category in the Brünken et al. (2003) taxonomy, *germane cognitive load*, relates to the effect that the *abilities* and *knowledge* of the individual have on the overall cognitive load of a task. Arguably, experience within a domain may moderate the level of cognitive load induced. This final feature can be reconciled with Wood’s conceptualisation of
‘agents’ and how they interact within their world ‘representation’ and could therefore be regarded as operator-task based cognitive load. Finally, Brünken et al. see these three forms of cognitive load as being additive, and that the resources available within the system can be determined by summing the load associated with each extraneous, intrinsic and germane together and examining whether any residual resources remain (see Figure 2.1).

![Figure 2.1](image)

Figure 2.1. Brunken et al (2003): Depiction of how all three forms of cognitive load (extraneous, intrinsic and germane) are additive and influence the amount of residual cognitive resources available dependent on the relative size of the load for each.

As reviews of the cognitive load literature do not disambiguate cognitive load (Sweller, 1988) and cognitive complexity (Woods, 1988) as predictors of cognitive load, this thesis will be concerned with appreciating how both concepts are related to one another and, in turn, how they impact performance. The thesis adopts the taxonomy presented by Brünken et al. (2003) as it elegantly details the various features of the operator-task relationship that are likely to impact the overall load of a given situation. Additionally, at this point, it will be accepted that manipulations of task complexity will likely similarly impact perceptions of load. For simplicity in this thesis, the following two conceptualisations of cognitive load will be used hereafter: Task-related load will be considered as imposed cognitive load (akin to Woods conceptualisation of cognitive complexity) and; the experience of load will be considered to be perceived cognitive load (akin to Sweller’s proposition of cognitive load).
However, given that imposed cognitive load will dominate this thesis, it will simply be referred to as *cognitive load*.

To capture the similar, yet conceptually different, approaches of Woods (1988) and Sweller (1988) cognitive load will be measured in terms of both the external features of the task and the experience of the individual operator. A number of objective and subjective approaches have been suggested in the literature (Brünken et al., 2003; Brünken, Steinbacker, Plass & Leutner, 2002), and while Brünken et al. (2002) claim these to all be measures of cognitive load – they are, in fact, measures of cognitive complexity and perceived cognitive load when accepting the definitions presented in this thesis. These four general categories of measurement are classified along two dimensions: the *objectivity* and the *causal relation* of the measure. Objectivity relates to whether an approach uses self-reported data (a subjective measure) – essentially measuring perceived cognitive load – or whether behaviour, physiological states or performance outcomes are evaluated (objective measures) – essentially measuring cognitive complexity. The causal relation dimension concerns whether the measure is directly related to the construct being measured or whether the construct is evaluated using a related phenomenon (an indirect measure). For instance, a direct measure of cognitive load may be the number of pieces of information to be processed in a given time period, while an indirect measure of the same load may measure the accuracy of the decision made once the pieces of information are processed. Thus, any measure of cognitive load may be: (1) direct and subjective; (2) direct and objective; (3) indirect and subjective or; (4) indirect and objective. A depiction of these four categories and examples of representative measures are provided in Table 2.2.
Table 2.2

*Cognitive Load Measurement Categorisation* and Example Measures for Each Category

<table>
<thead>
<tr>
<th>Causal Relationship</th>
<th>Indirect</th>
<th>Direct</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjective</strong></td>
<td>Example: Self-reported measure of mental effort exerted</td>
<td>Example: Self-reported difficulty of materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Objective</strong></td>
<td>Example: Physiological (heart rate, skin conductance, blood pressure); Behavioural and Learning Outcome measures.</td>
<td>Example: Brain activity measures (fMRI); dual-task performance.</td>
</tr>
</tbody>
</table>

Explanation: Self reporting of the individual’s experience make these measures subjective. The self-report of ‘difficulty’ is the perception of cognitive difficulty of the task itself and this is indirect. On the other hand the report of mental effort is indirect as the individual reports on their own feelings not the external task.

Explanation: These measures are objective because they can be taken without active involvement by the individual. Physiological and behavioural measures are considered indirect as while cognitive load may adversely affect these, it is possible another construct could be responsible. In contrast, dual-task performance and brain activity are unequivocal as direct measures of cognitive load.

* This categorisation technique is consistent with that documented by Brünken et al. (2003).
The appeal of using a concept like cognitive load to evaluate attentional resourcing is that a similar approach was used to aid the disambiguation of the early versus late debate addressed in Section 2.2. Several experimental and review papers (Forster & Lavie, 2009; Lavie, 1995; Lavie, 2005; Lavie, Hirst, de Fockert & Viding, 2004) have demonstrated that the reason that the early versus late selection model divide occurred was that, despite the veracity of the cognitive limitations advocated by each approach, they merely reflected different stages of processing. Thus, while the early and late positions were seen as antagonistic (in much the same way the multiple- and unitary-resource models are), a review of the two proposals made it clear that each, by way of their experimental methods, had effectively tapped separate yet interrelated constructs.

It is suggested that the review in this thesis of the multiple- versus unitary-resource divide in terms of experimental methods and the level of cognitive load imposed, might be as beneficial as the similar approach taken by Lavie and her colleagues (Lavie, 1995; Lavie, 2005; Lavie et al., 2004; Lavie & Tsal, 1994).

2.5 Thesis Rationale and Research Questions

The nature of the attentional resource has been debated for over 50 years (Alais et al., 2006). There appeared to be an early assumption that one model of attention would eventually dominate, as the results of carefully planned experiments came to light. However, this was not the case. Rather, the divide in the literature between multiple- and unitary-resource accounts arose, and continues (Arnell, 2006; Arnell & Jenkins, 2004; Jolicœur, 1999; Potter et al., 1998; Soto-Faraco & Spence, 2002). Nevertheless, in more recent times, it has been acknowledged that neither model alone appears capable of explaining all the results (Sinnett et al., 2006). Moreover, the depth and multidisciplinary breadth of the experimental
results indicates perhaps that neither model of attention is completely inaccurate (Burr & Alais, 2006; Ernst & Bulthoff, 2004; Spence & Driver, 1996).

While the multiple-resource model predicts perfect cross-modal performance, equal to performance if the auditory and visual tasks were completed independently of one another, the unitary-resource model predicts that cross-modal performance should equal performance when completing two of the same modality tasks. Conceivably, both models of attention are appropriate reflections of some level of human perceptual and cognitive processing, an outcome supported since strict adherence to one set of model predictions is not evident in the literature. For instance, while Treisman and Davies (1973) showed that two similar tasks (e.g., two auditory tasks or two visual tasks) interfere to a greater extent than two dissimilar tasks (auditory and visual), their data do not support the assertion that these two unalike tasks will be perfectly resource-shared (that is, performed without decrement from single-task performance). Indeed, Treisman and Davies indicated significant cross-modal performance reductions. As such, their results do not conform perfectly to a multiple-resource model (non-perfect time sharing) but neither to a unitary-resource model (equal within- and cross-modal decrements). Therefore, it might be suggested that past research has not been sufficiently rigorous in relation to the definitions of these models. For this reason, they have, to date, appeared to be conflicting and seemingly irreconcilable.

The review in this chapter has highlighted that the apparently incompatible findings may reflect less that the models are contradictory and more that the different experimental paradigms have unearthed two different aspects of resource allocation within a single attentional system. One aspect of this system is used for processing less cognitively complex stimuli and conforms to a multiple-resourcing structure; the other is involved in the processing of cognitively complex stimuli and operates in a unitary-resource manner. By accepting the potential for a continuum between the classic positions (see Figure 2.2) which
is, in part, explained by the notion of cognitive load, it is possible to understand how research results can be viewed to be complementary, rather than adversarial. In summary, it is proposed that the distribution of attentional resources is determined by the cognitive load involved in processing the stimuli. Resources are allocated in terms of processing demands: easily processed stimuli are handled by modality-specific resources whereas more difficult processing requires additional resources provided by a modality-independent resource.

**Figure 2.2.** The proposal. (a) The proposed continuum ranging from strict multiple resources through weaker multiple- (e.g., separate-but-linked account) and unitary-support, through to strong unitary-resource support. This proposed continuum functions as a function of task complexity and perceived cognitive load. (b) A proposal of a hierarchical model which spans the entire continuum. Both the continuum and the hierarchical model span task characteristics from simple to complex, and low to high perceived cognitive load (or mental effort).
In this thesis, three main issues will be investigated. The first issue concerns to degree to which cognitive load is, indeed, a factor which can influence attentional processing, and secondly, should cognitive load be demonstrated to influence attentional resourcing, there will be an evaluation of how the two classic models might interact to explain processing. The third issue concerns whether alternative approaches on offer (e.g., the hierarchical or staged models) suggested in the literature can be further developed to explain the interaction between two different processes, processes that have previously given rise to two separate resourcing models.
Chapter 3

EXPERIMENTAL DESIGN:

DEVELOPING A RESOURCE-MODEL-NEUTRAL METHOD FOR ASSESSING ATTENTIONAL RESOURCES AND COGNITIVE LOAD
3.1 **CHAPTER OVERVIEW**

This chapter will take a short departure from the momentum developed in Chapters 1 and 2 in which various attentional models were discussed and the hypothesis put forward that cognitive load may impact how resources are allocated during dual-task processing. This Chapter concerns the development of appropriate methods for the experimental investigation of such a cognitive load hypothesis. The argument advanced in Chapter 2 was that experimental methods have contributed to the empirical support for two seemingly incompatible models. Therefore, it is essential that the experimental methods adopted in the present series of experiments are model-neutral – that is, that the designs will not bias the experimental results towards one model over the other outside of the intended manipulations of cognitive load that are intended. The aim of this chapter is to explain how these manipulations are to occur to avoid bias.

3.2 **INTRODUCING THE DESIGN METHODS TO BE USED IN THIS THESIS**

Despite the apparent relationship between model support and experimental methods (the ‘model-method’ association) selected by researchers, as outlined in Chapter 2, it is necessary to investigate this claim within the confines of a single experimental design prior to any determination of causality. As a result, it is essential that an experimental framework be designed itself will not associated with either the multiple- or unitary-resource models (that is, a model-neutral design) and therefore, is unlikely to contribute further to the difficulty associated with the interpretations of outcomes.

3.2.1 **Independent Variables and Contrasts of Interest**

The most effective means of determining the presence of bias towards either modality-specific or modality-independent models of attention is to examine the
performance of individuals in both within- and cross-modal conditions. The presence of a cross-modal advantage (Rodway, 2005), indicates modality-specific resource limitations, an outcome consistent with a multiple-resource approach for tasks undertaken in different, or crossed, modalities will not compete for attentional resources (Alais et al., 2006). In contrast, should no such cross-modal advantage be found, it is most likely that a single modality-independent limitation is in operation, a finding that would support the unitary-resource framework for all tasks compete for the same resources (Spence & Driver, 1996). Across all experiments in this thesis, within-modal performance and cross-modal performance will be compared to establish whether a cross-modal advantage is present or absent, and whether the magnitude of any cross-modal advantage may alter.

Given that the auditory and visual modalities have been identified (see Section 1.3) as those most likely of all perceptual systems to be integrated at some level, they will be the modalities focussed on in this series of studies. As such, the experiments will be designed so that participants will (at a minimum) complete experimental trials consisting of two auditory tasks (AA); two visual tasks (VV); and one auditory together with one visual (AV) task. This independent variable will be referred to as ‘modality-combination’ comprising three levels (i) AA (ii) VV and (iii) AV.

To examine the cognitive load on attentional resourcing, it is necessary to establish several experimental conditions in which each type of cognitive load – extraneous, intrinsic and germane cognitive load (Brünken et al., 2003) – can be manipulated and examined. Details of the manipulations for each of these types of cognitive load will be described within the individual experiments where they are relevant. However, at an overview level, the aim is to use conditions that induce low through to high cognitive load to determine whether any changes in cognitive load are associated with corresponding changes in the degree of a cross-modal advantage.
While neither within- versus cross-modal comparisons, nor manipulations of cognitive load are especially novel in information processing experiments, to date the two have not been explicitly evaluated in concert. Therefore, it is necessary to develop a new method to examine the two concepts concurrently. This is not altogether undesirable for, as demonstrated in Chapter 2, current experimental methods may not be model-neutral, and developing a new approach to evaluate these concepts together can be done with careful consideration of the method also being model-neutral. The following sections are concerned with justifying an experimental method that is model-neutral – defined as a method via which support has been found for multiple-resource models in some and unitary-resources in other studies.

3.2.2 A task framework to examine multi-modal attention

A number of experimental approaches have been used to investigate attentional resourcing, including dual-task divided-attention tasks (2 x primary tasks), dual-task selective-attention tasks (1 x primary, 1 x secondary tasks), attentional blink, inattentional blindness, and repetition blindness (see Section 1.3). However, to date, no single experimental method has been established as a benchmark. Three main features are required to establish such a benchmark: firstly, the method must contrast same-modality (AA and VV) and cross-modality (AV) attentional processes; secondly, it must accommodate manipulations of cognitive load; and thirdly, in past research it should be associated with results supporting both unitary- and multiple-resource models equally. In Table 3.1 a number of methods and stimuli previously used in attentional resourcing research are identified, together with an indication of the proportion of studies using the particular method features (either paradigm, stimuli and/or response) and the model subsequently supported by the research as a way of establishing the neutrality of the
methods/stimuli. Specifically, a design feature will be considered model-neutral if in past research it has been equally associated with both classic resource models. Those methods selected, based on model-neutrality, for use in the present series of experiments will be described in the following sections.

Table 3.1

*Experimental methods, stimuli, responses and measures as used in those studies reviewed in Chapter 2 with a summary of the resource model subsequently supported by the study*

<table>
<thead>
<tr>
<th>Experimental Methods/Stimuli Employed</th>
<th>% of studies supporting the Multiple-Resource Model</th>
<th>% of studies supporting the Unitary-Resource Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual-task (divided)</td>
<td>58%</td>
<td>42%</td>
</tr>
<tr>
<td>Dual-task (selective)</td>
<td>57%</td>
<td>43%</td>
</tr>
<tr>
<td>Repetition Blindness</td>
<td>29%</td>
<td>71%</td>
</tr>
<tr>
<td>Attentional Blink</td>
<td>31%</td>
<td>69%</td>
</tr>
<tr>
<td>Word detection in spoken sentences</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>Word detection in text-based sentences</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>Spoken target letter in letter distractors</td>
<td>52%</td>
<td>48%</td>
</tr>
<tr>
<td>Spoken target numeral in letter distractors</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>Text-based target letter in letter distractors</td>
<td>51%</td>
<td>49%</td>
</tr>
</tbody>
</table>
Table 3.1 (cont.)

<table>
<thead>
<tr>
<th>Experimental Methods/Stimuli Employed</th>
<th>% of studies supporting the Multiple-Resource Model</th>
<th>% of studies supporting the Unitary-Resource Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based target numeral in letter distractors</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>Pitch changes</td>
<td>84%</td>
<td>16%</td>
</tr>
<tr>
<td>Tone pattern changes</td>
<td>82%</td>
<td>18%</td>
</tr>
<tr>
<td>Colour grating changes</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>Geometric changes</td>
<td>80%</td>
<td>20%</td>
</tr>
<tr>
<td>Simulated Driving</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>Interactive Conversations</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>Speech Shadowing</td>
<td>48%</td>
<td>52%</td>
</tr>
<tr>
<td>Speeded Response</td>
<td>55%</td>
<td>45%</td>
</tr>
<tr>
<td>Unspeeded Response</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td>Accuracy/Error Rates</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Reaction/Response</td>
<td>74%</td>
<td>26%</td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain mapping</td>
<td>66%</td>
<td>44%</td>
</tr>
</tbody>
</table>
**Dual Task Framework.** The dual task target monitoring and identification method fits the specifications set forth above when designing a benchmark experimental method; facilitating cross- and within-modality contrasts and systematic cognitive load manipulations and which is model-neutral. Dual task designs require participants to monitor two stimulus sources concurrently and to respond when a target occurs in either, or both, of the two sources. This method was selected over alternatives (such as, repetition blindness and the attentional blink) because it appears to be a model-neutral design (given its equal association with both models) and, additionally, because the measures used in the dual task method (e.g., reaction time to target presentation and accuracy of target identification) are direct and objective measures of cognitive load – which are considered the most reliable – particularly during early stage research like this (Brünken et al., 2003).

The dual-task framework considered most useful in the present studies was the divided-attention design, again because it facilitates all those features required of a benchmark. Of interest in these studies are the within- and cross-modal processing capabilities and limitations, making a divided attention task favourable over the alternative, selective-attention approach. In the latter secondary task paradigm, information about residual resources in the system will result in an outcome that cannot adequately provide details about a structure encompassing both modality-specific and modality-independent resources as residual resource measures do not allow a fine analysis of performance decrements in both tasks simultaneously.

**Task Stimuli.** Within this selected dual-task paradigm, alphabetic letters were considered the most appropriate stimuli to use as they can be presented in both auditory (spoken RAP sequences) and visual (text based RSVP sequences) modes, and, thus, presumably represent relatively equivalent stimuli in both modalities (see Appendix A for
a listing of all alphabetic letters considered in these experiments). Alternative stimuli (e.g., pitch discrimination (auditory), colour-switching (visual) used in previous studies (Alais et al., 2006; Mondor & Amirault, 1998; Ward, 1994) do not offer the same level of equivalence across the two modalities because the underlying cognitive processes involving two modalities may well be distinct and, could thus, contribute to bias towards explanations of one model over the other. In Chapter Two (Section 2.3), letter identification appeared to be model-neutral (see Spence & Driver, 1996 (separate-but-linked); Alais et al., 2006 (multiple-resources); Arnell & Larson, 2002 (unitary-resources) among others), presumably due to letter identification being a common day-to-day activity (e.g., reading). Nevertheless, as the task requires participants to discriminate a specific target from a milieu of distractors, the task itself is not one regularly performed and may not be automated at a cognitive level. Therefore, undertaking this type of task may not be extremely effortful, yet not entirely effortless. This, in all likelihood, is why it is not over represented in experiments in which support is provided for one model over the other (see Figure 2.1).

Response Format. The speeded response paradigm – in which participants are required to press a key as quickly as possible following the presentation of the target – was preferred over the unspeeded response paradigm – in which participants provide details of the target at the conclusion of the trial (e.g., whether the target was an alphanumerical character, exactly which character it was). Moreover, in the review conducted here, the speeded response paradigm was considered slightly more model-neutral than its unspeeded counterpart – an outcome confirmed by Hein, Parr and Duncan (2006) who concluded that, within the attentional blink (AB) domain, speeded response studies produced stronger ABs than those produced in studies where an unspeeded response format was adopted.
Dependent Variables. A number of objective performance measures, together with a series of subjective ratings items are used to evaluate the effects of cognitive load in the following studies. The dependent variables of interest were refined throughout the studies and as such, will be outlined prior to each study. Broadly, both reaction time (RT) and accuracy measures were recorded – and as is evident in Table 3.1, both of these are considered model-neutral. However, the finer details of these dependent variables will be addressed in the following sections where the manipulations and measures of cognitive load are discussed in more detail. The subjective items included in this series of studies have not previously been assessed within the studies reviewed, here presumably since attentional resourcing and cognitive load have not been investigated concurrently. Therefore, they do not appear in Table 3.1. These subjective items were included on the basis that the definition of cognitive load still remains ambiguous and providing converging data on the phenomena was considered preferable in this novel investigation.

3.2.3 Manipulation of Cognitive Load

In this series of experiments, three types of cognitive load will be investigated – extraneous (how information is presented), intrinsic (task-related) and germane (task-operator relationship) (Brünken et al., 2006). All three forms of intrinsic and germane cognitive load will be evaluated in isolation and concurrently and several different aspects of human behaviour that contribute to these forms of cognitive load (e.g., exposure and motivation) will be addressed in the first three experiments. Accordingly, the specifics of intrinsic and germane cognitive load are addressed in the rationale for each experiment, rather than in this chapter. However, extraneous cognitive load will be evaluated in all the experiments so it is addressed at this point.
As defined by Brünken et al. extraneous cognitive load concerns how materials or information are *presented* to the individual and the burden imposed by the presentation method. The question of interest across all the experiments is whether a cross-modal advantage is present (indicative of the multiple-resource model) and if so, the degree of this advantage, or absent (supporting the unitary-resource model). To determine the presence and magnitude of the cross-modal advantage comparisons will be drawn between performance when stimuli are *presented* cross-modally (i.e., auditory/visual) and within-modally (i.e., auditory/auditory or visual/visual). Consequently, extraneous cognitive load is just as much a manipulated variable (cross-modal vs. within-modal presentation methods), as it is a dependent variable of interest (comparing performance cross-modally to within-modally). Therefore, it is possible that extraneous cognitive load may confound the within- and cross-modal manipulations and, for this reason, the impact of the cognitive load manipulations will be measured not only through objective performance measures (reaction time and accuracy), but also through subjective measures (e.g., questionnaires focussing on mental effort, task complexity and difficulty, and self-monitoring of performance (confidence)).

Taking both objective and subjective measures of the effects of the cognitive load manipulations will mean that any objectively measured performance decrements can be compared to the load manipulations in place and, in turn, matched to participants’ experience of that load. This is important as the nature of cognitive load either as an imposed task feature or as a subjective experience remains uncertain (see Section 2.5). Providing both objective and subjective appraisals of the manipulations will enable more comprehensive conclusions concerning the imposition of cognitive load and moreover will provide a manipulation check; does the cognitive load manipulation used in the task actually impact participant perceptions of cognitive load?
3.2.3.1 Objective Measures of Cognitive Load

Brünken et al. (2003) presented a two-dimensional classification of measure of cognitive load (See Table 2.1). Therefore, to measure the effects of cognitive load objectively, both reaction time (RT) and accuracy measures, as suggested by Brünken et al., were considered most appropriate as cognitive resource overload should manifest in increased RT and decreased accuracy (Alais et al., 2006). Therefore, the benefits of using RT and accuracy to examine performance, under the defined experimental conditions, is that both represent methods for evaluating attentional resource limitations (Spence & Driver, 1998) and further that, the measurement and interpretation of RT and accuracy is relatively straightforward and is unlikely to be misinterpreted. Moreover, RT and accuracy have been used in studies where support has been found for multiple-resources (Schmitt et al., 2000), and is equally represented in those where support was found for unitary-resources (Ward, 1994). Therefore, the measures tend to model-neutral.

3.2.3.2 Subjective Measures of Cognitive Load

As shown in Table 2.1, cognitive load can also be measured using a variety of subjective measurements. Subjective methods of assessment suggested by Brünken et al. (2003) include self-reported measures of ‘invested mental effort’ (pp. 55) (indirect-subjective measure) and ‘self-reported difficulty of materials’ (pp. 55) (direct-subjective measure). Both are easily examined by requiring participants to complete questionnaires following each modality combination x cognitive load block of experimental trials.

Self-report ratings can be contrasted with the pre-determined levels of cognitive load to investigate whether the manipulations affect not only objective performance but whether such manipulations were detected by the participant. This is an important contrast given the earlier findings that task complexity (task-related cognitive load) does
not share a perfect positive relationship with perceived cognitive load (Reed et al., 1985). Determining that the cognitive load manipulations are experienced subjectively as anticipated is essential because unrecognised cognitive disengagement (performance decline due to participant deciding not to continue) could lead to unaccountably non-linear performance decrements i.e., decrements not in line with the a priori assumptions.

To ascertain individuals’ experience of the objectively-defined and manipulated cognitive load, participants will be asked to rate experimental trials in terms of: (i) complexity, (ii) difficulty, and (iii) confidence (see Table 3.2). Additional subjective questions were incorporated into later experiments to reflect independent variables progressively introduced (i.e., motivation).

3.4 Summarising the Method Framework to be Used in this Series of Studies

In the experiments in this thesis, the dual-task paradigm will be used with RT and accuracy as objective dependent variables, and self-ratings of complexity, difficulty and confidence as subjective dependent variables. The dual-task paradigm involves the presentation of alphabetic letter sequences containing a target and multiple distractor letters to which participants are asked to make a speeded response (a key press) on their identification of the target.
Table 3.2 Validating Cognitive Load Manipulation using Subjective Questions

<table>
<thead>
<tr>
<th>Target Concept</th>
<th>Questions and Response Scales</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complexity</strong> (subjective-Direct)</td>
<td>Q: How many letters were similar to the target letter OR How many letters required processing in this trial</td>
<td>Woods (1988) was interested by the load imposed as the objective and quantifiable features of a task altered. Asking participants about ‘complexity’ requires them to reflect on the features of the task and not their experience in completing the task.</td>
</tr>
<tr>
<td></td>
<td>R: 7 point likert scale from few → many.</td>
<td></td>
</tr>
<tr>
<td><strong>Difficulty</strong> (Subjective-Indirect)</td>
<td>Q: How difficult was this trial?</td>
<td>Sweller (1988) was concerned with the experience of cognitive load.</td>
</tr>
<tr>
<td></td>
<td>R: 7 point likert scale from not at all difficult → very difficult</td>
<td>Asking participants to rate ‘difficulty’ requires them to reflect on their experience when completing the trial.</td>
</tr>
<tr>
<td><strong>Confidence</strong></td>
<td>Q: Reflecting on your performance do you think that: (i) You missed all the targets (ii) You missed a portion of the targets (iii) You correctly identified both targets (iv) You have no idea</td>
<td>Given accuracy scores will be calculated on correct identification of all possible targets within a pre-determined time frame this rating was included to ascertain initially whether the RT window being imposed in the experiment may be too strict (e.g., participant believing they identified all targets, but RT cut off relegating some trials to be incorrect; and in later experiments to determine whether confidence was a predictor of performance.</td>
</tr>
</tbody>
</table>
Four sets of experiments are now reported in which the general research question is whether cognitive load influences how attentional resources are distributed. By this means, it will be determined whether: (i) both multiple- and unitary-resource models can be supported in a single experiment in which all aspects other than cognitive load are held constant; (ii) how different types of cognitive load (extraneous, intrinsic and germane), and other individual differences (e.g., level of motivation and/or exposure) may contribute to support for one or the other model; and (iii) whether a composite model – similar to the previously hypothesised hierarchical or staged attentional models (see Section 2.5) – may better account for the deployment of attentional resources in information processing.
Chapter 4

STUDY ONE

MANIPULATING COGNITIVE LOAD WITH MASKING
4.1 RATIONALE

In Chapter 2 it was shown that there is a strong link between the cognitive load of the task, the mental effort required of the individual and the resource model supported in the findings of the research. Additionally, in Chapter 2 it was shown that more simple experimental designs involving tasks which could be described as automatic were highly correlated with support for a multiple-resource model of attention (Alais et al., 2006; Horrey & Wickens, 2006; Soto-Faraco et al., 2005; Mondor & Amirault, 1998; Potter et al., 1998; Spence & Driver, 1996); whereas methods which require significantly more effort by the participant most regularly lead to support for the unitary-resource model of attention (Arnell & Larson, 2002; Eimer & Van Velzen, 2002; Jolicoeur, 1999; Ward, 1994; Farah et al., 1989). This first study is therefore designed to examine whether under varying levels of cognitive load outcomes consistent with both classic models of attention can be found in a single experiment. It is the presence or absence of any cross-modal advantage which has in the past dictated the need for two adversarial models (Wickens & McCarley, 2008). Support for both models would be demonstrated if under some cognitive-load conditions a cross-modal advantage can be documented, and under changing cognitive load conditions any such cross-modal advantage disappears.

The following questions will be addressed in this study:

1. Does the level of cognitive load affect from which attentional resources reservoir(s) resources will be drawn from during processing?

Hypothesis 1: A cross-modal performance advantage (i.e., shorter RTs and increased accuracy in AV compared to AA and VV) will be present and it will diminish as the level of cognitive load increases. The point at which any cross-modal advantage will disappear cannot be predicted at this stage.
What is the relationship between imposed cognitive load (aka. cognitive complexity; Woods, 1988) and perceived cognitive load (aka. cognitive load; Sweller, 1988)?

Hypothesis 2: Ratings of complexity will increase as cognitive load level increases.

Hypothesis 3: Ratings of difficulty will increase as cognitive load level increases.

4.2 Experimental Design

A dual-task target-identification paradigm was used with both auditory and visual representations of alphabetic letters (see Chapter 3). The degree of any cross-modal advantage was determined by contrasting the target letter identification scores when two same modality tasks were performed compared to when two different modality tasks were performed.

A 3 (modality-combination) x 3(cognitive-load) within-subjects factorial design was employed. For the 3 level modality-combination independent variable the first two levels were within-modality dual-tasks conditions in which participants completed either (i) two auditory tasks (AA) or (ii) two visual tasks (VV); and the third level was a cross-modal condition in which participants completed one auditory and one visual task (AV). In this experiment no single-task trials were included.

The three levels of load were (i) low cognitive load (LCL) (ii) medium cognitive load (MCL) and (iii) high cognitive load (HCL). These levels were determined by changing the similarity between the target and distractors letters. Previous research has demonstrated that the use of visually similar distractor letters increases difficulty of identifying a target letter in a visual single-task scenario (Scialfa, Esau & Joffe, 1998). A similar finding is found in a
psychoacoustic study suggesting that the more similar two channels of sound are, the more unlikely a listener will be able to differentiate and distinguish between them (Gerritsen, Slawinski & Eagle, 2006). Accordingly, in the low cognitive-load condition all distractor alphabetic letters were highly dissimilar to the target letter, in the medium-cognitive load condition half the distractor letters were highly similar to the target, while the other half were highly dissimilar to the target letter, and in the high cognitive-load condition all distractor letters were highly similar to the target letter. Presumably, in both the auditory (Brungart, 2001) and visual (Jacobson & Rheinlander, 1978) modalities the more similar the target and distractor letters are, the more likely they are to be confused due to feature masking.

The dependent variables were objective performance measures, and subjective ratings of task complexity and difficulty. The objective performance measures were (i) reaction time (RT) – response time from target presentation to target identification key-press and (ii) target identification accuracy – number of targets correctly identified (both, one or neither) within the pre-determined reaction time frame of 3,400 ms.

The subjective performance measures were pen and paper questions which immediately followed a subset of experimental trials asking participants to indicate the complexity of the trial, the difficulty of the trial, and their confidence in their performance on the trial. The form of these is shown in Appendix B.

4.3 Method

4.3.1 Participants

The participants were 47 undergraduate students from the University of Western Sydney: male \((n = 9)\) and female \((n = 38)\). The age range was from 17 to 54, with a mean age of 23.7 years and all reported normal (or corrected-to-normal) visual acuity and normal
hearing. All students received course credit for a first year psychology course in return for their participation.

4.3.2 Materials and Apparatus

The experiment was presented and data collected with DMDX software (Forster & Forster, 2003) run on a LG Laptop computer with a 17” widescreen colour monitor. The participants were seated before the computer at a set distance of approximately 55-60 cm (see Figure 4.1 for a depiction of the experimental set up), resulting in the visual stimuli subtending approximately 1° of the visual angle (with a median width of 1 cm). Both visual stimuli were presented at the top of the computer screen (to minimise offset between them and the auditory stimuli) and displaced at -8° and +8° from the central plane. This 16° separation distance was considered close enough to reduce the need for any eye movements, an outcome which might confound reaction time and in turn accuracy measures (Alais et al., 2006) but distant enough to reduce the possibility of conjunctures (perceptual fusions) between the left hand side (LHS) and right hand side (RHS) stimuli (Treisman & Schmidt, 1982). Auditory stimuli were presented via audio loud speakers (Genelec Analogue 50/60Hz) connected to the audio jack of the laptop computer and positions just above the top of the laptop screen. This vertical offset was required so that the visual stimuli would not occlude the auditory stimuli (that is, a visual screen was not placed in front of the auditory loud speaker obstructing the participants perception) and an offset on the vertical plane is less sensitive to location offsets than the horizontal plane (Broadley & Kirkland, 1979). Further, a vertical offset of 30 mm should not be easily localised meaning the origins of the auditory and visual stimuli could be said to be coincident.
Participants used the laptop keyboard to key through the experiment instructions using the [spacebar] and to register identification of the target letter during the experimental trials (LHS and RHS [shift] keys) which were colour coded blue and red for ease of discrimination (see Figure 4.2 for a depiction of the stimulus presentations).

Following completion of the 18 (2 trials from each of the 9 experimental conditions) practice trials and 324 (36 trials from each of the 9 experimental conditions) experiment trials participants were asked to complete a 3-item ratings questionnaire on completion of an additional 18 trials: (i) a question asking participants to rate on a 7-point Likert scale (“few” to “lots”) the similarity of the target letter to the surrounding alphabetic letters that had been presented during the trial, (ii) a question asking participants to indicate on a similar 7-point scale (“not difficult” to “very difficult”) how difficult they felt the trial has been for them, and (iii) a question asking participants to select the statement most representative of their performance on the trial, including statements “missed both targets”, “missed at least one
target”, “identified both targets” or “I don’t know”. A full copy of this questionnaire can be viewed in Appendix B.

Figure 4.2. Depiction of the presentation of the auditory and visual stimuli: (a) Auditory/auditory (AA) condition; (b) Visual/visual (VV) condition, and; (c)+(d) Auditory/visual (AV) condition.

4.3.3 Stimuli and Pilot Studies

In this experiment participants were required to monitor two strings (a sequence) of alphabetic letters presented to them simultaneously. One string was presented to the LHS headphone (auditory) or the LHS of the computer monitor (visual), the other presented to the
RHS headphone (auditory) of the RHS of the computer monitor (visual). The participants were tasked with viewing the dual alphabetic letter strings and to press either the left or right [shift] key when the target letter was seen/heard on the LHS or RHS. In order to create the stimuli for use in the experiment it was necessary to run several preliminary investigations. Firstly, determination of the compression/stretching of all auditory stimuli to ensure presentation duration of all visual and auditory letters were matched; secondly, target and distractor letters had to be selected following similarity and dissimilarity ratings; Thirdly, development of letter string sequences and their temporal parameters; Fourthly, validation of the cognitive load manipulations. These pilot studies and preliminary investigations are set out below.

4.3.3.1 Alphabetic Letter Presentation and Duration

To select the auditory and visual alphabetic letters to be used in this and subsequent experiments a pilot test was conducted to determine the appropriate duration presentation of the auditory letters. Auditory presentation duration is important given the temporality of auditory stimuli, and, excessively compressed or stretched may appear unnatural and distorted to the listener while similar distortion may not be as apparent in the visual presentation of text-based letters. Auditory compression and stretching is regularly undertaken in such studies so that all auditory, and subsequently visual, stimuli can be presented for equal durations and thus avoiding any confounding effects of unequal stimulus presentation durations.

Past studies have unfortunately presented auditory letters for a number of different durations (Potter et al., 1998; Soto-Faraco et al., 2005; Jolicoeur, 1999; Arnell & Jolicoeur, 2002) and a number of these compression durations seem to leave a number of alphabetic letters unrecognisable to the researcher. Therefore, a standard presentation duration for the
alphabetic letter stimuli is not available, so a pilot study was conducted to determine the most appropriate duration for this, and subsequent, experiments.

In Pilot Study One (Auditory Letter Presentation Duration) (see Appendix C for a full report of the study) participants were required to rate the most appropriate auditory letter presentation duration from a number of possible letter presentation durations (56 ms, 80 ms, 115 ms, 150 ms, 175 ms, 190 ms, 300 ms) all of which have been used in previous studies. All 26 alphabetic letters (see Appendix A for a listing and pronunciation guide for all 26 letters) were compressed to these durations in a text-to-speech program (TextAloud 2.0) in an Australia male voice. Ten participants (6 female and 4 male, $M$ age of 26.1 years) were asked to listen to all letter recordings at each of the six durations and to name each letter. Participants were also asked to make any comments they felt pertinent about the presentation of each letter under each duration. The results of this pilot study suggested the letter “W” was not appropriate for compression – presumably due to its natural length being considerably longer than all other letters – and thus the quality of the recording and the identifiably of the letter “W” was compromised much sooner than other alphabetic letters, with participant comments suggesting it sounded strange, e.g., “like a chipmunk” (Participant 16) during nearly all compression durations. Once the letter “W” was removed from analysis it was revealed that presenting letters for 150 ms was the maximum compression that still yielded all remaining 25 alphabetic letters identifiable to participants (see Appendix D for recordings of all 25 letters at 150 ms durations).

For consistency, visual stimuli included all 25 alphabetic letters (minus “W”) determined appropriate in the auditory letter presentation duration determination in Pilot Study 1. All 25 visual letters were presented on the computer monitor in 24 point black Times New Roman font on an off-white background screen. At this size the letters subtended
approximately 1.0° of the visual angle. On the basis of the reasoning above visual letters were presented for the same duration as the auditory letters, 150 ms.

4.3.3.2 Selection of Target and Distractor Letters

For the dual letter identification task it was necessary to establish a target letter which was highly contrastive to a series of alternative ‘distractor’ letters, while highly similar to a series of alternative letters in order to facilitate evaluation of cognitive load based on target-distractor similarity. Further, it was decided that the target letter should be identical in both the auditory and visual identification tasks to reduce any task-switching costs (Potter et al., 1998), however, distractor letters would need to be different across modalities as similarity ratings would be based on the visual features of the letter in the text-based presentation, and on phonological features of the letters in the auditory-based presentation. To determine the appropriate target letter, and modality specific distractor letters a second Pilot Test was conducted.

In Pilot Study Two (Target-Distractor Similarity Ratings) (see Appendix E for a full report), participants were required to listen to all of the 25 auditory stimuli (randomised) and to then view all of the 25 visual text stimuli (randomised). Study participants (9 female and 6 male, $M$ age of 27.6 years) were fully briefed about the purpose of their task – in terms of how the value of their classification determinations of targets and distracters would be used for subsequent experiments. Participants were told they could view/hear the auditory and visual letters as many times as they wished, and as the auditory and visual stimuli were accessed on different computers they could switch between the computers as often as required in order to select a single alphabetic letter they believed would have highly similar and highly dissimilar counterparts in each of the auditory and visual modes. Once participants
had selected a single target letter they were then asked to concentrate on categorising highly similar and highly dissimilar letters to it in each the auditory and visual modalities separately.

Following this pilot testing “B” was selected as the target letter common to the auditory and visual modalities. For the auditory stimuli, the set “V, E, D, C and G” was identified as highly similar to the target “B” (thus, becoming the a highly similar auditory distractor letters) and “U, X, Z, I and L” as highly dissimilar to “B” (thus, becoming the highly dissimilar auditory distractor letters). For the visual stimuli, the set “F, E, D, P and R” were rated as highly similar to the target (thus, becoming the highly similar visual distractor letters), while “Y, J, T, N and M” were identified as being highly dissimilar to the target (thus, becoming the highly dissimilar visual distractor letters).

4.3.3.3 Developing the Auditory and Visual Letter String Sequences

With the target and modality-specific distractor letters determined the presentation sequences of these were developed. This involved determining letter strings consisting of 1 target and 7 distractor letters, which were then developed into pairs of either two auditory 8-letter strings (AA trials), two visual 8-letter strings (VV trials) or an 8-letter auditory and 8-letter visual strings (AV trials). Therefore, each letter string sequence consisted of 16 letters (2 targets and 14 distractors).

A letter sequence template was developed, from which the appropriate auditory and visual letter stimuli were input to satisfy the parameters of each specific trial. The letter-sequence template (see Figure 4.3) sees a single letter presented to the participant every 170 ms (i.e., 6Hz, hereafter presentation rate in ms) and the sequence continues via alternation between the participant’s LHS and RHS. For instance, the sequence may begin with one letter being presented to the LHS of the particular participant for 150 ms, this will be followed by a 20 ms period of silence, followed by a letter being presented to the RHS of the participant for
150 ms, followed by 20 ms of silence. This pattern is repeated until 8 letters are presented in each of the LHS and RHS. Presentation commencement was randomised between the LHS and RHS of each particular participant across all trials.

**Figure 4.3.** Template for letter presentation sequences for the left hand side (LHS) and right hand side (RHS) string pairings – top and bottom horizontal lines respectively. On both the LHS and RHS an auditory/visual prime will be presented at the 200 ms mark (duration of 150 ms) to ready participants for the trial commencement. Each letter (150 ms in duration) is represented above by the rectangular boxes rising from the horizontal lines. As can be seen, no more than one letter is presented at the same time. The dashed vertical line indicates the point at which the first target *may* appear. The break (~) shown towards the end of both horizontal lines indicates that a period of silence was been omitted from this visual depiction of the temporal parameters of the trial.
The target-distractor letter arrangement within this template will see the target letter only ever presented after the first 570 ms of the trial – to ensure participants have time to properly engage in the trial. Following the 570 ms mark the target may randomly appear in any of the following letter positions on both the LHS and RHS (that is, the target will be present in both letter strings). All other letter positions will then be occupied with a distractor letter, appropriate to the modality in question and appropriate to the cognitive-load level being examined. As such, within any single experimental trial a participant may be required to process two auditory letter strings (AA), each in rapid serial visual presentation (RSVP) arrangement, or two visual letter strings (VV), each in rapid auditory presentation (RAP) arrangement, or a single RSVP string on one presentation side, and an RAP string on the alternative presentation side (AV).

In the low-cognitive-load conditions the target letters will be surrounded by 7 highly dissimilar distractor letters (as determined in the Pilot Test 2), in the medium-cognitive-load conditions the target letters will each be surrounded by 3 highly similar and 4 highly dissimilar distractor letters, and in the high-cognitive-load condition the target letters will each be surrounded by 7 highly similar distractor letters. An example of the possible resulting AA, VV and AV trials within the letter sequence template can be seen in Figure 4.4. To validate the cognitive load manipulations Pilot Study 3 (Cognitive Load Manipulation Validation) asked participants to view a series of experimental trials where imposed cognitive load was manipulated using target-distractor similarity, and asked to rate on a 3-point Likert scale whether the distractor letters were highly similar or dissimilar to the target. As expected, the LCL condition trials were all rated as having dissimilar distractors, the MCL trials as having some similar and some dissimilar distractors and all the HCL trials as having similar distractors. It was concluded that the target-distractor similarity manipulation was appropriate for manipulating the imposed load in Study One.
Figure 4.4. Example experimental trials (a) AA trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load (b) VV trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load (c) AV trials (i) low cognitive load (ii) medium cognitive load (iii) high cognitive load.
4.3.4 Procedure

Participants were seated in front of the laptop computer in a small laboratory testing room, such that the computer screen was at approximately chest height and loud speakers were presented laterally coincident and slightly above where the visual letters (at 16° separation) would display (see Figure 4.1). The experiment was divided into several sections to ensure the participant did not become fatigued and to allow time-out periods for the researchers to address any technical problems encountered by participants prior to the experimental trials of interest being attempted. Participants first completed 18 practice trials (2 trials from each experimental condition) so that they would be familiar with the auditory and visual stimuli, the timing and progression of the experiment, and the required responses. The experimenter then reviewed their practice responses to ensure performance was appropriate (i.e., they were responding to both targets, not just one, and were responding immediately following the target presentation and not at the conclusion of the trial). After this brief evaluation procedure participants began the 324 experimental trials (36 trials from each of the 9 experimental conditions). None of the practice trials sequences appeared in the experimental trials.

The participants initiated each block of trials (including the practice trials, and subsequent experimental trials) by pressing the [space bar]. They were then presented with a series of hashes (#####) across the centre of the monitor for 150 ms at the beginning of each trial. Following the offset of the hashes, there was a blank (and silent) period of 100 ms, and then the series of 16 alphabetic letters (RSVP and/or RAP) were presented. Participants were instructed that when they saw/heard the target letter appear on their LHS to press the left [shift] key and when seen/heard on the RHS to press the right [shift] key. Shift keys were coloured coded blue and red respectively to aid discrimination. Inter-trial interval was set at 5 secs after which the next trials commenced automatically. After every 27 item trials,
participants were given a rest period which terminated once they pressed the [spacebar] to begin the next block of trials. In total participants completed 12 trial blocks (324 trials), which took approximately 45 mins.

Following the experimental trials participants were asked to complete a further 18 trials, identical to the 18 practice trials, each followed by a 3-item rating questionnaire (see Appendix B). The participant commenced each rating trial by pressing the [spacebar] and completing the trial, including the button presses as was required during the experimental trials. The program then presented participants with a screen instructing them to complete the pen and paper ratings questionnaire for that trial. When ready, participants pressed the [spacebar] to initiate the next trial. This procedure continued for all 18 rating items with this second stage of the experimental taking approximately 10 mins.

4.4 RESULTS

4.4.1 Data Reduction

Two sets of data were of interest; participants responses (RT and accuracy) and their ratings in the final 18 trials. For the participants’ response data cleaning involved any trial in which more than a single key press was recorded for each [shift] key. For all remaining trials the DMDX output was input by a batch script into a Microsoft Excel spreadsheet in which formulae were used to derive reaction times, in milliseconds (ms), from the onset of each target to the button press. Accuracy data were derived by coding the data to reflect valid responses – that is of the two key presses both were made within the acceptable time frame window (2-ID), where only one of the two was within the acceptable time frame window (1-ID), or were neither were made within the acceptable time frame (0-ID). In this study a valid response occurred when the correct [shift] key was pressed 150 ms-435 ms after the presentation of the target letter. This period for valid responses was established by established
by accepting a lower limit of 150 ms previously indicated by Spence and Driver (1996) to be the fastest possible reaction time from presentation onset to key press identification. The upper limit of 1,500 ms used in the Spence and Driver study was not appropriate here as a window of 150 ms-1,500 ms could mean that a single key press could in fact be a response to up to three possible letters being presented. Therefore, an upper limit of 425 ms was determined for a valid response. This limit was derived by reviewing the rating trials of all participants in which a confidence rating of “I am very confident I identified both targets” and taking a median measure of the RTs for these trials. The median for these trials across all participants and cognitive load levels was 372 ms (SD = 26.4) and the upper limit of valid responses (425 ms) was deemed to be 2 standard deviations above the median.

Each participant completed 36 trials in each of the 9 experimental conditions. Thus there were 9 average RTs calculated for each participant. This mean was determined by averaging only those trials in each condition in which two valid key presses (2-ID) had been made. Trials where one or no (1-ID and 0-ID) valid key presses were made were not included in the RT analysis. Similarly, 9 accuracy scores were determined as the percentage correct for the 2-ID trials in each condition.

Participant rating questionnaires were coded to allow quantitative analysis. As each of these trials consisted of 2 trials under each of the 9 experimental conditions, average ratings were derived. The first question, regarding task complexity involved 7-point Likert scale ratings so means were derived with a score of 1 representing low complexity and a score of 7 high complexity. Likewise for the second question, regarding difficulty, average difficulty was computed for each condition directly from the Likert scale rating. For the third question as responses were ordinal, rather than as intervals, the four categories were reverse scored so that responses of “I identified both targets” were classified as ‘highly confident of positive

---

1 The median, rather than the mean, was used as previous research (Spence & Driver, 1996; Ward, 1994) has suggested that means tend to be skewed in such tasks and that medians better reflects the central tendency.
performance’ (rating of 4), responses of “I think I identified one target” were classified as ‘somewhat confident’ (rating of 3), response of “I’m not sure” were classified as ‘unconfident’ (rating of 2) while responses of “I missed both targets” were classified as ‘very unconfident” (rating of 1).

4.4.2 Data Screening

Data screening using the Statistical Package for the Social Sciences (SPSS – v.15) was conducted. Inspection of participant RTs revealed that one participant had made single key-presses on every trial between the onset of the trial and the 602 ms time mark. Given that no target stimulus was ever presented prior to the 590 ms it was decided that a serious participation attempt had not been made and the data for this participant were removed. In addition, three outliers were identified during normality testing after raw scores were converted into z scores (Tabachnick & Fidell, 2008). In each of these participants had initially performed reasonably well, for approximately the first 20% of the trials. Although they then ceased to maintain concentration with key presses being discontinued after the fourth (two participants) and fifth (one participant) blocks of experimental trials. Despite these three participants re-engaging during the ratings items it was deemed appropriate to remove all these participants’ data to maintain consistency across the objective and subjective measures.

Following removal of these four outliers, all assumptions of normality were met although there did appear to be some level of skewness in the RT data. Nevertheless, this does not indicate an abnormal distribution as RTs tended to cluster toward the upper limit of the experimentally imposed 150-425 ms time frame, and so did not appear to reflect a natural tapering off of RTs. Therefore, this violation of the assumption of normality was not considered to be a concern (Hills, 2005).
Chapter 4 – Study One

4.4.3 Data Analyses Overview

In this and all subsequent analyses planned contrasts were tested for each factor. For modality combination the two orthogonal contrasts were within- vs. cross-modality conditions (AA+VV vs. AV), and auditory vs. visual within-modality conditions (AA vs. VV). For the 3-level cognitive load factor, linear and quadratic trends were tested to determine any changes in performance across the three cognitive load levels. See Appendix F for details of the planned contrasts conducted.

A repeated measures 3 x 3 factorial ANOVA, with alpha set at .05, was conducted to compare performance across all nine experimental conditions for both RT and accuracy scores. Additional 3 x 3 factorial ANOVAs were conducted, using the same planned contrasts, to evaluate each of the 3 ratings questions.

4.4.4 Analyses

4.4.2.1 Performance: Modality Combination and Cognitive Load

Reaction Time. Reaction time data are schematically presented in Figure 4.5. Despite some indication of an advantage for cross- vs. within-modality conditions, especially in the low cognitive load conditions, there was no linear or quadratic effect for either modality combination ($F_{AV vs. AA+VV}(1,46)=.63, p=.43, \text{partial } \eta^2=.01; F_{AA vs. VV}(1,46)=.19, p=.67, \text{partial } \eta^2=.01$) or cognitive load ($F_{Lin}(1,46)=.01, p=.97, \text{partial } \eta^2=.01; F_{Quad}(1,46)=.00, p=.98, \text{partial } \eta^2=.00$). Nor were any interactions between these two conditions significant. This indicates that reaction times to the target presentation were similar across all three modality combinations (AA, VV and AV), and that increases in cognitive load did not affect the speed of these target identifications. Additionally, the standard errors for all modality combinations are substantial, however, of interest is that the standard error for the cross-
modal condition is considerably smaller than for the two within-modal conditions – again suggesting there is some benefit to cross-modal processing despite the non-significant results.

Figure 4.5. Mean participant RTs (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).

Accuracy: Accuracy data are schematically presented in Figure 4.6. Inspection of Figure 4.6 and statistical analyses show that cross-modal (AV) presentation generally facilitates more accurate performance than does within-modal (AA or VV) presentation ($F_{AV}$ vs. $AA+VV(1,46)=26.22, p<.001$, partial $\eta^2=.36$) while there is no significant difference between the two within-modal presentation styles which facilitate similar levels of accuracy performance ($F_{AA\,\text{vs.}\,VV(1,46)}=.60, p=.44$, partial $\eta^2=.01$).
There were no significant overall linear or quadratic effects for cognitive load suggesting that either the changes in the task did not appropriately impact cognitive load and in turn performance, or that increases in the level of load were not matched by an inverse decrease in accuracy performance ($F_{Lin}(1,46)=1.23, p=.27$, partial $\eta^2=.03$; $F_{Quad}(1,46)=.39$, $p=.54$, partial $\eta^2=.01$).

No interactions between modality combination and cognitive load level were evident in these accuracy data. In particular the AV vs. AA+VV contrast did not interact with cognitive load and so the cross-modal advantage is consistent over all load conditions.

![Figure 4.6. Mean participant accuracy percent scores (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).](image-url)
4.4.2.2 Complexity, Difficulty and Confidence Ratings

Task Complexity. A depiction of the task complexity ratings can be seen in Figure 4.7. Planned contrast analyses of modality combination and cognitive load indicated that despite the complexity of material presented cross-modally appearing to be somewhat lower than the complexity of the material when it is presented within-modally no significant effect was in fact seen ($F_{AV \text{ vs. } AA+VV}(1,46)=.41, p=.35$, partial $\eta^2=.02$). Additionally, there is little difference in the level of perceived complexity between when two auditory tasks or two visual tasks are completed simultaneously ($F_{AA \text{ vs. } VV}(1,46)=.31, p=.58$, partial $\eta^2=.01$). There was a general linear, but no quadratic effect of cognitive load, suggesting that perceived complexity increased in a steady and linear fashion across the increasing levels of cognitive load ($F_{Lin}(1,46)=34.70, p<.001$, partial $\eta^2=.43; F_{Quad}(1,46)=.22, p=.64$, partial $\eta^2=.01$). As there was no interaction between modality combination and cognitive load conditions the linear trend seen for cognitive load is consistent across all modality combinations.

![Figure 4.7](image.png)

Figure 4.7. Mean ratings of cognitive complexity (+SE) for each of the modality combination and cognitive load conditions.
Task Difficulty. The relative ratings of difficulty for each of the experimental conditions can be seen in Figure 4.8. While it might appear that cross-modal presentation was associated with lower levels of perceived difficulty than within-modal presentation, and further, that over increasing levels of cognitive load perceived difficulty increases, statistically speaking there were no significant modality combination ($F_{AV vs. AA+VV}(1,46)=1.62, p=.21, \text{ partial } \eta^2=.03$) ($F_{AA vs. VV}(1,46)=3.58, p=.02$) or cognitive load ($F_{Lin}(1,46)=1.66, p=.20, \text{ partial } \eta^2=.04$) ($F_{Quad}(1,46)=.06, p=.80, \text{ partial } \eta^2=.01$) effects, nor any interaction between the two. This suggests that participants did not perceive any of the modality combinations conditions to be more difficult that the others. More importantly, the three increasing cognitive load conditions were not associated with corresponding increases in ratings of difficulty suggesting that the cognitive load manipulations did not in fact lead to a change in the experience of this load.

Figure 4.8. Mean ratings of task difficulty (+SE) of all modality combination (AA, VV and AV) and perceived cognitive load (low, medium and high) conditions.
Confidence. Mean ratings of confidence depicted in Figure 4.9. As can be seen there was no evidence that cross-modal perceived confidence was different to within-modal performance ($F_{AV vs. AA+VV}(1,46)=2.35, p=1.33, \text{ partial } \eta^2=.05$), nor were there any differences between perceptions of confidence in the two within-modal conditions ($F_{AA vs. VV}(1,46)=1.00, p=.32, \text{ partial } \eta^2=.02$). Likewise no differences were seen in perception of confidence across the cognitive load conditions ($F_{Lin}(1,46)=1.69, p=.20, \text{ partial } \eta^2=.04$; $F_{Quad}(1,46)=.26, p=.61, \text{ partial } \eta^2=.01$). No interactions between modality combination and cognitive load were evident in perceptions of confidence. These findings suggest that confidence is not affected by how the stimuli were presented to the participant (AA, VV or AV) and the increasing levels of cognitive load did not impact perceptions of confidence either generally, or as a function within any modality combination condition.

![Confidence Rating Chart]

*Figure 4.9.* Mean ratings of task confidence (+SE) of all modality combination (AA, VV and AV) and perceived cognitive load (low, medium and high) conditions.
Of interest on 5% of AV trials, 2% of VV trials and 1.5% of AA trials ratings of ‘high confidence’ were matched with performance data where only one valid response was recorded. Moreover, on virtually every one of these trials (98.64%) a key press was recorded within 100 ms of the valid time frame window concluding. That is, a key press – coded as invalid – was actually made somewhere between 426 ms and 525 ms after the target was presented.

4.5 DISCUSSION

It was hypothesised that there should be a cross-modal performance advantage (i.e., shorter RTs and increased accuracy in AV compared to AA and VV) and that any such cross-modal advantage would diminish as the level of cognitive load increased. This interaction hypothesis was not supported in this experiment. However, it would be premature, for a number of reasons, to conclude that the results here support the predictions of a multiple-resource model.

Firstly, while RT data here are similar across all conditions it is possible that the prescribed time window in this experiment was too short and created a truncated RT distribution. This claim is supported by the confidence ratings which show there are a number of cases where data was coded as invalid due to only one key press occurring in the time window when in fact the participant had made a response only just outside the window – but well within the limits suggested by Spence and Driver (1996) – and were highly confident of identifying both targets. Including these responses may well reduce the skewness of the data and in turn affect the RT analysis outcomes particularly given that such a large percentage of these ‘invalid’ responses were associated with cross-modal trials. Moreover, Figure 4.6 does suggest some modality-combination x cognitive load effect even though this effect was not significant.
Secondly, in terms of performance accuracy there was a small but low (Tabachnick & Fidell, 2007) cross-modal advantage. A similar weak cross-modal advantage was demonstrated by Spence and Driver (1996) however, in their subsequent work this weak advantage became stronger as features of their experimental design altered (Soto-Faraco & Spence, 2002; Spence et al., 2000). In the earlier study (Spence & Driver, 1996) the auditory and visual stimuli were presented from spatially coincident locations. Under such conditions participants would be placed under significant load to discriminate between the stimuli whether cross- or within-modality because physical separation aids in discriminability (Brungart & Simpson, 2002) and intelligibility (Morrison, Burnham & Carlile, 2010 – see Appendix G). However, in their later experiments (Soto-Faraco & Spence, 2002; Spence et al., 2000) spatial separation was much larger – a condition which would aid discriminability and in turn reduce cognitive load. As such, the weak cross-modal advantage in this experiment may well be due to the cognitive load of the experiment being moderate (rather than extremely high or extremely low) and most importantly not varying across conditions as anticipated.

In this regard the cognitive load manipulation in this experiment did not appear sufficiently substantial to impact performance (null findings across both RT and accuracy data) or ratings of cognitive load (null finding for difficulty ratings). In line with the experimental design of Spence and Driver (1996) the two stimulus sequences were presented from spatially non-coincident locations thus perhaps reducing cognitive load. As this level of cognitive load was maintained across all conditions both in terms of the experimental

---

2 In this study ratings of complexity did not appear to correlate with performance as well as ratings of difficulty; like the performance data ratings of task difficulty were consistent across all levels of experimentally imposed load. Therefore, the terminology introduced in Chapter 2 will be maintained (for the sake of simplicity) – that is, cognitive load can be imposed by the features of the task, but ultimately cognitive load is actually a perception of the difficulty of those task features as experienced by the individual. However, for all future studies the relationship between the imposition of load and the experience of load will continue to be evaluated to ensure that the experimental manipulations of load is actually valid.
manipulations and in participants’ ratings of load then it is of little surprise that a significant cross-modal advantage sans cognitive load effect was found.

It appears then that a better manipulation of cognitive load is required. In Pilot Study Two (see Appendix E) participants’ significant differences in ratings of the similarity between the targets and distractors across the three levels of imposed cognitive load (in their ratings of task complexity) was believed to be evidence of perceived differences in cognitive load. However, here in Study One the task difficulty ratings seemed to indicate that these similarity relationships did not in fact impact the participants’ experience – that is, target-distractor (dis)similarity did not affect the experience of load.

A review of the data here do indicate that a cognitive load difference may have been emerging; RTs appeared fastest across all modality combinations in the low load condition, and slowest in the high load conditions and accuracy appeared best in the low load conditions and worst in the high load conditions. Thus, it appears that the target-distractor similarity manipulation did have some effect but not great enough to significantly impact performance. Therefore, future experiments should perhaps consider the use of other manipulations of cognitive load. One such alternative may be a manipulation of the number of items to be processed within a given time frame, a suggestion made by Woods (1988) as a feature of a task that is likely to impact overall task difficulty. Similar timing manipulations have been used to influence cognitive load in a number of previous studies (Morrison, Wiggins & Porter, 2010). Therefore, this timing manipulation will be used in Study Two in order to manipulate cognitive load.

With respect to time, the confidence ratings show a proportion of trials may have been excluded from analysis by not meeting the RT time window criterion (150 ms to 425 ms) where in fact participants were certain that they accurately identified the targets. This RT window was considerably shorter than in previous experiments (e.g., Spence & Driver,
The ratings data compared to the performance data suggest that extending the window to an upper limit of 525 may be worthwhile. While such a time frame of 150 ms to 525 ms would allow key press responses to represent more than one letter presentation, in this experiment no participant actually responded prior to the 205 ms mark. Accordingly, extending the upper limit may be less problematic if the lower limit could be similarly adjusted.

While the hypothesis set out at the beginning of this chapter was not supported this may be due to an unsuccessful attempt to manipulate cognitive load. Study Two will employ a ‘time’ manipulation of cognitive load, a manipulation consistent with Woods’ (1988) conceptualisation of cognitive complexity. A time manipulation allows for an objective qualification of the cognitive complexity of the task and therefore a relative gauge of intrinsic cognitive load. Additionally, it has been shown that such a time manipulation of cognitive load impacts both performance and participant experience (Szalma, Hancock & Quinn, 2008).
STUDY TWO\textsuperscript{1}

MANIPULATING COGNITIVE LOAD WITH TEMPORAL RATE OF STIMULUS PRESENTATION

\textsuperscript{1} This study has been submitted for publication (Morrison, Burnham, Wiggins & Dean, 2010) (see Appendix H).
5.1 RATIONALE

The aim of this study is identical to that of Study One: to examine whether support for both the multiple- and unitary-resource models can be found when levels of cognitive load are manipulated. However, as highlighted in the discussion at the conclusion of Study One a temporal rate of stimulus presentation timing manipulation may be more effective than the target-distractor similarity mode used in that study, and such a timing manipulation was employed here.

Consistent with those proposed in Chapter 4 – Study One the hypotheses in Study Two here are:

Hypothesis 1: A cross-modal performance advantage (i.e., shorter RTs and increased accuracy in AV compared to AA and VV) will be present and it will diminish as the level of imposed cognitive load increases. The point at which any cross-modal advantage will disappear cannot be predicted at this stage.

Hypothesis 2: Ratings of complexity will increase as imposed cognitive load level increases.

Hypothesis 3: Ratings of difficulty will increase as imposed cognitive load level increases.

5.2 EXPERIMENTAL DESIGN

The design of this study is identical to that in Study One with the exception of the method of cognitive load manipulation. Again a dual-task letter identification task will be used within a 3 (modality-combination) x 3 (cognitive-load) within-subjects factorial design. Two independent variables were of interest. The first, modality combination, has three levels; (i) two auditory tasks - AA, (ii) two visual tasks - VV and (iii) one auditory and one visual
task – AV. The second, cognitive load, also has three levels; (i) low cognitive load (LCL) (ii) medium cognitive load (MCL) and (iii) high cognitive load (HCL).

A response alternative rate related manipulation of cognitive load (Woods, 1988) introduced in this study. While the trial duration remained at 4,500 ms, the number of distractors processed in this period altered between cognitive load levels. The LCL condition trials consisted of a left and right hand side (LHS and RHS) alphabetic string each containing one target “B” and two auditory or visual distractors. In the MCL condition trials the LHS and RHS strings each contained one target and five distractors and in the HCL condition both the LHS and RHS strings consisted of one target and eleven distractors.

The dependent variables of interest were objective performance measures of reaction time (RT) and percent accuracy, and the subjective measures of cognitive complexity, difficulty, and confidence as in Study One.

5.5 Method

5.4.1 Participants

There were 25 participants, 15 males and 10 females, with a mean age of 27.8 ($SD=6.2$) for males and 25.3 ($SD=3.8$) for females. All participants reported normal hearing and normal (or corrected-to-normal) visual acuity. Participants provided informed consent and received $50 reimbursement for their travel.

5.4.2 Materials and Apparatus

As in Study One, the experiment was conducted on a LG 17” widescreen laptop with attached loud speakers (Genelec Analogue 50/60Hz), with stimulus presentation and participant responses run and recorded in DMDX software (Forster & Forster, 2003). Auditory stimuli were presented to participants via the loud speakers, which were vertically
offset (by 30 mm) from the presentation of the visual stimuli at the top of the computer screen.

Again, participants used the [spacebar] to key through onscreen instructions and the left (colour coded blue) and right (colour coded red) shift keys to indicate identification of the target letter presentations during each trial.

A 3-item ratings questionnaire was included after each of the final 18 trials in the experiment. These items remained the same as those in Study One: (i) ratings of cognitive complexity (ii) ratings of trial difficulty and (iii) an indication of trial performance. A full copy of this questionnaire can be viewed in Appendix B.

5.3.3 Stimulus Materials

The auditory and visual letters selected as the target and distractor stimuli in Study One were reused in this experiment. As such, the letter “B” was retained as the target letter in both the auditory and visual stimulus strings. Given that the distractor letters did not adequately impact cognitive load in Study One, it was decided to retain all of these (Auditory: V, E, D, C, G, U, X, Z, I and L; Visual: F, E, D, P, R, Y, J, T, N and M) given the large number of distractor letters required in this study (a maximum of 11 here, compared to 8 in the previous).

With cognitive load being manipulated by the temporal rate of stimulus presentation in this study a template of timing sequences was developed to represent the three different cognitive load levels. In the LCL condition one target letter was presented on each of the LHS and RHS in the presence of two distractor letters on each presentation side. In the MCL condition one target letter was presented on each the LHS and RHS in the presence of five distractor letters on each presentation side. Finally, in the HCL condition on target letter was present on each the LHS and RHS in the presence of 11 distractor letters on each presentation
side. Within each presentation side the individual letters were evenly spaced across the duration of the trial while across presentation sides the letters were temporally offset so that no stimuli were presented at the exact same time. Therefore, with each letter being presented for a 150 ms duration, this meant that in the LCL condition a letter was presented every 660 ms (resulting in an inter-stimulus-interval (ISI) of 510 ms), in the MCL condition a letter was presented every 480 ms (ISI = 330 ms) and in the HCL condition a letter was presented every 165 ms (ISI = 15 ms). A graphical representation of these three timing sequences is shown in Figure 5.1.

The auditory and visual letter stimuli were placed in these templates to create two rapid auditory presentation (RAP) strings (in the AA condition), two rapid visual serial presentation (RSVP) strings (in the VV condition) and then one RAP and one RSVP string (in the AV condition) to formulate string pairing stimuli in much the same fashion as in Study One. An example of the possible resulting AA, VV and AV trials within the letter timing sequence template can be seen in Figure 5.1.

Single-task trials were included in this experiment to provide a baseline measure of participant performance. In any cases where participants failed to record two valid responses within the predetermined RT window all their data was excluded from further analysis.

5.3.4 Procedure

As in Study One, participants were seated in front of the laptop computer in a small lit laboratory testing room, so that the computer screen and accompanying auditory loud speakers were at approximately chest height. Firstly, the participants completed 18 practice trials so that the participant would be familiar with both the auditory and visual stimuli, the timing sequences, progression of the experiment, and the responses they were required to make during each trial before commencing the experimental trials. Following the practice
Figure 5.1. Template for letter presentation sequences under: (a) Low cognitive load (as shown for an AA trial). (b) Medium cognitive load (as shown for a VV trial). (c) High cognitive load (as shown for an AV trial).
trials the experimenter reviewed their responses to ensure their performance was appropriate (i.e., they responded to two targets, not just one and they had responded immediately following target presentation and not at the conclusion of the trial). After this short evaluation procedure participants began the experimental trials.

The experiment was organised in experimental blocks, with 29 trials present in each block. Participants initiated the beginning of each block by pressing the [spacebar] and with each subsequent trial beginning automatically 10 seconds after the previous one concluded. At the conclusion of each trial block, participants were given a break of self-determined duration and pressed the [spacebar] to commence the next experimental block. In total participants completed 12 trial blocks (348 trials – 324 dual-task trials and 24 single-task trials) taking approximately 50 minutes to complete.

As in Study One, participants completed a further 18 rating trials – the same as the practice trials – after which they completed a section of the questionnaire. This second stage of the experiment took approximately 10mins.

5.5 RESULTS

5.4.1 Data Reduction

Data reduction for the reaction time (RT) and accuracy data, and the complexity, difficulty and confidence ratings data was the same as it was in Study One.

5.4.2 Data Screening

Firstly, participant single-task trial were screened for sub-acceptable performance (<80%). The data for three participants were excluded on this basis with one participant failing to make any key presses at all, and the other two participants recording accuracy scores on these items at 49% and 72%.
Following removal of these data no univariate or multivariate outliers were apparent and the data for the remaining 22 participants was retained. All assumptions of normality were met although, as in Study One, there again appeared to be skewness present in the RT data. This skewness appeared to be smaller than that evident in Study One and it was not thought to be a serious violation of normality (Tabachnick & Fidell, 2007).

5.4.3 Data Analyses Overview

In this and all subsequent analyses planned contrasts were tested for each factor in the same way they were performed in Study One. See Appendix F for an outline of the planned contrasts conducted.

5.4.4 Analyses

5.4.4.1 Performance: Modality Combination and Cognitive Load

Reaction Time: As can be see in Figure 5.2, there appears to be no differences between the speed of processing over modality conditions. The analyses bear this out; there were no significant differences between cross- and within-modal conditions ($F_{AV vs. AA+V1}(1,24)=3.12, p=.09, \text{ partial } \eta^2=.12$) or between the two within-modal conditions ($F_{AA vs. V1}(1,24)=.07, p=.80, \text{ partial } \eta^2=.01$). Further, the results show that increasing levels of cognitive load does not alter the speed of processing ($F_{Lin}(1,24)=.07, p=.80, \text{ partial } \eta^2=.01$; $F_{Quad}(1,24)=.43, p=.52, \text{ partial } \eta^2=.02$). No interactions between modality combination and cognitive load were evident.

Taken together these results suggest that participants’ reaction times to the presentation of the target stimuli were similar across all experimental conditions in this experiment. This could indicate that, at least with the task used here, changes in stimulus presentation do not affect reaction time, or that the valid reaction time windows (that is, the
truncated RTs) imposed in this experiment did not allow for a complete evaluation of the longer reaction times of participants.

Individual differences were also evident in the data, with 23% of participants responding on the majority of trials within the first 100 ms, and where responses were not made inside this 100 ms they in fact did not respond at all during the trial in 92% of cases. In contrast, a further 17% of participants consistently responded within the final 100 ms and often times made responses just outside of the pre-established valid time frame.

**Accuracy Performance:** As can be seen in Figure 5.3, AV performance is consistently better than AA or VV performance; generally cross-modal presentation is associated with more accurate performance than within-modal presentation, while no differences in accuracy
can be seen between the two within-modal conditions ($F_{AV vs. AA+VV}(1,24)=42.77, p<.001$, partial $\eta^2=.64$; $F_{AA vs. VV}(1,24)=3.63, p=.07$, partial $\eta^2=.13$). In terms of cognitive load there is a linear, but not a quadratic, decrease in accuracy performance across the increasing levels of cognitive load ($F_{Lin}(1,24)=146.57, p<.001$, partial $\eta^2=.86$; $F_{Quad}(1,24)=2.15, p=.16$, partial $\eta^2=.08$). As such, this indicates that generally as cognitive load increases accuracy performance steadily decreases.

An interaction between cross- versus within-modal and the linear cognitive load trend is also seen with cross-modal performance decreasing more rapidly across the increasing levels of cognitive load compared to the shallower decline in within-modal performance across the same levels of cognitive load ($F_{AV vs. AA+VV*Lin}(1,24)=14.18, p=.001$, partial $\eta^2=.37$).

Figure 5.3. Mean participant accuracy percent scores (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).
This interaction shows that while performance for AA and VV trials steadily decreases across the load conditions, in the AV condition performance remains quite consistent between the LCL and MCL conditions, but quickly declines between the MCL and HCL conditions. In fact, the overall reduction in performance for cross-modal processing is substantially larger than the overall reduction in performance in the two within-modal conditions.

Given the individual differences apparent in the RT data it follows that for some participants who regularly made responses just outside of the valid response window accuracy scores were somewhat lower than other participants. There was no evidence to suggest that those reactions made just outside the valid window were more likely to belong to any of the modality combination or cognitive load conditions more than others. As such any individual differences seen were felt equally across all experimental conditions.

5.4.4.2 Task Complexity and Task Difficulty

Task Complexity. Illustrated in Figure 5.4 there are no differences in perceptions of complexity for the cross- and both within-modal conditions ($F_{AV, AA+VV}(1,24)=.61, p=.44$, partial $\eta^2=.03$; $F_{AA, VV}(1,24)=.03, p=.87$, partial $\eta^2=.01$). This suggests that participants did not appraise the complexity of cross- and within-modal presentation formats differently.

A significant linear and quadratic effect for cognitive load ($F_{Lin}(1,24)=1925.39, p<.001$, partial $\eta^2=.99$; $F_{Quad}(1,24)=71.46, p<.001$, partial $\eta^2=.75$) indicates that perceived complexity generally increases as cognitive load increases and more specifically that this increase is initially quite steep (between LCL and MCL conditions) before it begins to decelerate and flatten somewhat (between MCL and HCL conditions). Given that no interactions were evident the results suggest that increases in rated complexity across cognitive load levels was consistent for all modality combinations.
Figure 5.4. Mean ratings of task complexity (+SE) across the three modality combination and three cognitive load levels.

Task Difficulty. As graphically illustrated in Figure 5.5 cross-modal presentation is generally associated with lower levels of perceived difficulty than is within-modal presentation ($F_{AV \ vs. \ AA+VV}(1,24)=370.45, \ p<.001, \ partial \ \eta^2=.94$), while presentation of information in either AA or VV format results in similar perceptions of difficulty ($F_{AA \ vs. \ VV}(1,24)=.07, \ p=.79, \ partial \ \eta^2=.01$). There is also an overall linear, but not quadratic, effect for cognitive load indicating that across conditions, as cognitive load increases, a matched steady increase in perceived difficulty is evident ($F_{Lin}(1,24)=1366.51, \ p<.001, \ partial \ \eta^2=.98$; $F_{Quad}(1,24)=3.98, \ p=.57, \ partial \ \eta^2=.14$).
Nevertheless, while no overall quadratic trend was apparent for cognitive load there was a cross versus within-modality quadratic effect ($F_{Cross\ vs.\ Within*Quad}(1,24)=69.37, p<.001$, partial $\eta^2=.74$). This inspection of Figure 5.5 shows different and opposing quadratic trends for the cross- and within-modal conditions: in the cross-modal condition as cognitive load increases perceptions of difficulty at first only minimally increase (between LCL and MCL) before a sudden and substantial increase in perceived difficulty (a U-shaped quadratic trend with a strong linear component); whereas for within-modal conditions there is a relatively smooth increasing linear effect with what little quadratic component there is having an inverted-U-shape.
5.4.4.3 Task Confidence: Establishing Appropriate RT Windows

Results are graphed in Figure 5.6 confidence ratings for cross-modal and within-modal across the three cognitive load levels. The results indicate that generally there is no difference between cross-modal and within-modal ratings of confidence \((F_{AA+V1}(1,24)=2.62, p=.12, \text{ partial } \eta^2=.10)\), however, a difference can be seen between the two within-modal conditions \((F_{AA+V1}(1,24)=12.42, p=.002, \text{ partial } \eta^2=.34)\). This indicates that confidence in dual-task visual-only conditions is significantly superior to dual-task audio-only, and most likely to audio-visual processing also.

![Graph showing confidence ratings](image)

**Figure 5.6.** Mean ratings of confidence for the three modality combinations and three cognitive load conditions.
Significant linear and quadratic effects are also evident for cognitive load $(F_{Lin}(1,24)=162.21, p<.001, \text{partial } \eta^2=.87; F_{Quad}(1,24)=6.40, p=.018, \text{partial } \eta^2=.21)$ suggesting that as cognitive load increases perceived confidence decreases, and that initially this decrease is more subtle (from LCL to MCL) before declining more sharply (from MCL to HCL). An interaction between the two within-modal conditions and cognitive load is also evident $(F_{AA \text{ vs. } VV*Quad}(1,24)=5.68, p=.025, \text{partial } \eta^2=.19)$ whereby confidence in visual-only processing is initially higher than for auditory-only processing, but as cognitive load continues to increase this benefit begins to erode and confidence dips sharply to approach a similar level to that of AA and AV processing at high levels of cognitive load.

As in Study One there appeared to be a number of trials where participants indicated high confidence in identifying both targets. However – one, or both, of their key presses fell just outside the RT window established at the outset of this study following adjustments made in Study One. It appeared across all trials in this study that the same six participants were responsible for this mismatch between confidence and performance in at least 90% of cases. While for the majority of participants the 205-425 ms window appeared reasonable, for a few it appeared a more appropriate RT window would be between 235 ms and 455 ms.

5.5 DISCUSSION

It was hypothesised at the outset of this chapter that cross-modal performance advantages (indicated by shorter RTs and increased accuracy) should be present when individuals are performing under conditions of low cognitive load, and that any such cross-modal advantage would diminish as the level of cognitive load increased. These hypotheses were supported in this study, only for identification accuracy, RT measures did not provide any useful indication of performance variance.
Specifically, the results demonstrate that under conditions of low cognitive load cross-modal accuracy was close to ceiling at over 94%, while at a same level of cognitive load within-modal performance was substantially lower (80.16-81.22%). Such a superior outcome for cross-modal performance is consistent with the expectations of a multiple-resourcing account of attention. However, under conditions of high cognitive load, while cross-modal performance remained superior to within-modal performance (a ‘cross-modal advantage’) the cross-modal advantage was no longer statistically significant – an outcome inconsistent with the predictions of a multiple-resourcing account (Alais et al., 2006; Wickens, 2002; Duncan et al., 1997). Instead, the results suggest that, under higher levels of cognitive load, percent correct performance is similar irrespective of whether two tasks are undertaken in the same or different modalities – an outcome consistent with the claims of unitary-resource supporters (Soto-Faraco et al., 2005; Arnell & Jenkins, 2004; Jolicœur 1999).

While it was anticipated that a cross-modal advantage would be evidenced by increasing percent accuracy accompanied with decreasing reaction times this prediction was not supported. While heightened cross-modal accuracy was seen to be superior to within-modal accuracy there appeared to be no significant difference in RTs. This outcome does not obviate the claim that a cross-modal advantage was demonstrated under conditions of lower cognitive load as due to the small reaction time window used to define valid responses it may well be that any valuable differences between cross- and within-modal processing are evident only when a full spectrum of RTs are evaluated – as in Spence and Driver (1996). Moreover, numerous attentional resource studies have used accuracy scores, in the absence of RT data, to measure performance in cross- and within-modal dual task scenarios (Sinnet et al., 2006; Soto-Faraco et al., 2005; Potter et al., 1998).

Given the same outcomes in this study, as in Study One with respect to RT data, and given that the methods used here will be maintained in the forthcoming studies, it see ms
fruitless to continue to assess RT data given the null findings to date. Therefore, with the understanding that performance can be inferred from accuracy data (Spence, 2001) and that such measures of performance are used in similar studies that investigate attention (Alais et al., 2006; Soto-Faraco et al., 2005; Arnell & Larson, 2002; Spence et al., 2000; Jolicœur, 1999) performance accuracy will be the principal indicator of performance throughout the following studies in this thesis.

In a similar fashion to the relationship between cognitive complexity and cognitive load demonstrated in Study One, the results of this study further show why having both objective and subjective measures of cognitive load is essential. In this study participants were able to accurately reflect on the cognitive complexity of the individual trials; their ratings of complexity coincided well with the manipulations of cognitive load in the study. However, as could be seen here while the experience of load (as measured by rated task difficulty) was reflective of cognitive complexity (ratings of task complexity) – that is, heightened ratings of complexity was generally associated with high levels of rated difficulty – this relationship was not perfect or linear. For instance, when comparing within- and cross-modal ratings of task complexity and task difficulty. In the within-modal condition the medium cognitive load manipulation appeared to moderate equally between the low and high cognitive complexity measures and likewise somewhere equidistant between the low and high difficulty measures. In contrast, in the cross-modal condition the medium condition moderated equally between low and high measures of complexity – but in terms of perceived difficulty participants rated it more like the low than the high load condition. These outcomes suggest that cognitive complexity does provide a neat manipulation of load, however the relationship between task complexity (and thus the experimental manipulation) and the experience of difficulty (the cognitive load) are not perfect and therefore it is essential to
always take both objective and subjective measures to ensure the experience of load is indeed being examined.

The confidence ratings in this experiment provided an avenue by which to evaluate the RT window used in this study to define valid trial responses. With several participants being identified as requiring slightly different RT windows to others (205-425 ms vs. 235-455 ms) it sees necessary in future studies to institute personalised RT windows for each individual. Such RT windows are regularly used in the literature given the substantial individual differences often found in the attentional field particularly where speeded motor responses are required (Tipper & Baylis, 1987). The fact the some participants did react at repeatedly at different speeds to others suggests alternative factors might be responsible for the individual differences evident in both reaction time and accuracy data.

The individual differences could result from some participants being more motivated than others at the outset with anecdotal evidence suggesting that a number of participants who performed poorly had been pre-identified by the researcher as appearing disinterested prior to them commencing the experiment. Additionally, the role of motivation was implicated as a number of participants made comment at the conclusion of their participation that they had received no feedback along the way and as such felt their performance deteriorated. Two participants commented that the lack of any feedback, beyond the experimental trials, on their performance lead to them making slower responses – comments that possibly reflect these participants responding just outside the valid time frame – as they were unaware of the parameters of a “good responses”. While a further three participants felt their performance deteriorated as a result of the researcher not evaluating their performance causing them to lose interest and the desire to continue performing their best.

In addition to motivation it is possible that task familiarity may have impacted performance as a few participants in this Study had previously participated in Study One –
although at the outset of this experiment the target samples were drawn from seemingly different populations. These repeat participants were recorded by the researcher and it was revealed that in general their performance was high. As data from Study One were de-identified it was not possible to review whether that had performed as highly then, or whether significant improvements were evident from their original exposure in Study One to their participant in this study.

Feedback (Kluger & DeNisi, 1996), motivation (Dweck, 1986) and practice (Collie, Maruff, Darby & McStephen, 2003) are all active features of learning theory. Learning involves acquiring, enhancing, or changing of one’s knowledge and skills (Illeris, 2000; Ormorod, 1995). It is thought that learning new cognitive practices or ‘short cuts’ can be achieved when a person is either intrinsically (interest in personal non-tangible gains) or extrinsically (the opportunity to gain a tangible asset) motivated, when they are provided with feedback about their current progress, and particularly when they are repeatedly exposed to a particular task as occurs when gaining expertise in any given area. The introduction of ‘short cuts’ when completing a task is thought to be an information management strategy which can reduce cognitive load (Morrison et al., 2010; Gigerenzer & Goldstein, 1996). Therefore it makes intuitive sense that those people who did perform well included individuals who either appeared to be highly motivated at the outset of their participation or those who had participated previously and heightened performance is likely a result of reduced cognitive load – as is the central proposition of this thesis. As such, subsequent studies will need to review these features to determine whether the individual differences noted in this and the previous study may be accounted for by learning principles.

In summary, this study provides strong evidence that both models can be supported in the one experimental design, suggesting that the two mechanisms may actually coexist. One such explanation which may account for such a co-existence is that attentional resources are
arranged in a uni-modal-plus-supramodal (Spence & Driver, 1996), staged (Arnell & Duncan, 2002), or hierarchical (Posner, 1990) framework. The proposition of such a hierarchical model is not novel in the literature (Burr & Alais, 2006; Schwartz, Berthommier, & Savariaux, 2004; Spence & Driver, 1996; Posner, 1990; Posner, 1980). However, the interactivity or relationship between the two stand-alone models has not been investigated until now. Certainly the results of this study suggest a somewhat modality-specific allocation structure at low levels of load (cross-modal advantage present), yet a more modality-independent network as the levels of cognitive load increase. However, how these two structures interact remains somewhat ambiguous at this point. This interaction will be further addressed in Chapter 7 – Study Four when additional cognitive load levels and tri-task scenarios are evaluated in order to contrast two possible interacting structures.

Before any evaluation of the hierarchical model proposal is undertaken (in Chapter 7) it is first necessary to evaluate the individual differences apparent in this experiment. Determining the underlying nature of these individual differences is essential to understanding just how cognitive load – in its extraneous, intrinsic and germane for ms – impacts performance. So far both extraneous and intrinsic load have been identified as impacting cognitive load given the apparent modality-combination and cognitive load effects (and their interactions) demonstrated in this study. Specifically, under the lowest levels of both extraneous and intrinsic load individuals appear to perform their best but as these two forms of load increase this performance begins to deteriorate. In the next study the role of germane load (the abilities, skills, knowledge brought to a task by the individual) will be investigated in terms of task motivation and task exposure.
Chapter 6

STUDY THREE

EXTRANEOUS, INTRINSIC AND GERMANE COGNITIVE LOAD
6.1 RATIONALE

In Study Two it was shown that extrinsic cognitive load significantly impacted capacity limitations; under low cognitive load there appeared to be no competition for resources cross-modally compared to within-modally, however, under conditions of high cognitive load there was competition and the cross-modal advantage was eliminated. Moreover, this performance based cross-modal advantage was accompanied by participants reporting overall lower levels of difficulty in the cross-modal compared to within-modal conditions. This indicates that in addition to extraneous load being affected by the manipulated cognitive load conditions, the presentation format additionally affected intrinsic cognitive load. These outcomes are in clear support of the central hypotheses of this thesis regarding the relationship between dual-task attentional resourcing and cognitive load. Nevertheless in Study Two there did appear to be significant individual differences in objective performance and subjective self-report experience of cognitive load. It is these individual differences which are of interest in this study.

Study Two revealed that some participants responded to the presentation of the target much more quickly than other participants, and in turn some participants showed evidence of a more substantial cross-modal advantage than others. A review of the participant information and data indicated that some participants had participated previously in Study One (thus, potentially gaining experience within the task) and others appeared to be more cognisant of the response requirements of the dual-task than others and the overall purpose of the study (thus, potentially being more motivated to perform well). Additionally, in the subjectively rated questions some participants also appeared more confident than others in their responses – although without confidence ratings for all experimental trials it was not possible to determine any real correlation between confidence and performance.
The potential role of confidence, motivation and/or exposure in moderating cognitive performance is certainly not novel – there is extensive literature demonstrating that increasing any of these aspects is likely to positively impact performance on a given task (Roberts, Hann & Slaughter, 2006). Confidence, motivation and exposure (or otherwise known as ‘training’ or ‘experience’) are thought to aid in the learning processes involved when individuals gain mastery in a domain – as they begin to automatise the task (or sub-tasks) (Sweller, 1994). Importantly, increasing motivation and/or exposure are thought to be two of several ways to increase the knowledge and skills of the individual and in turn reduce the experience of germane cognitive load. As such, it is not only valuable to investigate the individual differences evident in Study One and Study Two but in doing so it is possible to examine the final aspect of the 3-part taxonomy of cognitive load proposed by Brünken et al. (2003).

The aim of this study is to investigate whether confidence, motivation and/or exposure may affect cognitive load and/or performance. It is proposed that investigation of these features may provide some added explanation of the individual differences demonstrated in Study Two. For instance, increases in motivation may lead to subsequent increases in cognitive load, at least initially, as the individual will presumably exert more effort in participation. Given time and exposure to the task this motivation would be expected to encourage individuals to develop strategies or ‘short cuts’ that enable them to perform well while using fewer resources, that is, as their skills become automatised through rehearsal (Paas, Renkl & Sweller, 2003). As such, it would be expected that over time participants will experience less perceived cognitive load on the same trials which were previously perceived to have higher levels of cognitive load. In this study both motivation and exposure will be directly manipulated, while confidence ratings will again be taken and then, along with the ratings of task complexity and task difficulty, used to predict performance accuracy.
To manipulate motivation and to determine whether it impacts cognitive load, or perceived cognitive load, in this study half of the participants will be provided with a monetary incentive to perform their best (the ‘motivated’ group), while the other half of participants will receive the same reward but will be told the incentive is not contingent on their performance (the ‘non-motivated’ group). To manipulate exposure, participants will be required to undertake the experiment on two occasions (pre- and post-test sessions). Half of the participants will then be exposed to an ‘exposure’ task on three additional occasions between pre- and post-test sessions (the ‘exposure’ group) while the other half of participants will not have this additional exposure (the ‘no-exposure’ group). As such, participants will complete the same basic experiment as in Study Two, although this time participants will additionally be allocated to one of four experimental groups (motivated/exposure, motivated/no-exposure, non-motivated/exposure, non-motivated/no-exposure).

The following questions will be addressed in this experiment:

(1) Is there a modality combination x cognitive load effect, as seen in Study Two be replicated in this study?

Hypothesis 1: The degree of cross-modal advantage will be inversely proportional to cognitive load.

(2) How do motivation and exposure affect performance between pre- and post-test sessions, do they interact with the modality combination x cognitive load relationship (as in (1) above), and do they impact perceived cognitive load?

Hypothesis 2: Perceived cognitive load will decrease over time (from pre- to post-test) as a positive function of motivation and/or exposure.

Hypothesis 3: Any cross-modal advantage will increase over time (pre- to post-test) as a positive function of motivation and/or exposure.
(3) How do complexity, difficulty and confidence relate to one another and what are their roles in predicting performance?

Hypothesis 4: Complexity, difficulty and confidence are independent constructs.

Hypothesis 5: Complexity, difficulty and confidence will predict performance.

6.2 EXPERIMENTAL DESIGN

This experiment has a $2 \times 2 \times (3 \times 3 \times 2)$ factorial design. The same two within-subjects independent variables used in Studies One and Two, are again examined; modality combination consisting of three levels (i) auditory/auditory (AA) (ii) visual/visual (VV) and (iii) auditory/visual (AV) together with cognitive load (i) low cognitive load (LCL) (ii) medium cognitive load (MCL) and (iii) high cognitive load (HCL). An additional independent variable – Test Version consisting of two levels (i) pre-test and (ii) post-test was included to evaluate the effects of test exposure. The within-subjects (extraneous) cognitive load manipulation remains the same as that used in Study Two with the number of items to be processed across the standard trial duration altering cognitive load (in accordance with Woods, 1988) – three, six and eleven items respectively. Participants will therefore complete trials falling in all nine of these within-subjects variable conditions.

The first of the between-subjects independent variable is motivation consisting of two levels (i) motivated – in which participants are instructed that an additional monetary prize will be awarded if they perform in the top 25% of participants, and (ii) non-motivated – in which participants are told they will receive an extra payment at the conclusion of the experiment irrespective of their performance. The second of the between-subjects independent variables is task exposure again consisting of two levels (i) no-exposure – in
which participants complete only pre- and post-test session and (ii) exposure – where participants complete an additional three sessions (the same as the experimental sessions) between pre- and post-test. Therefore, participants were allocated to one of four possible between-subjects group: motivated/exposure, motivated/no-exposure, non-motivated/exposure, non-motivated/no-exposure. The need to evaluate the effect of exposure required a 3-day longitudinal design requiring participants to take part in a pre-test session on day one followed by a post-test session on day three and for the exposure group a series of exposure sessions on day two.

To allow comparisons to be drawn between studies the same stimuli (from Study Two) were used in this experiment, and the same dependent variables were again assessed with the omission of RT (see Chapter 4, Discussion). Therefore, in this study cross- and within-modal performance is assessed via target identification accuracy – an objective and direct measures of cognitive load (Brünken et al., 2006).

Self-report measures of trial difficulty, trial complexity and confidence will again be assessed in a 3-item rating questionnaire at the conclusion of the pre- and post-test sessions. These items are of most interest in evaluating the relationship between cognitive complexity (as measured by task complexity ratings) and perceived cognitive load (as measured by task difficulty ratings) and in turn how they, together with confidence, may predict accuracy performance.

6.3 Method

6.3.1 Participants

There were 63 participants in the pre-test sessions, 27 males (mean age 25.0 years, $SD=10.1$) and 35 females (mean age 23.8 years, $SD=6.7$). Presumably due to the semi-longitudinal design numbers dropped to 49: 26 males (age: $M=25.1$ years, $SD=10.1$) and 23
females (age: $M=25.2$ years, $SD=5.2$) in post-test (attrition rate = 8.6%). All participants reported normal hearing and normal (or corrected-to-normal) visual acuity. Participants provided informed consent and were made aware they could withdraw this consent at any time without penalty and would receive any incentives on offer at a pro-rata rate. Participants received course credit for their participation in the pre-test session, and were paid $30 to return for the post-test session (no payment was provided during the exposure sessions). Participants were also awarded a further $25, by way of a Caltex Fuel Card\(^1\), following completion of each the pre- and post test sessions – with those in the motivated groups believing the award of these fuel cards was contingent on performance. All participants were debriefed at the conclusion of the post-test session that the Fuel Card reward would be given irrespective of their performance.

6.3.2 Materials and Apparatus

As in the previous studies, the experiment was conducted on a LG 17” widescreen laptop with attached loud speakers (Genelec Analogue 50/60Hz), with the stimuli presentation and participant responses run and recorded in DMDX software (Forster & Forster, 2003). Auditory stimuli were presented to participants via loud speakers, which were vertically offset (by 30 mm) from the presentation of the visual stimuli at the top of the computer screen.

Again, participants used the keyboard to progress through the experiment; [spacebar] to key through onscreen instructions and the left (colour coded blue) and right (colour coded red) shift keys to indicate identification of the target letter presentations during each trial.

\(^1\) A Caltex Fuel Card is issued by Caltex Petrol (Gas) Stations throughout Australia which can be pre-loaded with a specific dollar value. This card can then be used to redeem fuel or other merchandise available through Caltex Petrol Stations (excluding cigarette products).
A 3-item rating questionnaire was included in the final stages of both the pre- and post-test sessions in the same fashion as it was conducted in the previous experiments, yielding ratings of: (i) complexity (ii) difficulty and (iii) confidence. (see Appendix B).

6.3.3 Stimuli

The auditory and visual letter stimuli and letter sequence strings used in Study Two were again used in this study; the target letter remained “B” and the distractor letters (Auditory: V, E, D, C, G, U, X, Z, I and L; Visual: F, E, D, P, R, Y, J, T, N and M) were also reused. Moreover, the auditory and visual letter sequence pairings used to manipulate extrinsic cognitive load in Study Two were again used in this study (see Figure 5.1). The AA, VV and AV trials were constructed in the same manner as in Study Two.

Single-task trials were included to provide a baseline measure of participant performance. In comparison to Study Two, the number of single-task trials was increased (from 24 to 48) so that valid trial response windows could be individually determined. Rather than imposing a pre-established response time frame on all participants in this experiment individual response time windows were calculated for each participant based on their target identification performance on single-task trials. This method was introduced here given that in Study Two it appeared that some participants were consistently faster than others and a single response window meant that at times it could be difficult to determine which of two letters the participants was responding too. Using this individualised method, it is anticipated that the response time windows can be made considerably smaller and thus more accurate.

6.3.4 Procedure

As in the previous studies participants were seated in front of the laptop computer within a small lit laboratory testing room, so that the computer screen and accompanying
auditory loud speakers were at approximately chest height. The same experimental setup was used for each the pre-, post and exposure-test sessions.

Upon arrival at the pre-test session participants were allocated to one of the four between-subjects experimental groups. Those allocated to the exposure groups were asked if they would be willing to return for an additional two days of testing, and those in the no-exposure group were asked to return two days later for the post-test session. If participants were not willing to return they were still required to complete the pre-test session to receive the course credit on offer. Participants in the motivated experimental group were then told that their performance would be monitored and if they were able to perform in the top 25% of participants they would be awarded a $25 fuel card for each the pre- and post-test sessions. Participants in the non-motivated group were instructed they would receive the fuel card at the conclusion of the post-test sessions, and further that they were not to discuss this reward with anyone until the study was concluded. Withholding the fuel cards until after the post-test session was necessary to maintain the illusion of performance calculations being conducted for the motivated group. Participants who did not return for the post-test sessions were still awarded the fuel card for participating in the pre-test session.

At the commencement of the pre-test sessions participants completed 18 practice trials with two trials from each of the nine within-subjects conditions. These practice trials were used so that participants could become familiar with the auditory and visual stimuli, the three time sequences, the progression of the experiment and the responses they would be required to make. Following completion of these 18 practice trials, the researcher reviewed participants’ responses and provided some feedback regarding their performance – letting them know whether they were identifying the targets within a reasonable RT window (205 ms-425 ms the window used in Study Two as a guide). If a participant’s responses did not appear reasonable to the researcher (i.e., they were responding to both targets, not just one,
and were responding immediately following the target presentation and not at the conclusion of the trial) they again explained the task and they participant re-took the practice trials until a reasonable performance level was achieved.

The first phase of the pre-test session saw participants complete the 324 dual-task trials originally tested in Study Two (36 trials for each experimental condition) together with 48 single-task trials. These trials were randomly presented within 12 experimental blocks, each consisting of 31 trials. Each trial block was initiated by pressing the [spacebar], which was followed by a period of silence, a series of hashes (###) together with white noise and then a second period of silence before the letter sequences commenced 400 ms after the [spacebar] strike. Within each block, the trials began automatically 500 ms after the conclusion of the previous trial.

The second phase on the pre-test session required participants to complete 36 ‘rating trials’ – the same as those completed in the practice trials repeated twice. Each trial was initiated with the participant pressing the [spacebar] the trial being presented and responded to as in the first phase. Following each trial participants completed a 3-question reflection questionnaire.

Participants in the exposure group were required to return the follow day to complete the 372 experimental trials on three separate repetitions. The rating items were not completed during these exposure sessions.

On the third day participants from both the exposure and no-exposure groups returned to take part in the post-test session. This session was run in the same fashion as on day one, including the 18 practice trials (to re-engage and re-familiarise participants), the 372 experimental trials and the 36 rating items.

The pre- and post test sessions took approximately 1 hour and 20mins to complete, and each of the exposure sessions took approximately 50mins each to complete.
6.4 RESULTS

6.4.1 Data Reduction

Data reduction was undertaken in the same manner as it was in Studies One and Two for accuracy, and the complexity, difficulty and confidence ratings.

Like the coding in the first two studies each participant’s experimental trials was coded to reflect the valid responses – either two valid key presses (2ID), one valid key press (1ID) or no valid key presses (0ID). In this study a valid response occurred if the appropriate [shift] key was pressed within an individualised RT window following presentation of the target letter. Individualised RT windows were established by evaluating all the single-task trials in which a single [shift] press was recorded for each participant and establishing a mean RT for these trials. The valid RT window was then established to be $\pm 2SD$ around this mean (see Appendix I for a table of all participant valid RT windows in this experiment). The RT window established in the pre-test session was retained for the post-test session.

The data for a participant was entirely excluded from analysis if they did not complete both the pre- and post-test sessions, or if their error rate was over 20% on the single-trial conditions.

Participant rating items were coded to allow for quantitative analysis. The data reduction for these three items was conducted in the same manner as in Studies One and Two.

6.4.2 Data Screening

The data for any participant who did not complete both pre- and post-test sessions was excluded from analysis. On this basis the pre-test data for 12 participants were removed from further analysis. Frequency analyses indicated there was no difference in accuracy performance in each of the nine experimental conditions between those participants who did
or did not return for the post-test sessions. This indicates that no identifiable biases may impact the post-test outcomes as a result of the changing participant pool. Surprisingly the drop-out rate was evenly spread across the four between-subjects groups (it was anticipated that the exposure group drop out rate would be higher than the no-exposure group given the need to return without incentive on day two), and as such no additional recruitment was necessary to balance the conditions prior to inferential analyses being conducted.

Data screening using SPSS revealed one univariate outlier. According to the recommendations of Tabachnick and Fidel (2007) the outlying scores were adjusted to be one unit more extreme than the next closest. No multivariate outliers were identified and as such data for the remaining 49 participants was retained. All univariate and multivariate assumptions of normality were met for all performance and self-rated dependent variables to be examined in the following analyses of variance, factor analysis and regression analysis.

6.4.3 Data Analyses Overview

In this study the planned contrasts in the previous two studies for modality combination and cognitive load were again conducted. Additional planned contrasts (simple comparisons and their interactions) were occasioned by the introduction of the three new two-level independent variables (these are set out in Appendix J).

For the objective dependent variable, accuracy, firstly, a (3 x 3) repeated measures factorial ANOVA, with alpha set at .05, was conducted on just the pre-test trials to determine whether the results demonstrated in Study Two could be duplicated here before further analyses were conducted. secondly, a 2x2x(2) mixed factorial ANOVA was conducted to evaluate whether the germane load manipulation of motivation and exposure enhance performance from pre- to post-test sessions. Finally, a 2x2x(3x3x2) mixed factorial ANOVA is used to examine the effects of motivation and exposure between pre- and post-test sessions
for each of the three modality combination conditions (AA, VV and AV) across the three load levels (LCL, MCL and HCL).

As this is the first study to incorporate all three sub-types of cognitive load (intrinsic, extraneous and germane) a factor analysis, correlations and a number of multiple-regression analyses are performed to evaluate correlations between objectively manipulated cognitive load and the perceptions of complexity, difficulty and confidence, whether these latter three are individual constructs, and how there later three, the subjective measures, predict objective performance accuracy.

6.4.4 Analyses

6.4.4.1 Performance: Modality Combination and Cognitive Load

Prior to evaluating the effects of germane cognitive load (via motivation and exposure manipulations) it is important to determine whether the trends found in Study Two and their interactions with cross- and within-modality conditions for cognitive load could be duplicated. To this end, a repeated measures ANOVA was conducted on just the pre-test data for modality combination and cognitive load collapsed over the motivation and exposure variables. The results are presented in Figure 6.1.

The cross-modal versus within-modality contrast revealed an overall cross-modal advantage \( F_{AV \ vs. \ AA+VV}(1,24)=179.03, \ p<.001, \ \text{partial } \eta^2=.79 \), while no difference in performance accuracy was evident between dual-auditory and dual-visual performance \( F_{AA \ vs. \ VV}(1,24)=2.00, \ p=.16, \ \text{partial } \eta^2=.04 \). There was also a significant overall linear and quadratic effect for cognitive load \( F_{Lin}(1,24)=994.90, \ p<.001, \ \text{partial } \eta^2=.95; \ F_{Quad}(1,24)=58.64, \ p<.001, \ \text{partial } \eta^2=.55 \) indicating decrements in performance as a function of cognitive load accelerate as cognitive load increases. These overall linear and quadratic effects of load also interact with the cross- versus within-modal contrast \( F_{AV \ vs. \ AA+VV}(1,24)=179.03, \ p<.001, \ \text{partial } \eta^2=.79 \).
$AA+VV*Lin(1,24)=71.19, p<.001$, partial $\eta^2=.60$; $F_{AV vs. AA+VV*Quad(1,24)}=15.29, p<.001$, partial $\eta^2=.24$). As can be seen in Figure 6.1 while within-modal performance tends to decrease at a consistent relatively shallow rate across increasing levels of cognitive load, cross-modal performance initially decreases at the same rate (from LCL to MCL) before a sharper decrease (between MCL and HCL) such that accuracy diminishes to a point similar to both AA and VV performance. These patterns of the differences between cross- and within-modal accuracy decreases across increasing levels of cognitive load are consistent with the outcomes of Study Two.

![Figure 6.1](image_url)

*Figure 6.1.* Mean participant accuracy percent scores across (+SE) all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).
6.4.4.2 Effects of Germane Cognitive Load, Motivation and Exposure, on Performance

The planned comparisons are depicted in Figure 6.2. These revealed that overall performance increased from pre- to post-test sessions, suggesting a general improvement in performance independent of manipulations of motivation and exposure ($F_{\text{Test}}(1,45)=93.65$, $p<.001$, partial partial $\eta^2=.68$). In fact, neither motivation alone, nor exposure alone, nor a combination of the two were directly associated with performance collapsed over pre- and post-test trials ($F_{\text{Mot}}(1,45)=.59$, $p=.45$, partial $\eta^2=.01$; $F_{\text{Exp}}(1,45)=1.11$, $p=.30$, partial $\eta^2=.03$; $F_{\text{Mot*Exp}}(1,45)=.18$, $p=.67$, partial $\eta^2=.01$). Nevertheless, there were significant interactions of both motivation and exposure with pre- versus post-test.

![Figure 6.2](image_url)

*Figure 6.2. Mean participant accuracy percent scores (+SE) for pre- to post-test sessions for the two motivation and two exposure conditions.*
First, individuals who were motivated demonstrated significantly greater improvements from pre- to post-test than did those in the non-motivated groups \((F_{Test*Mot}(1,45)=21.38, p<.001, \text{partial } \eta^2=.32)\). Second, those who received exposure sessions between pre- and post-tests demonstrated significantly greater improvements in accuracy compared to those who had no additional sessions \((F_{Test*Exp}(1,45)=16.84, p<.001, \text{partial } \eta^2=.27)\). Despite these facilitatory effects of both motivation and exposure, and the indication in Figure 6.2 that the motivated/exposure group had greatest improvements from pre- to post-test sessions, the interaction of motivation and exposure from pre- to post-test was not significant \((F_{Test*Mot*Exp}(1,45)=.18, p=.67, \text{partial } \eta^2=.01)\). This suggests that the facilitatory effects of both motivation and exposure are independent of each other; the two manipulations of germane cognitive load do facilitate performance and do so independently.

### 6.4.4.3 Interactions of Extraneous and Germane Cognitive

The results of the factorial ANOVA are depicted in Figure 6.3. The results here for the pre-test, shown in section 6.4.4.1 show similar cross- and within-modal effects across the levels of cognitive load to those in Study Two (section 5.4.4.1). Specifically, cross-modal performance is only minimally affected between LCL and MCL with a more rapid deterioration in performance between MCL and HCL conditions, while in both within-modal conditions the performance decrement is shallower and consistent.

In post-test conditions the same pattern as in pre-test conditions for each motivation by exposure condition is evident with the exception of the motivated/exposure group. Here cross-modal performance which had at pre-test seen a shallow followed by a steeper decrease in accuracy over cognitive load levels\(^2\), now at post-test show a shallow decrease across all

\(^2\) \(F_{AV vs. AA+VV*TextV*Lin}(1,12)=13.75, p=.003, \text{partial } \eta^2=.53; F_{AV vs. AA+VV*TestV*Quad}(1,12)=4.96, p=.046, \text{partial } \eta^2=.29\)
Figure 6.3. Changes to modality combination by cognitive load interaction from pre- to post-test as a function of motivation and exposure.
levels of increasing load\textsuperscript{3}. The sharp drop in AV performance between MCL and HCL conditions is no longer present at post-test leading to cross-modal performance being significantly superior to within-modal performance across all levels of cognitive load, with similarly-sloped parallel functions for cross- and within-modal performance.

These outcomes seem to suggest that with both motivation to improve and exposure (or training) to the task cross-modal performance maintains its superiority to within-modal performance under conditions of high cognitive load. It would seem that the presence of these two features somehow immunises the individual against performance deterioration.

6.4.4.4 Predicting Performance from Complexity, Difficulty and Confidence

A mixed measures ANOVA was conducted to assess any changes from pre- to post-test perceptions for complexity, difficulty and confidence as a function of motivation and exposure.

*Task Complexity.* As can be seen in Figure 6.4 there was no overall significant effect from pre- to post-test sessions indicating that ratings of task complexity were constant over sessions \((F_{\text{Test}}(1,45)=.17, p=.68, \text{partial } \eta^2=.01)\). Moreover, no overall effects of motivation or exposure were evident, nor any interactions between these indicating that perceptions of complexity were not affected by the presence of motivation and/or exposure factors:

- \(F_{\text{Mot}}(1,45)=.48, p=.49, \text{partial } \eta^2=.01\);
- \(F_{\text{Exp}}(1,45)=1.02, p=.32, \text{partial } \eta^2=.02\);
- \(F_{\text{Mot} \times \text{Exp}}(1,45)=.17, p=.68, \text{partial } \eta^2=.01\).

\footnote{\(F_{\text{AV vs. AA+VV\*TestV\*Lin}}(1,12)=.56, p=.47, \text{partial } \eta^2=.04\); \(F_{\text{AV vs. AA+VV\*TestV\*Quad}}(1,12)=.16, p=.70, \text{partial } \eta^2=.01\)}
Figure 6.4. Mean participant complexity ratings (+SE) for pre- to post-test sessions for the two motivation and exposure conditions.

Task Difficulty. As can be seen in Figure 6.5 there is a general decrease in perception of difficulty between pre- and post-test conditions independent of any other manipulations of motivation and exposure ($F_{\text{Test}}(1,45)=90.65$, $p<.001$, partial $\eta^2=.65$). While there were no effects of motivation and exposure of their interaction collapsed over pre- and post test\(^4\) it does appear that these factors differentially affected decrements in ratings from pre- to post-test. Indeed, individuals who were motivated demonstrated significantly greater decreases in perception of difficulty from pre- to post-test ($F_{\text{Test*Mot}}(1,45)=54.78$, $p<.001$, partial $\eta^2=.42$), and those who received exposure sessions between pre- and post-tests also demonstrated

\(^4\) $F_{\text{Mot}}(1,45)=.49$, $p=.48$, partial $\eta^2=.01$; $F_{\text{Exp}}(1,45)=1.88$, $p=.18$, partial $\eta^2=.03$; $F_{\text{Mot*Exp}}(1,45)=.15$, $p=.70$, partial $\eta^2=.01$
significant decreases in their ratings of difficulty compared to those who had no additional sessions ($F_{Test*Exp}(1,45)=32.45$, $p<.001$, partial $\eta^2=.35$). Despite the indication in Figure 6.5 that the group of participants in the motivated/exposure group had the greatest decrease from pre- to post-test sessions the interaction of motivation and exposure from pre- to post-test was not significant ($F_{Test*Mot*Exp}(1,45)=3.51$, $p=.07$, partial $\eta^2=.11$).

![Figure 6.5. Mean participant difficulty ratings (+SE) for pre- to post-test sessions for the two motivation and exposure conditions.](image)

**Task Confidence.** As can be seen in Figure 6.6 there is an overall improvement in ratings of confidence from pre- to post-test sessions ($F_{Test}(1,45)=12.75$, $p=.001$, partial $\eta^2=.22$). There were no effects of motivation or exposure collapsed over time ($F_{Mot}(1,45)=.87$, $p=.36$, partial $\eta^2=.01$; $F_{Exp}(1,45)=2.21$, $p=.14$, partial $\eta^2=.01$), however, an interaction
between test version and exposure suggests that ratings of confidence increase more following extended exposure to a task \( F_{Mot\times Exp}(1,45)=31.55, p<.001, \text{ partial } \eta^2=.21 \). Overall, this suggests that confidence increases as task exposure increases, while motivation appears to have little impact on the perception of confidence.

![Figure 6.6](image.png)

**Figure 6.6.** Mean participant confidence ratings (+SE) for pre- to post-test sessions for the two motivation and exposure conditions.

### 6.4.4.5 Comparison of Subjective Measures and Objective Accuracy

Comparing the outcomes of these subjective rating items with the objective accuracy performance results (see 6.4.4.2) some interesting findings are apparent. Generally it appears that as accuracy performance decreased quite strong decreases in perceptions of difficulty were apparent which were generally accompanied by an overall increase in perceived confidence. Moreover, while increases in accuracy and matched decreases in perceived
difficulty were a result of motivation and exposure only the presence of exposure saw somewhat of a match between increases in both accuracy and confidence. No changes were evident for task complexity. Given the appearance of these relationships it was considered appropriate to conduct a factor analysis and a series of multiple regressions to determine the independence of these constructs and in turn the power of these measures to predict accuracy performance.

A principal components factor analysis (with varimax orthogonal rotation) was performed to examine whether confidence ratings are indeed a measure independent of the two subjective load items, and whether task complexity and task difficulty are also two independent constructs. Three components (factors) with eigenvalues greater than one were extracted, which together accounted for 57% of the variance. The factor loadings, communalities \((h^2)\) (the amount of variance of the variable accounted for by the factors), and percentage of variance explained after varimax rotation are shown in Table 6.1. Component loadings less than .30 have been suppressed to aid interpretation (as per Hills, 2005).

As can be seen in Table 6.1, each of the three subjective ratings are predominantly represented by a single separate factor. That said, there is some commonality between complexity and difficulty over Factors 1 and 2 but Factor 1 is predominantly a ‘complexity’ factor and Factor 2 a ‘difficulty’ factor. Factor 3 is clearly a separate ‘confidence’ factor.

Multiple standard regressions were then conducted to evaluate how each of the three self-rated variables (complexity, difficulty and confidence)\(^5\) contributed to accuracy variance of pre-test accuracy percent scores of participants. Pre-test scores were evaluated to avoid confounds that may have been introduced by collapsing pre- and post-test scores and the manipulations of motivation and exposure. Three separate regressions were conducted for each modality combination (AA, VV and AV) to determine whether different combinations

\(^5\) As the factor analysis revealed three relatively distinct constructs, original ratings measures were used in these regressions rather than calculating new comparative measures based on the produced factors.
of factors may differentially affect performance. The results of this analysis should be
interpreted with caution as the sample size could be considered small ($N=49$) (Hills, 2005).
However, according to Tabachnick and Fiddel (1989) such a sample does meet the absolute
minimum size required for a multiple regression analysis, that is, with three variables of
interest a minimum of 15 participants would be required to maintain power and reliability.

Table 6.1

Varimax Rotated Component Loadings for the Three Subjectively Rated Questions in the
Reflections Questionnaire

<table>
<thead>
<tr>
<th>Item</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Rating of Complexity</td>
<td>.58</td>
</tr>
<tr>
<td>Rating of Difficulty</td>
<td>.30</td>
</tr>
<tr>
<td>Rating of Confidence</td>
<td>…</td>
</tr>
</tbody>
</table>

% of variance 20.00 34.09 21.89 75.98

Label

Cognitive Load Confidence

Table 6.2 shows the descriptive statistics and variable intercorrelations for each of the
modality combinations. As could be expected from the factor analysis previously performed
(see section 6.4.6) across all modality combinations ratings of trial complexity and trial
difficulty were highly correlated (AA: $r= .536$; VV: $r= .437$; AV: $r= .460$).
Table 6.2

Means and Intercorrelations between Accuracy Percent Scores, Ratings of Complexity, Difficulty and Confidence for AA, VV and AV trials

<table>
<thead>
<tr>
<th>Variables</th>
<th>Accuracy % (Criterion)</th>
<th>Trial Complexity</th>
<th>Trial Difficulty</th>
<th>Trial Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AA Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>-.196</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-.717*</td>
<td>.536*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>.005</td>
<td>-.002</td>
<td>.107</td>
<td></td>
</tr>
<tr>
<td>M (SD)</td>
<td>79.71 (2.44)</td>
<td>4.61 (.93)</td>
<td>4.59 (.49)</td>
<td>5.10 (.79)</td>
</tr>
<tr>
<td><strong>VV Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>-.196</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-.450*</td>
<td>.437*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>-.807</td>
<td>.213</td>
<td>.356</td>
<td></td>
</tr>
<tr>
<td>M (SD)</td>
<td>79.71 (2.44)</td>
<td>4.61 (.91)</td>
<td>4.53 (.51)</td>
<td>1.63 (.52)</td>
</tr>
<tr>
<td><strong>AV Trials</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>-.053</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-.749*</td>
<td>-.460**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>.108</td>
<td>.239</td>
<td>-.139</td>
<td></td>
</tr>
<tr>
<td>M (SD)</td>
<td>81.55 (2.24)</td>
<td>4.63 (.85)</td>
<td>4.29 (.46)</td>
<td>5.10 (.78)</td>
</tr>
</tbody>
</table>

* \( p \leq .001 \), ** \( p \leq .005 \)

Table 6.3 displays the unstandardised regression coefficients (\( B \)) and standardised regression coefficients (\( \beta \)) for the predictors of interest within the modality combination groups.
Table 6.3

*Standard Multiple Regression of Trial Complexity, Trial Difficulty and Trial Confidence as Predictors of Accuracy Performance*

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>B SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA Trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>-.069</td>
<td>.278</td>
<td>-.026</td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-3.537</td>
<td>.525</td>
<td>-.720*</td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>.251</td>
<td>.318</td>
<td>.082</td>
</tr>
<tr>
<td>Constant</td>
<td>94.993</td>
<td>2.815</td>
<td></td>
</tr>
<tr>
<td>VV Trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>.015</td>
<td>.159</td>
<td>.006</td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-7740</td>
<td>.308</td>
<td>-.746**</td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>-3.960</td>
<td>.298</td>
<td>-.137*</td>
</tr>
<tr>
<td>Constant</td>
<td>89.302</td>
<td>1.422</td>
<td></td>
</tr>
<tr>
<td>AV Trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial Complexity</td>
<td>.043</td>
<td>.126</td>
<td>.017</td>
</tr>
<tr>
<td>Trial Difficulty</td>
<td>-4.687</td>
<td>.232</td>
<td>-.853*</td>
</tr>
<tr>
<td>Trial Confidence</td>
<td>-.081</td>
<td>.137</td>
<td>-.029</td>
</tr>
<tr>
<td>Constant</td>
<td>101.854</td>
<td>1.352</td>
<td></td>
</tr>
</tbody>
</table>

* p<.001. ** p=.027

Before discussing the results of these three regression analyses in details it can be noted that while perceived difficulty is a predictor of performance in all three modality combinations, complexity did not appear to add any explanatory power above and beyond the
contribution of perceived difficulty. Further, difficulty alone predicts for AA and AV performance but in VV difficulty and confidence are predictors.

Two 3-step stepwise multiple regressions (1. complexity → difficulty → confidence; 2. difficulty → complexity → confidence) were then performed (on collapsed modality combination data given the similarities across modalities) to determine whether, given the correlations between complexity and difficulty, complexity may well be a predictor of load—a finding possibly hidden in the standard regression. Indeed, the first stepwise regression indicated cognitive complexity to be a predictor of performance at Step 1 with the addition of difficulty at Step 2 reliably improving the predictive power of the equation, R=.94, $R^2=.88$, $F(2,46) = 78.45$, $p < .001$. However, the second stepwise regression results indicated that difficulty alone predicted performance at Step 1 ($R=.994$, $R^2=.89$, $F(1,47) = 69.09$, $p < .001$) and that the addition of complexity did not reliably improve $R^2$.

Returning to the original standardised multiple regression results, where difficulty was alone a predictor in AA and AV, but both difficulty and confidence were predictors in VV. Descriptive and frequency analysis of accuracy as a function of confidence for VV compared to AA and AV indicate that while participants were more confident of their performance in the AV ($M=5.78$, $SD=.15$) compared to the AA ($M=4.12$, $SD=1.11$) and VV ($M=4.25$, $SD=.74$) conditions, this more confident self-assessment was often in error presumably why confidence, although most superior in AV was not a predictor of AV performance, For instance, as shown in Table 6.4, in the VV trials participants self-assessed their performance very accurately (e.g., ratings of “identified both targets” were accompanied by errors in actually identifying them in only 6.9% of cases, with 1.5% of ratings being “unsure”). In contrast in AV trials participants tended to overrate their performance (e.g., “identified both targets” was accompanied by errors in actual identification in 12.4% of cases and 1.2% of “unsure” ratings) and in AA trials participants tended to underrate their
performance (e.g., ratings of “unsure” were in 8.1% of cases where in fact participants identified both targets in over 60% of these “unsure” trials).

Table 6.4

*Participant Ratings of Target Identification Confidence and Constitution of Actual Performance on These Trials*

<table>
<thead>
<tr>
<th>Modality Combination</th>
<th>AA</th>
<th>VV</th>
<th>AV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error % of trials rated as “I identified both targets detected”</td>
<td>0%</td>
<td>6.9%</td>
<td>12.4%</td>
</tr>
<tr>
<td>% of trials rated as “unsure of my performance”</td>
<td>8.1%</td>
<td>1.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>% of “unsure” trials where both targets actually identified</td>
<td>60%</td>
<td>9%</td>
<td>12%</td>
</tr>
</tbody>
</table>

6.5 **Discussion**

It was hypothesised that germane cognitive load would impact the way in which attentional resources are allocated during multi-task processing. Specifically, it was predicted that both motivation and exposure to the task would lead to decreases in germane cognitive load and result in a sustained cross-modal advantage. That is, for those individuals in the +motivation and +exposure groups post-test ratings of difficulty should be reduced, with post-test accuracy higher. In addition a cross-modal advantage was hypothesised as for the earlier studies. These hypotheses were all supported in this study, with the modality combination (AA+VV vs. AV) by cognitive load cross-modal advantages similar to those in Study Two (greatest advantage for low, moderate advantage for medium, and no advantage for high cognitive load). Moreover, individuals who were in the +motivated and/or +exposure groups
reported, between pre- and post-test sessions, reduced ratings of difficulty and greater accuracy under high cognitive load.

Therefore, it can be seen that both extraneous and intrinsic cognitive load (Brünken et al., 2003) do impact how attentional resources are distributed during auditory and visual processing. Additionally, the interactions between modality combination, cognitive load manipulation, motivation and exposure suggest that germane cognitive load (Brünken et al.,) moderates the overall experience of load – in this case reducing overall load and bringing about an improvement in target identification accuracy, most obviously during AV processing.

In terms of complexity and difficulty ratings participants’ perceptions of complexity remained consistent from pre- to post-test sessions while ratings of difficulty appeared to change, and most obviously for those individuals in the ‘motivation and/or ‘exposure groups. That complexity ratings remained constant is promising given that the actual trials did not alter between test conditions. Therefore, this complexity rating is highly likely to be a measure of exactly the dimension it was intended to measure – the features of the task which were manipulated to affect perceptions of load. In contrast, the reductions seen in the ratings of difficulty suggest that with motivation and opportunity to improve individuals did manage to reduce the overall load experience to facilitate performance improvements. These slightly disparate results of trends between ratings of complexity and difficulty were further evaluated in the factor analyses where it was revealed that the two are independent constructs which, although are often correlated (factor loadings greater than 30% on each others primary factor), do not completely overlap. This independence was evident in the subsequent multiple regression analyses which indicated that complexity predicted performance in the forced 3-step (complexity → difficulty → confidence) stepwise regression but that difficulty ratings
were a far stronger predictor of performance (the enter-all regression revealed only difficulty to be a predictor).

Over and above the hypotheses addressed in this study an unexpected finding concerning confidence ratings was revealed. Confidence was a significant predictor specifically of VV performance. This seems to suggest that participants are most appropriately confident when they are not engaging in any auditory processing; accuracy was highest in the VV condition, lowest in the AA condition and moderate in the AV condition. That individuals’ most accurately reflected on visual, as opposed to auditory, processing in this study would seem consistent with findings suggesting humans tend to use their visual modality for 90% of all stimulus processing (Sivak, 1996). With this level of day-to-day exposure it would seem logical that self-assessment of this processing would be more accurate than auditory self-assessment, which is less regularly engaged and practiced. It would therefore be of interest in future studies to determine whether extensive training in auditory tasks could improve confidence and in turn performance in AV and AA conditions. Improving auditory processing self-assessment is crucial for in many workplace environments performance on the job is measured not only in terms of reaction time and decision accuracy (the dependent variables of interest in this series of studies) but additionally in the ability to monitor a situation and to intercede only when necessary (Wickens & Hollands, 2000; Endsley & Kiris, 1995; Moray, 1986).

In this study the role of an attentional system which incorporates features of both the multiple- and unitary-resource models was again supported. However, in past studies it was uncertain whether fluctuating support for the pure multiple- and unitary-resource models resulted from changes in task features alone or whether such fluctuations were the result of the cognitive load experienced by the individual. To put this in the form of a question, is the direct influence of cognitive complexity responsible for a change in resource allocation – for
instance, do the task requirements, the number of stimuli requiring simultaneous processing, cause a shift to a unitary-resource given the sheer volume of processing or, does the shift from multiple- to unitary-resources result from the indirect experience of cognitive load – for instance, do the task requirements impose a burden on the individual and, dependent upon the magnitude of this burden and how it is moderated by the skills of the individual, determine how the task will be resourced. The results of this experiment demonstrate that germane cognitive load (Brünken et al., 2003) can indeed influence performance (↓ germane load → ↑ accuracy) and as germane load arises in part from the knowledge and abilities of the individual operator and can therefore moderate only the experience of the task and not the features of the task itself, this study shows that it is cognitive load which moderates how resources are allocated.

With three studies now reliably demonstrating that cognitive load affects how attentional resources are distributed the model-continuum proposed in Chapter 2 is supported with a slight qualification. In the continuum, both task complexity and perception of difficulty were used to explain fluctuations of support for the two pure resource models. With the findings presented here it would now be more correct to propose that task complexity influences the perception of difficulty (which can be moderated by the skills/knowledge of the individual) and that this overall experience of difficulty is what activates movements along the unitary-/multiple-resource model continuum. In any case the model-continuum notion provides a heuristic for explaining why in this series of studies support could be found for both models (manipulating load leads to a change in resource allocation), and additionally provides insight as to why in past studies support has been found for both the multiple- and unitary-resource models.

Now that the role of cognitive load in moderating resource allocation has been resolved, at least within the confines of the series of studies here, it is of interest to determine
exactly how this shift between the two pure models occurs. That is, cognitive load is the reason *why* the shift occurs, but it is also of interest exactly *how* this shift occurs. Does it depend upon a two-tiered model consisting of discrete multiple- and unitary-resources whereby a categorical switch between the two is experienced at some point along the ‘low cognitive load/cross-modal advantage’ versus the ‘high cognitive load/no cross-modal advantage’, if so, then this is a *coordinated hierarchical resource model*. Alternatively, given that there appears to be a diminishing cross-modal advantage (rather than a presence/absence dichotomy) it may be that the two systems function together and that gradually a unitary-system contributes more and more resources to processing and that the relatively smaller contribution of the modality-specific resources leads to the cross-modal advantage becoming, at a behavioural level, negligible. If so then an *interactive hierarchical resource model* is more appropriate.

Such hybrid model hypotheses represent a significant step forward in this field, for while the potential for such hierarchical models have been indicated in the literature (Alais et al., 2006; Spence & Driver, 1996; Sinnett et al., 2006; Posner, 1980) no such elaborations have been made about how a model might be structured. Examining the potential for such a structure is addressed in the fourth and final study here by introducing additional levels of cognitive load allowing a more intricate picture of the model-continuum shift to be developed. Moreover, single-, dual- and tri-task trials will be introduced in the next study to provide a means to differentiate between the coordinated and interaction hierarchical models.
Chapter 7

STUDY FOUR

A HIERARCHICAL MODEL OF ATTENTIONAL RESOURCING
7.1 RATIONALE

The outcomes of the first three studies suggest that human information processing is supported by attentional resources which conform to neither the structure of the multiple- nor unitary-resource models in isolation. Instead the results appear to indicate that attentional resources may be supplied via different reservoirs (either modality specific or general purpose) dependent upon the requirements of the task being completed.

This study introduces two modifications. Firstly an additional two levels of cognitive load, resulting in five levels of cognitive load will be examined here. In Study Three it was shown that performance decrements are quite linear for the within-modal conditions, yet for the cross-modal condition performance appeared to hold across the LCL and MCL conditions, and then decline steeply between the MCL and HCL conditions. To obtain a clearer understanding of the performance decline as a function of cognitive load two new levels, one in between LCL and MCL conditions, and the other between the MCL and HCL conditions are introduced.

Secondly, the question of theoretical and practical interest is how exactly the two systems work together – in interactive tiers or whether they are isolated structures that work in a complementary fashion, in coordinated tiers. In the former, resources would always be supplied direct to each modality via modality-specific reservoirs and when a task requires more resources than are available in the modality-specific reservoir, a general-purpose unitary-resource ‘tops up’ the modality specific reservoirs which then feed through these additional resources to the cognitive areas engaged in the task. In contrast, in the latter the unitary-resource system would not rejuvenate the modality-specific reservoirs but instead would assume direct contact with the cognitive areas engaged in the task. An example of these two structures is presented in Figure 7.1.
Figure 7.1. Depiction of the various attentional resource supply. (a) Two LCL tasks resourced by separate modality specific resources. (b) The two LCL tasks increase in difficulty (becoming HCL tasks) whereby the additional resources are appropriated by: (i) modality-specific stores ‘topped up’ by the modality-independent store (ii) a modality-independent store which assume absolute responsibility for resourcing.
The dual-task stimuli used in Studies One, Two and Three will be used with additional single-task and tri-task trials included to allow for comparisons to be made between the actual resource limitations empirically observed and the limitations predicted by each of these models. Specifically, while in previous studies the presence of a cross-modal advantage has been used to support the presence of modality-specific resource reserves, it has not yet been conclusively shown whether cross-modal presentation facilitates performance that is equal to if the auditory and visual tasks were actually performed in isolation, or, whether, cross-modal presentation merely diffuses some level of the interference seen in within-modal processing. The incorporation of single-task trials in this study will allow comparisons between single-task auditory and visual processing within dual-task cross-modal processing to answer this question. While both these outcomes still support the presence of modality-specific reserves from a practical perspective it is worth, while in the processing of developing a working model, knowing whether the modality-specific reserves do entirely, or only in part, efface the detrimental effects of multi-task processing. Additionally, the inclusion of the tri-task trials will allow comparisons to be made between dual-task cross-modal processing and tri-task processing to permit performance evaluations which will clarify whether an interactive or coordinated hierarchical model is the most accurate representation of the human attentional resource system.

It is predicted that if the relationship between the two hierarchical tiers is interactive the addition of a new auditory or visual task, when the modality-specific reservoirs are already overloaded (as in 7.1b(i) above), should result in greater overload and thus a reduction in performance of both the auditory and visual tasks already being completed. In contrast, should the two tiers be coordinated, the same addition of a new auditory or visual task should not affect completion of either of the primary tasks and the new task should draw
only from the modality-specific resources and thus not complete for resources with the primary tasks (as in 7.1b(ii)).

Therefore, the following aims of this study are:

(1) In the performance decrement associated with increases in cognitive load, to identify accurately the point at which unitary-resources are required to facilitate processing.

Hypothesis 1: A cross-modal advantage should be evidenced under low levels of cognitive load which will then diminish as cognitive load increases.

(2) To determine how multiple- and unitary-resource tiers work together to facilitate multi-task processing under varying levels of load.

Hypothesis 2a: If the modality-specific and modality-independent resources are arranged in an interactive hierarchical structure then dual-task performance should be superior to tri-task performance.

Hypothesis 2b: If the modality-specific and modality-independent resources are arranged in a coordinated hierarchical structure then dual- and tri-task performance should be equivalent.

Participant perceptions of task complexity, task difficulty and task confidence will again be evaluated in this study so that comparisons can be with previous studies. Moreover, the repeated systematic collection of these data if accompanied by similar outcomes, in this and previous studies, will provide a strong basis on which to draw conclusions about human attentional processing.
7.2 EXPERIMENTAL DESIGN

The design of this experiment is a partial 5 x 7 factorial design with two independent variables: Modality Combination and Cognitive Load. In a similar fashion to the previous three experiments the modality combination independent variable represents the differing single-, dual- and tri-task combinations of letter sequences required to examine the various resource limitations in processing. In this experiment an additional number of single- and tri-task trials are added to the dual-tasks to allow additional contrasts to be made. The modality combination IV will consist of seven levels: (i) a single auditory task (A); (ii) a single visual task (V); (iii) two auditory tasks (AA); (iv) two visual tasks (VV); (v) an auditory and a visual task (AV); (vi) two auditory and one visual tasks (AAV), and; (vii) one auditory and two visual tasks (AVV). The cognitive load independent variable will be manipulated similarly the manipulation in Studies 2 and 3 (number of items within a given time frame) and will here consist of five levels of load: (i) Low Load (LCL) consisting of 2 stimulus sequences with three letters in each; (ii) Low-Medium Cognitive Load (LMCL) consisting of 1 sequence with three letters and 1 sequence of 6 letters); (iii) Medium Cognitive Load (MCL) consisting of 2 sequence with six letters in each; (iv) Medium-High Cognitive Load (MHCL) consisting of 1 sequence with six letters and 1 sequence of twelve letters, and; (v) High Cognitive Load (HCL) consisting of 2 stimulus sequences with twelve letters in each.

In this experiment a truncated but carefully focussed set of planned contrasts will be conducted as follows. Firstly, as the single-task trials only consist of 3 levels of cognitive load¹, contrasts between single- and dual-task trials will be conducted using just three levels of cognitive load. Secondly, in the tri-task conditions it is only necessary to evaluate primary

¹ All single-, dual- and tri-task trials were constructed based on 3-letter, 6-letter and 12-letter strings. Accordingly, in dual-task conditions participants were faced with either 2x3-letter (LCL), 1x3-letter+1x6-letter (LMCL), 2x6-letter (MCL), 1x3-letter+1x12-letter (MHCL) and 2x12-letter (HCL). In single-task trials only 1x3-letter (LCL), 1x6-letter (MCL) and 1x12-letter (HCL) were possible arrangements. As such, in single-versus dual-task comparisons only those shared cognitive load levels (LCL, MCL and HCL) were valid contrasts.
cross-modal conditions where high load is known to occur (as per previous studies) and to contrast performance on these trials with and without the addition of a secondary single-task trial known to impose load. Therefore, the main set of analyses will involve (i) investigating the dual-task modality combination conditions (AA, VV and AV) across the five levels of cognitive load (LCL, LMCL, MCL, MHCL and HCL), followed by (ii) analyses comparing single- to dual task trials (A, V, AA, VV and AV) across three levels of cognitive load (LCL, MCL and HCL) and then (iii) analyses comparing cross-modal dual task processing with tri-task processing (AV, AVa and AVv). To aid understanding the trials of interest are displayed in Table 7.1.

Subjective questions regarding ratings of task complexity, task difficulty and task confidence will again be collected in this study. Given the role of motivation in moderating performance (Study 3) a fourth question was included in the now 4-item rating questionnaire so that any participants rating significantly different\(^2\) to others on this question could be excluded from the study so as not to confound results. As in the previous studies the DVs of performance (as measured by target detection accuracy) together with the ratings on the task complexity, task difficulty and task confidence questions were evaluated.

7.3 Method

7.3.1 Participants

In this study 28 participants took part of these there were 16 males with a mean age of 25.3 years (SD=2.2) and 12 females with a mean age of 26.4 years (SD=3.8). All participants reported normal hearing and normal (or corrected-to-normal) visual acuity. Participants provided informed consent and were made aware they could withdraw this consent at any time without penalty. Participants received course credit for their participation in this study.

\(^2\) As per outlier calculation in SPSS.
Table 7.1

Outline of Modality Combination x Cognitive Load Trials of Interest

<table>
<thead>
<tr>
<th>Trial Name</th>
<th>Modality Combination</th>
<th>Cognitive Load Level</th>
<th>Stimulus Characterisation</th>
<th>A₁ letters</th>
<th>A₆ letters</th>
<th>A₁₂ letters</th>
<th>V₃ letters</th>
<th>V₆ letters</th>
<th>V₁₂ letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>ST - Auditory Only</td>
<td>LCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>ST – Auditory Only</td>
<td>MCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>ST – Auditory Only</td>
<td>HCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V1</td>
<td>ST - Visual Only</td>
<td>LCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>ST – Visual Only</td>
<td>MCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V5</td>
<td>ST – Visual Only</td>
<td>HCL</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA1</td>
<td>DT – Auditory/Auditory</td>
<td>LCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA2</td>
<td>DT – Auditory/Auditory</td>
<td>LMCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA3</td>
<td>DT – Auditory/Auditory</td>
<td>MCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA4</td>
<td>DT – Auditory/Auditory</td>
<td>MHCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA5</td>
<td>DT – Auditory/Auditory</td>
<td>HCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV1</td>
<td>DT – Visual/Visual</td>
<td>LCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV2</td>
<td>DT – Visual/Visual</td>
<td>MLCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV3</td>
<td>DT – Visual/Visual</td>
<td>MCL</td>
<td>✓ ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.1 (Cont.)

Outline of Modality Combination x Cognitive Load Trials of Interest

<table>
<thead>
<tr>
<th>Trial Name</th>
<th>Modality Combination</th>
<th>Cognitive Load Description</th>
<th>Stimulus Characterisation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ST=single-task, DT=dual-task, TT=tri-task</td>
<td></td>
<td>A 3 letters</td>
</tr>
<tr>
<td>VV5</td>
<td>DT – Visual/Visual</td>
<td>HCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV1</td>
<td>DT – Auditory/Visual</td>
<td>LCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV2</td>
<td>DT – Auditory/Visual</td>
<td>LMCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV3</td>
<td>DT – Auditory/Visual</td>
<td>MCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV4</td>
<td>DT – Auditory/Visual</td>
<td>MHCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV5</td>
<td>DT – Auditory/Visual</td>
<td>HCL</td>
<td></td>
</tr>
<tr>
<td>AV5+A1 (aka. AVa)</td>
<td>TT – Auditory/Visual + Auditory</td>
<td>HCL + LCL</td>
<td>✓</td>
</tr>
<tr>
<td>AV5+V1 (aka. AVv)</td>
<td>TT – Auditory/Visual + Visual</td>
<td>HCL + LCL</td>
<td>✓</td>
</tr>
</tbody>
</table>
7.3.2 Materials and Apparatus

As in the previous studies, the experiment was conducted on a LG 17” widescreen laptop with attached loud speakers (Genelec Analogue 50/60Hz), with the stimuli presentation and participant responses run and recorded via DMDX software (Forster & Forster, 2003). Auditory stimuli were presented to participants via the loud speakers which were vertically offset (by 30 mm) from the presentation of the visual stimuli at the top of the computer screen. Again, participants used the keyboard to progress through the experiment; [spacebar] to key through onscreen instructions and the left (colour coded blue) and right (colour coded red) [shift] keys to indicate identification of the target letter presentations during each trial.

A reflections questionnaire was included in the final stages the experiment with trials and subsequent questionnaire completion being conducted in the same manner as the previous experiments. The questions required ratings of: (i) cognitive complexity (ii) trial difficulty (iii) target identification confidence and (iv) motivation. A full copy of this questionnaire can be viewed in Appendix K.

7.3.3 Stimuli

The auditory and visual letter stimuli and letter sequence strings used in Study Two were again used in this study. As such, the target letter “B” remained and the distractor letters (Auditory: V, E, D, C, G, U, X, Z, I and L; Visual: F, E, D, P, R, Y, J, T, N and M) were also reused. The auditory and visual letter strings used to manipulate extrinsic cognitive load in the previous study were again used in this study. In this study the strings could be presented in either single-task strings, dual-task string sequences or tri-task string sequences. As such, single letter strings consisting of each 3, 6 and 12 letters used in Study 3 were presented to participants independently, these strings were then matched into pairs of AA, VV or AV
sequences (using the same template as in Study 3), and then from these string sequences tri-task trials were created by adding a single letter string (consisting of 3 letters) to the AV sequences (consisting of 12 letters on each presentation side).

The timing sequences used in Study 3 were again used in this experiment, however, an additional timing sequences was required in order to present the tri-task trials. In this new sequence the standard AV high cognitive load sequence (2 x 12 letters) was matched with a single A or V low cognitive load string (1 x 3 letters). The AV task was considered the primary task and the additional LCL string was considered the secondary task. The LCL string was presented to the participant on the opposite side of the screen/loud speakers to that of the same-modality primary task. That is, if the primary auditory task was being presented to the LHS loud speakers and the primary visual task presented on the RHS of the computer screen then a secondary auditory task would be presented to the RHS loud speaker, or, a secondary visual-task to the LHS of the computer screen. In keeping with the timing sequences used in previous trials this meant that the letters presented in the LCL secondary task would need to overlap with letters in the primary task. To minimise confusion the simultaneous presentation of the LCL secondary task coincided with the presentation of the opposite modality letters (as having 2 auditory letters presented at the exact same onset may obscure the clarity of each). Moreover, the three target letters being presented were offset so that multiple responses were not required at the same time. For clarification of these timing sequences see Figure 7.2.

7.3.4 Procedure

As in the previous studies participants were seated in front of the laptop computer within a small laboratory testing room, so that the computer screen and accompanying auditory loud speakers were at approximately chest height. Firstly, the participants completed
Figure 7.2. Template for letter presentation sequences in the tri-task trials. (a) A primary AV task together with a secondary visual task. Note that the primary visual task is on the RHS, while the secondary visual task is on the LHS and that letter for letter these do not onset at the same time. (a) A primary AV task together with a secondary auditory task. Note that the primary auditory task is on the LHS, while the secondary visual task is on the RHS and that letter for letter these do not onset at the same time.
36 practice trials (1 trials from each of the trial types being evaluated in this experiment) so that they would be familiar with the stimuli, timing sequences, experiment progression, and the required responses. Secondly, the experimenter reviewed their practice trials responses to ensure their performances were appropriate (i.e., they were responding to the appropriate number of targets and they were responding immediately following the target presentation and not at the conclusion of the trial). After this small evaluation procedure participants’ began the experimental trials once the experimenter was happy the participants understood what was required.

The experiment was organised in experimental blocks, with 29 trials present in each block. Participants initiated the beginning of each block by pressing the [spacebar] and with each subsequent trial beginning automatically 10 seconds after the previous one concluded. At the conclusion of each trial block, participants were given a self-determined break and pressed the [spacebar] to commence the next experimental block. In total participants completed 360 trials arranged into 12 trial blocks (30 trials per block) taking approximately 1 hour to complete.

Following completion of the 12 trial blocks participants completed a further 36 ‘reflection’ trials – the same as the practice trials – after which they completed a section of the questionnaire specific to each individual trial. This second stage of the experiment took approximately 20 minutes.

7.4 RESULTS

7.4.1 Data Reduction

Data reduction was undertaken in the same manner as it was in the previous studies for the accuracy data, and the complexity, difficulty and confidence ratings data. The responses to the motivation trials were then used to decipher any individuals who differed
greatly from the mean motivation scores who would then be excluded from analysis given the role motivation may play in moderating cognitive load (as found in Study Three).

In this study valid responses were calculated using the individualised RT window method established and used in Study Three: evaluating all the single-task trials where a single [shift] press was recorded by the participant and establishing a mean RT for these trials. The valid RT window was then established to be $\pm 2SD$ around this mean (see Appendix L for a table of each participants’ valid RT window).

Accuracy in single- and dual-task trials was calculated by comparing the number of trials in which the maximum number of targets were correctly identified to those where the maximum number of targets were not identified to give a percent accuracy score. In the tri-task trials accuracy was determined by comparing only those trials where the secondary task target was correctly identified (this ensures participants were actually partaking in the tri-task nature of the trial) and then comparing those where both primary targets were identified compared to those where both targets weren’t identified to give a percent accuracy score.

7.4.2 Data Screening

The data for any participant who recorded non-standard ratings of motivation were removed from the analysis. To enable this motivation scores for each of the 36 ratings items were entered into SPSS and a frequency and descriptives analysis was conducted to determining any outlying scores. On this basis the data for one participant were removed from further analysis as ratings of over 5.5 (on a 7 point likert scale) were recorded for all trials while in comparison the average response rating from the rest of participants was 3.45 ($SD=1.1$).

Data screening using the Statical Package for the Social Sciences (SPSS) was then undertaken prior to conducting inferential analyses. This process revealed no univariate or
multivariate outliers based on the accuracy scores resulting in the data for 27 participants being retained. The assumptions of normality were considered satisfactory although slight evidence of kurtosis was evident for the accuracy data. However, given the robustness of ANOVAs to assumption violation with this number of participants this minor violation was not considered enough to warrant the use of non-parametric analyses.

7.4.3 Data Analyses Overview

To evaluate modality combination by cognitive load a 3 x 5 repeated measures ANOVA was conducted with the same contrasts as in Study Three were except that as the cognitive load factor now includes five levels, linear, quadratic and higher order trends over load were tested.

To compare single- and dual-task trials a 5 x 3 repeated measures ANOVA was conducted. The 5-level modality combination factor allows contrasts between the two single-task conditions and the three dual-task conditions: A, V, AA, VV and AV. As only 3 levels of cognitive load could be manipulated in the single-task conditions only the LCL, MCL and HCL dual-task trials were contrasted in this analysis. Therefore, linear, quadratic and higher order trends were evaluated in the modality combination conditions, while only linear and quadratic trends were tested for cognitive load. However, given the experimental hypotheses concentrating on only shallow and steep trend changes over the increasing cognitive load levels focus will be given predominantly to linear and quadratic trend analyses, with higher order trend outcomes being evaluated in terms of the linear and quadratic trend hypotheses.

To compare dual- and tri-task linear and quadratic contrasts were conducted to compare AV with AVa and AVv trials, and later AVa and AVv. As these comparisons were conducted to examine the effect of an additional low load task when an individual was
already engaging a high load task no comparisons across all the levels of load were necessary.

For the task complexity, task difficulty and task confidence one-way ANOVAs (with planned comparisons) were conducted for each of these subjective measures to compare ratings on the 3 task types (single-, dual- and tri-tasks). Following this, a 3 x 5 repeated measures ANOVA was conducted on each of the ratings scores to compare cross- and within-modal dual-task difficulty ratings across the five cognitive load levels.

Details of all the planned comparisons in this experiment can be found in Appendix M.

7.4.4 Analyses

7.4.2.1 Modality Combination and Cognitive Load: Performance Across Five Levels of Load

As shown in Figure 7.3 overall cross-modal presentation results in far superior performance to either of the within-modal presentations \( F_{AV \ vs. \ AA+VV}(1,27)=80.64, \ p<.001, \) partial \( \eta^2=.75; \ F_{AA \ vs. \ VV}(1,27)=.94, \ p=.342, \) partial \( \eta^2=.03).\) The deleterious effect of cognitive load on accuracy increases as cognitive load conditions increases, starting more shallow under lower load conditions, before a sharp increases in evident under the higher levels of load \( F_{Lin}(1,27)=441.53, \ p<.001, \) partial \( \eta^2=.94; \ F_{Quad}(1,27)=43.01, \ p<.001, \) partial \( \eta^2=.61).\)

Turing to the modality combination and cognitive load interaction while both within-modal conditions show steady decrements across the five levels of cognitive load\(^3\) an entirely different pattern is seen for AV. The cross-modal performance decrements remain quite shallow across the first four levels of cognitive load (LCL through to MHCL) followed by a very sharp decline in performance between MHCL and HCL \( F_{AV \ vs. \ AA+VV*Lin}(1,27)=14.15, \ p=.001, \) partial \( \eta^2=.34; \ F_{AV \ vs. \ AA+VV*Quad}(1,27)=839.19, \ p<.001, \) partial \( \eta^2=.59).\)

\(^3\) \( F_{AA \ vs. \ VV*Lin}(1,27)=.42, \ p=.522, \) partial \( \eta^2=.02; \ F_{AA \ vs. \ VV*Quad}(1,27)=.72, \ p=.705, \) partial \( \eta^2=.03; \ F_{AA \ vs. \ VV*Cubic}(1,27)=.27, \ p=.611, \) partial \( \eta^2=.01; \ F_{AA \ vs. \ VV*Quartic}(1,27)=2.58, \ p=.120, \) partial \( \eta^2=.09\)
To investigate the exact point of inflection for the cross- compared with the within-modal conditions (given the significant cubic and quartic trends) two additional contrasts were included to test the shape of the curves from LCL to MCL, and a second set of contrasts to test the shape of the curve from MCL to HCL conditions.

For the LCL to MCL portion of the curve overall cross-modal performance is higher than within-modal performance ($F_{AV \text{ vs. } AA+VV}(1,27)=65.55, p<.001$, partial $\eta^2=.71$) while AA and VV performance are relatively the same as one another ($F_{AA \text{ vs. } VV}(1,27)=.69, p=.414$, partial $\eta^2=.03$); and there were also overall linear and quadratic effects of cognitive load indicating that overall decreases in cognitive load are steeper to begin with (LCL $\rightarrow$ LMCL).
and then shallower (LMCL $\rightarrow$ MCL) ($F_{Lin}(1,27)=45.49$, $p<.001$, partial $\eta^2=.62$; $F_{Quad}(1,27)=9.97$, $p=.004$, partial $\eta^2=.27$).

For the MCL to HCL portion of the curve cross-modal was higher than within-modal performance ($F_{AV vs.AA+VV}(1,27)=90.21$, $p<.001$, partial $\eta^2=.77$) while AA and VV performance are relatively the same as one another ($F_{AA vs.VV}(1,27)=2.51$, $p=.124$, partial $\eta^2=.09$); and there were also overall linear and quadratic effects of cognitive load ($F_{Lin}(1,27)=279.84$, $p<.001$, partial $\eta^2=.91$; $F_{Quad}(1,27)=50.89$, $p<.001$, partial $\eta^2=.65$). These linear and quadratic effects of load are qualified by interactions of the linear and quadratic trend effects of load with the cross-modal versus within-modal contrast\(^4\), but no such interactions of trends between the within-modal contrasts\(^5\). As can be seen in Figure 7.3 both AA and VV performance decrease at the same rate from MCL to HCL whereas AV performance initially follows a similar shallow pattern (from MCL $\rightarrow$ MHCL) before a very sudden and steep decline in performance from MHCL to HCL conditions.

Taken together these results suggest that the higher order effects demonstrated in the original (3x5) analysis was likely an effect of the minimal blip in the VV modality at the MCL level -- rather than evidence for an unusual decreasing pattern for the AV compared to AA+VV which might have suggested alternative appraisals of the multiple- plus unitary-resource system proposed here.

### 7.4.4.2 Single- vs. Dual-Task Performance

Accuracy performance for single- (A and V) and dual- (AA, VV and AV) task trials across three increasing levels of cognitive load is shown in Figure 7.4. Overall it would appear that completing a single task allows near perfect performance (at least for those stimuli used here) while performing two tasks simultaneously leads to a substantial overall

\(^4\) $F_{AV vs.AA+VV*Lin}(1,27)=44.46$, $p<.001$, partial $\eta^2=.62$; $F_{AV vs.AA+VV*Quad}(1,27)=82.74$, $p<.001$, partial $\eta^2=.75$

\(^5\) $F_{AA vs.VV*Lin}(1,27)=.75$, $p=.395$, partial $\eta^2=.03$; $F_{AA vs.VV*Quad}(1,27)=2.52$, $p=.124$, partial $\eta^2=.09$
decrease in performance \( (F_{Single\ vs.\ Dual}(1,27)=246.36, p<.001, \text{ partial } \eta^2=.90) \). However, if there are two tasks to be completed simultaneously, then cross-modal presentation results in substantially better performance than does within-modal presentation \( (F_{AV\ vs.\ AA+VV}(1,27)=73.28, p<.001, \text{ partial } \eta^2=.73) \) with no difference between the both auditory and both visual dual-tasks \( (F_{AA\ vs.\ VV}(1,27)=.99, p=.382, \text{ partial } \eta^2=.04) \) and auditory (AA+A) and visual (VV+V) tasks more generally \( (F_{Auditory\ vs.\ Visual}(1,27)=2.74, p=.11, \text{ partial } \eta^2=.20) \).

As is evident in Figure 7.4, and consistent with previous studies, while the effect of cognitive load remains quite consistent for the two different within-modal dual-task presentations, AA and VV, \( (F_{AA\ vs.\ VV*Lin}(1,27)=.39, p=.539, \text{ partial } \eta^2=.02; \ F_{AA\ vs.\ VV*Quad}(1,27)=3.34, p=.079, \text{ partial } \eta^2=.11) \), for these single modality tasks the trend over
load is consistently linear, whereas for cross-modal dual-task presentation cognitive load increases initially have a moderate effect on performance (between LCL and MCL) before performance suddenly and dramatically deteriorates (from MCL to HCL) \((F_{AV vs. AA+VV*Lin}(1,27)=38.46, p<.001, \text{partial } \eta^2=.58; F_{AV vs. AA+VV*Quad}(1,27)=26.09, p<.001, \text{partial } \eta^2=.49)\). Single-task performance decrements appear to produce a curve with an opposite shape to that in the cross-modal condition, with performance most affected between the two lower levels of cognitive load before a flattening of the performance decrement between the two higher levels \((F_{Dual vs. Single*Lin}(1,27)=63.90, p<.001, \text{partial } \eta^2=.70; F_{AV vs. Dual vs. Single*Quad}(1,27)=30.72, p<.001, \text{partial } \eta^2=.53)\). In fact in reviewing Figure 7.3 performance decrements between the LCL and MCL conditions is similar for all cross-, within- and single-task trials (albeit with differing starting performance), the differences in performance changes manifest primarily between the MCL and HCL conditions with no interaction effects present \((F_{A+AA vs. V+VV*Lin}(1,27)=0.24, p>.05, \text{partial } \eta^2=.10; F_{A+AA vs. V+VV*Quad}(1,27)=3.56, p>.05, \text{partial } \eta^2=.11)\).

7.4.4.3 Dual- vs. Tri-Task Performance

As can be seen in Figure 7.5 performance was better in the dual-task cross-modal condition than in either of the tri-task conditions. It appears that the addition of the tertiary task (which imposes minimal cognitive load) does substantially reduce performance on both the auditory and visual aspects of the primary AV task \((F_{Dual vs. Tr}(1,27)=84.69, p<.001, \text{partial } \eta^2=.76)\). Moreover, the effects of the tertiary task are more deleterious if this secondary low cognitive load task is auditory as opposed to visual \((F_{AV_a vs. AV_v}(1,27)=12.89, p=.001, \text{partial } \eta^2=.33)\).
Figure 7.5. Mean participant percent accuracy scores (+SE) in cross-modal (HCL) trials without (dark grey) and with (light grey) the addition of a secondary (LCL) task.

7.4.4.4 Complexity, Difficulty and Confidence Ratings

Given the importance of the subjective measures of task variables in predicting performance (see Chapter 6) it was considered important to obtain participants’ ratings while completing the varying single-, dual- and tri-task trials.

Task Complexity. As can be seen in Figure 7.6 single-task trials are generally rated as being less complex than are dual-task trials ($F_{\text{Single vs. Dual}}(1,27)=481.83$, $p<.001$, partial $\eta^2=.95$), while dual-task trials are rated in turn as being less complex than the tri-task trials ($F_{\text{Dual vs. Tri}}(1,27)=941.24$, $p<.001$, partial $\eta^2=.97$). Overall, visual-only tasks are rated as less complex than their auditory-only counterparts ($F_{\text{A+AA vs. V+VV}}(1,27)=52.44$, $p<.001$, partial
η²=.66) but when only dual-task within-modal trials are compared this difference is no longer apparent ($F_{AA \text{ vs} VV}(1,27)=.02$, $p=.882$, partial $\eta^2=.01$) suggesting that the main variance is within the single-task trials. Moreover, the he dual-task cross- and within-modal trials are rated to be equivalent in terms of complexity ($F_{AV \text{ vs} AA+VV}(1,27)=1.61$, $p=216$, partial $\eta^2=.06$) suggesting that the critical factor for complexity in dual-tasks is simple the number of tasks. Finally, the addition of the tertiary auditory task increases rated complexity more than when this tertiary task is visual ($F_{AVa \text{ vs} AV}(1,27)=15.68$, $p<.001$, partial $\eta^2=.37$).

![Figure 7.6](image.png)

*Figure 7.6.* Mean ratings of task complexity (+ SE) for each of the modality combinations. Single-task conditions are shown in light grey, dual-task conditions in darker grey and tri-task conditions in black.

Overall, this seems to suggest that participants associate greater complexity with auditory stimuli – although this effect was not present for dual-task trials – and that
increasing the number of stimulus strings within the trial is associated with perceptions of increasing task complexity.

In turning to the modality combination and 5-level cognitive load contrasts as can be seen in Figure 7.7 there is no modality effect for complexity of dual tasks; AA, VV and AV trials are rated similarly overall ($F_{AV, AA+VV}(1,27)=1.61, p=.216$, partial $\eta^2=.06$; $F_{AA, VV}(1,27)=.02, p=.882$, partial $\eta^2=.01$). There is a linear, quadratic and cubic effect of load indicating that overall perceptions of complexity increase as cognitive load increases although this increase is not perfectly linear ($F_{Lin}(1,27)=1705.15, p<.001$, partial $\eta^2=.98$; $F_{Quad}(1,27)=58.38, p<.001$, partial $\eta^2=.68$; $F_{Cubic}(1,27)=7.26, p<.012$, partial $\eta^2=.21$). As there are no interactions between modality combination and cognitive load these trends apply to the increasing perceptions of complexity across all three dual tasks (AA, VV or AV).

![Figure 7.7](image-url)  
*Figure 7.7. Mean Ratings of task complexity (+SE) for the dual-task conditions across the 5 levels of cognitive load manipulated.*
Task Difficulty. As can be seen in Figure 7.8 single-task trials are rated as being less difficult than are dual-task trials ($F_{Single \ vs. \ Dual}(1,27)=3462.70, p<.001$, partial $\eta^2=.99$), and in turn dual-task trials are rated as being less difficult than the tri-task trials ($F_{Dual \ vs. \ Tri}(1,27)=2326.00, p<.001$, partial $\eta^2=.99$). Generally, visual-only tasks are rated as being similarly difficult to auditory-only tasks, both overall ($F_{A+AA \ vs. \ V+VV}(1,27)=.50, p=.49$, partial $\eta^2=.02$) and in dual-task conditions ($F_{AA \ vs. \ VV}(1,27)=.84, p=.368$, partial $\eta^2=.03$). As such, the differences between the within-modal and cross-modal dual tasks (AV versus AA and VV)\(^6\) and the two tri-tasks (AVa x AVv)\(^7\) cannot be explained by features of the visual stimuli themselves, but rather can only be explained in terms of the load imposed during multi-task processing.

\[\text{Figure 7.8. Mean ratings of task difficulty (±SE) for all modality combination trials. Tri-task conditions are shown in light grey, dual-task conditions in darker grey and single-task conditions in black.}\]

\(^6\) $F_{AV \ vs. \ AA+VV}(1,27)=372.19, p<.001$, partial $\eta^2=.93$

\(^7\) $F_{AVa \ vs. \ AV}(1,27)=7.73, p=.010$, partial $\eta^2=.22$
The modality combination (AA, VV and AV) by cognitive load (five levels) interactions are depicted in Figure 7.9. Overall it appears that cross-modal presentations are rated to be less difficult than within-modal presentations ($F_{AV vs. AA+VV}(1,27)=335.11, p<.001$, partial $\eta^2=.93$), while both AA and VV presentations are rated to have similar difficulty levels ($F_{AA vs. VV}(1,27)=.91, p=.349$, partial $\eta^2=.04$). Linear and quadratic effects of cognitive load were present whereby increasing the level of cognitive load leads to an increase in the level of perceived difficulty ($F_{Linear}(1,27)=1053.17, p<.001$, partial $\eta^2=.98$; $F_{Quad}(1,27)=15.58, p=.001$, partial $\eta^2=.39$).

![Figure 7.9. Mean Ratings of task difficulty (+SE) for the dual-task conditions across the 5 levels of cognitive load manipulated.](image-url)
While no differential linear, quadratic or higher order trend components were evident in the within-modal contrast\(^8\) there were differential linear and quadratic load effects for the cross-modal versus within-modal contrast (\(F_{AV vs AA+VV^{\text{Lin}}}(1,27)=89.09, p<.001\), partial \(\eta^2=.79\); \(F_{AV vs AA+VV^{\text{Quad}}}(1,27)=96.19, p<.001\), partial \(\eta^2=.80\)). As can be seen in Figure 7.9 these results indicate that while AA and VV difficulty ratings increase at a steady and equivalent level, AV difficulty ratings initially rise quite slowly before beginning a sharper increase from MCL through to HCL. Reflecting on the accuracy data presented in Figure 7.3, the data here are almost a mirror image, presumably reflecting the strong negative predictive relationship between perceived difficulty and accuracy shown Chapter 6.

\textit{Trial Confidence}. As can be seen in Figure 7.10 there is an effect for confidence as a result of the number of tasks being simultaneously completed; highest confidence ratings are for single-task, moderate for dual-task and lowest for tri-task trials (\(F_{\text{Single vs. Dual}}(1,27)=18.67, p<.001\), partial \(\eta^2=.41\); \(F_{\text{Dual vs. Tri}}(1,26)=628.07, p<.001\), partial \(\eta^2=.96\)). In terms of the dual-task trials there is no difference between cross-modal and within-modal ratings of confidence (\(F_{AV vs. AA+VV}(1,27)=3.62, p=.07\), partial \(\eta^2=.10\), however, ratings of confidence are higher for the visual than for the auditory within-modal conditions (\(F_{AA vs. VV}(1,27)=15.79, p<.001\), partial \(\eta^2=.37\)). Finally for the tri-tasks, given that the results so far suggest that confidence decreases as the number of auditory tasks to be performed increases, it is of little surprise that the addition of a tertiary auditory task is more detrimental than the addition of a tertiary visual task (\(F_{AVa vs. AVi}(1,27)=169.00, p<.001\), partial \(\eta^2=.86\)).

---

\(^8\) \(F_{AA vs VV^{\text{Lin}}}(1,27)=.213, p=.649\), partial \(\eta^2=.01\); \(F_{AA vs VV^{\text{Quad}}}(1,27)=.400, p=.533\), partial \(\eta^2=.02\); \(F_{AA vs VV^{\text{Cub}}}(1,27)=.213, p=.649\), partial \(\eta^2=.01\); \(F_{AA vs VV^{\text{Quart}}}(1,27)=3.55, p=.072\), partial \(\eta^2=.13\)
As can be seen in Figure 7.11 in the 3 (modality combination) x 5 (cognitive load) interactions for dual-tasks confidence for cross-modal tasks was overall rated as being similar to that for within-modal confidence, while overall VV confidence was greater than AA confidence ($F_{AV \ vs. \ AA+VV}(2,25)=3.16, p<.001$, partial $\eta^2=.58$; $F_{AA \ vs. \ VV}(2,25)=3.15, p<.001$, partial $\eta^2=.47$). Only a linear effect of cognitive load was found ($F_{Lin}(2,25)=138.38, p<.001$, partial $\eta^2=.84$; $F_{Quad}(2,25)=1.64, p=.212$, partial $\eta^2=.06$), suggesting that as cognitive load increases confidence steadily decreases. No interactions were found suggesting that the greater confidence in VV trials is consistent across all levels of load.

Figure 7.10. Mean ratings of task confidence (+SE) for all modality combination trials. Single-task trials are shown in light grey, dual-task trials in dark grey and tri-task trials in black.
Descriptive and frequency analyses of accuracy as a function of confidence were conducted to determine whether, as in Study Three, performance was affected by this confidence particularly in conditions where auditory stimuli were absent. See Table 7.2 for a summary and comparison of participant confidence ratings and matched performance. While participants were more confident of their performance in the AV condition – which was seemingly matched within increased performance – in fact sometimes this confidence was unfounded e.g., in cases in which participants were very confident that they had identified both targets they, in fact, identified only one (and in two cases no) target(s). Given these high confidence ratings it was considered important to re-evaluate why during data reduction these trials had been coded as either 1ID or 0ID – in case perhaps the cut off window was again

*Figure 7.11.* Mean ratings of trial confidence (+SE) for all the dual-task conditions across the 5 levels of cognitive load manipulated.
truncating and excluding otherwise valid responses. This review revealed that of these ‘high confidence’ 1ID and 0ID trials only three cases could possibly be supported by the truncated response window argument. Rather it is the case that most often these trials were coded as 1ID and 0ID because a response had come before the target had even been presented, or well after the trial had concluded.

In contrast to the apparent over-confidence in the AV condition, participants often under-estimated performance in the AA condition and were most accurate in monitoring their performance in the VV conditions. Additionally, accuracy in performance monitoring in the AVv condition was superior to the AVa condition again upholding the suggestion that monitoring performance on visual tasks is more accurate than on auditory tasks.

Table 7.2

*Participant Ratings of Target Identification Confidence and Constitution of Actual Performance on These Trials*

<table>
<thead>
<tr>
<th>Modality Combination</th>
<th>A</th>
<th>V</th>
<th>AA</th>
<th>VV</th>
<th>AV</th>
<th>AVa</th>
<th>AVv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error % of trials rated as “I identified both targets detected”</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>6.9%</td>
<td>12.4%</td>
<td>13.2%</td>
<td>13.1%</td>
</tr>
<tr>
<td>% of trials rated as “unsure of my performance”</td>
<td>1%</td>
<td>0%</td>
<td>8.1%</td>
<td>1.2%</td>
<td>1.7%</td>
<td>2.5%</td>
<td>1.6%</td>
</tr>
<tr>
<td>% of “unsure” trials where both targets actually identified</td>
<td>2%</td>
<td>0%</td>
<td>60%</td>
<td>9%</td>
<td>12%</td>
<td>16%</td>
<td>5%</td>
</tr>
</tbody>
</table>
7.5 **DISCUSSION**

The general aim of this study was to probe the details of the modality-specific and modality-independent resources that may be involved in the processing of complex tasks. The first hypothesis – that a cross-modal advantage will diminish as cognitive load increases – was supported, as in the previous two studies. Moreover, with the inclusion of the additional two cognitive load manipulations (LMCL and MHCL) the maintenance of this cross-modal advantage was seen to remain shallow before a very sudden decrease in performance was seen. This indicates that the involvement of modality-independent resources are likely to occur rapidly only once modality-specific resources are, or are near, expended, rather than their involvement gradually increasing as cognitive load increases. Furthermore, with comparative single-task performance benchmarks it is possible to conclude that cross-modal processing does not merely moderate performance levels somewhere between single-task and within-modal dual-task processing but rather the patterns of cross-modal processing is more similar to single-task than to dual-task within-modal performance. This outcome further supports the notion that under conditions of low cognitive load modality-specific resources do facilitate maximum performance.

The second hypothesis (alternative (a)) was supported with dual-task performance being superior to tri-task performance suggesting that an interactive hierarchical model to be the most accurate reflection of human attentional resourcing. The results suggest that initially tasks are resourced by modality-specific reserves, before the resources necessary to sustain performance become greater than the resources available within the modality-specific stores. It was shown here during tri-tasks that the addition of a low cognitive load-inducing auditory or visual tertiary task impacted on the processing of the primary AV task. This shows that once modality-independent resources are required they are used to supplement the modality-specific reserves rather than assuming responsibility directly for task processing from the
modality-specific stores. Should the modality-independent system have assumed all responsibility for the primary AV task this would release the modality-specific stores to process other low cognitive load tasks, however, in this study it was demonstrated that an additional low cognitive load task (shown in the previous studies to be supported by modality-specific reserves) was not processed independently from the primary AV task.

The hybrid model propositions put forth by Spence and Driver (1996) and Posner (1990) were vague with respect to how such a multi-staged model would work – other than it would have limitations consistent with both the multiple- and unitary-resource models. As suggested by Burr and Alais (2006) although their results unequivocally supported modality-specific resources, “depending upon the nature of the task demands, the most sensible strategy might well be to employ a supramodal attentional resource for a given task” (pp. 256). While this suggestion was qualified to mean that a unitary-resource system may be required to process auditory and visual stimuli when they form a part of the same “event” (that is, they supply independent information corresponding to the same stimulus origin and together may aid interpretation as in multi-modal tasks like speech comprehension) the results of this study suggest that such a relationship between the auditory and visual stimuli is not necessary – only that within a given period of time the overall demands of the task become more than can be handled by a modality-specific reservoir.

In the past it has been argued that a cross-modal advantage may occur because the additive difficulty of the separate tasks does not require sufficient modality-specific resources to cause any significant resource competition – that is, the visual and auditory tasks are not equivalent (Navon, 1977). In Studies One, Two and Three here it has been shown that processing of the AA and VV tasks remains equivalent for each level of imposed cognitive load. In this Study the aim was to map the allocation of attentional resources and investigate that claim that auditory and visual stimuli exert unequal demands on cognitive resources. To
this end, the single-task versus dual-task contrasts were included. The results clearly show
that cross-modal processing, in the dual-task, does not produce differential performance
outcomes than if the individual auditory and visual components are processed individually (in
the single-task cases) This result shows that multiple modality-specific reservoirs are
involved in processing such low-level tasks. Moreover, the measures of perceived task
difficulty further support this claim with low-level cognitive load AV tasks being rated to be
of similar difficulty to low level single-task auditory and visual tasks.

The results presented here certainly support the claim of a cross-modal advantage –
but it must not be overlooked that even though dual-task cross-modal processing is
significantly superior to dual-task within-modal processing it still remains somewhat inferior
to single-task processing. It is true that the current IHR Model cannot indisputably defend
such a situation that there appears to be a level of interference not entirely accounted for
during cross-modal processing. However, it is not the argument in this thesis that the hybrid
model is a definitive model of human attentional processing – but merely an important
collaboration between the multiple-resource and unitary-resource approaches currently on
offer. Certainly, in addition to strict auditory and visual attentional resources there may be
other for ms of resourcing. For instance, the Multiple-Resource Model (Wickens, 1982),
while reflected upon primarily in this thesis as a model of separate auditory and visual
resources (modalities), also proposed alternative mechanisms from which interference may
arise; stages (perception vs. cognition vs. response), responses (manual spatial vs. vocal
verbal), codes (spatial vs. verbal) and visual processing (focal vs. ambient). According to the
Multiple Resources Cube (Wickens & McCarley, 2007) one other mechanism which may
contribute to the results in these experiences is the spatial-verbal code. This may be an
alternative way of understanding why some level of interference is still present during dual-
task cross-modal situations. In such cases, irrespective of the processing modalities, the tasks
are both language based and therefore may compete for such ‘verbal’ resources. The competition resulting from this code well explains why there remains some level of competition during cross-modal processing (compared to single task processing). This competition can not be explained by reference to the modality dichotomy of the Cube alone. Investigating these alternative sites of resource competition will be important to the future development of the IHR Model.

More generally the results of this study support those of Studies Two and Three by showing similar results for the dual-task modality combination by cognitive load interactions. This is a valuable outcome; the reliability of the interaction between modality combination and cognitive load has been consistently shown across all three these studies. By including the additional two levels of cognitive load in this study a much clear picture emerges regarding the movement from modality-specific to modality-specific+independent processing. As far as the present stimuli and cognitive load manipulations allow, the switch appears to be quite sudden – rather than a gradually developing situation. It would seem that the modality-specific reserves handle processing until they are expended and then suddenly are replenished by modality-independent resources – rather than a situation in which they are continuously topped up even when only minimal resources are expended.

In the final chapter the results of all four studies will be reviewed especially with regard to conclusions concerning the Interactive Hierarchical Resource Model of Auditory and Visual Attention and the roles of each perceived complexity, difficulty and confidence in moderating extraneous, intrinsic and germane cognitive load.
Chapter 8

GENERAL DISCUSSION
8.1 **CHAPTER OVERVIEW**

The aim of this thesis was to determine whether the apparent divide in support between the multiple- and unitary-resource models could be explained by the levels of cognitive load engendered by the experimental methods employed in those studies. In general this aim was achieved with the results of a series of four studies generally showing that cognitive load does impact the presence and size of any cross-modal advantage – the main point of difference between the predictions of the two resource models. Specifically, across the studies conducted here where stimulus presentation rate was used to manipulate imposed load (Studies 2-4), under conditions of low cognitive load a cross-modal advantage was present – and outcome consistent with the predictions of a multiple-resource model (Wickens, 2007). However, as cognitive load increased the size of this advantage correspondingly decreased until there was no longer a significant effect – an outcome consistent with the predictions of the unitary-resource model (Arnell & Duncan, 2002).

Over and above this general finding there were many fine nuances of the relationship between cognitive load (and its sub-types), cognitive complexity and performance found in these experiments which should not go unrecognised. These will be considered in order to gain a fuller perspective of the load-model relationship ahead of a full appraisal of the Interactive Hierarchical Resource (IHR) Model. Consideration will then be given to the limitations of the studies and their results with proposals for how future research may seek to rectify such aspects. The chapter will conclude with a discussion of the implications of these outcomes for cognitive psychological research, human factors research and more applied fields such as workplace design.
8.2 Overview of Experimental Results

In Study One there appears to be little change in the size of the cross-modal advantage across the three levels of cognitive load manipulation, in fact the presence of the advantage itself was questionable for although cross-modal was consistently better than within-modal performance this difference was not significant in this experiment. These results may have been taken to indicate that processing was not facilitated by modality-specific resources at all, however the subjective ratings of complexity and difficulty suggested that perhaps the cognitive load manipulation was ineffective in producing the expected effect; while changes in the features of the task were recognised across the three manipulation levels (via appropriate changes in cognitive complexity ratings) these were not matched by changes in perceived difficulty. This minimal modality advantage in cross- compared to within-modal trials when paired with a slightly decreased rating of difficulty for cross- vs. within-modal trials suggests that extraneous cognitive load may indeed impact performance. That is, when information is presented across two modalities cognitive load is slightly lower and performance slightly superior to when information is presented to only one modality.

The cognitive load manipulation was consequently altered in Study Two - from a manipulation of target/distractor (dis)similarity to the number of targets and distractors presented within the trial duration, that is, rate of letter presentation. In Study Two a robust cross-modal advantage was evident in the low load condition, a reduced but still significant advantage in the medium cognitive load condition, and no significant effect in the high cognitive load condition. Moreover, complexity and difficulty ratings in this study were positively related to one another and both were inversely related to performance. It was concluded that the cognitive load manipulation in Study 2 was far more effective than that used in Study One. Additionally, the more robust findings in Study Two can also be attributed to the additional intrinsic cognitive load manipulations (the alterations in letter
presentation rate across imposed cognitive load levels) over and above to the effects of extraneous cognitive load.

The purpose of Study Three was to evaluate any additional effects of the contribution of germane cognitive load. The importance of evaluating such load, that imposed by the operator-task relationships, was warranted both for theoretical reasons – germane cognitive load forms the final factor of the Brünken et al. (2003) taxonomy – but also because the substantial individual differences observed in Study Two were reasoned to possibly result from some individuals’ prior task exposure and/or engagement in the experimental task. The cross-modal advantages found in Study Two were repeated in Study Three and additionally it could be seen that individuals who were both motivated and exposed to the task (compared to those unmotivated and less exposed) recorded significantly lower perceived difficulty ratings in post- compared to pre-test conditions and corresponding improved performance from pre- to post-test. Moreover, as there were no pre- to post-test changes in ratings of task complexity it can be concluded that participants did not believe any features of the task itself changed even though their experience of difficulty did. As such, the contribution of germane cognitive load (Study Three) was seen to impact performance over and above extraneous load (Studies One and Two) and intrinsic load (Study Two). The outcomes Study Three suggest that the human attentional system has limitations at both modality-specific and modality-independent levels, as would be proposed in a hierarchical resource model.

Of additional interest in this study were the contrasts between ratings of complexity, difficulty and confidence. While it was of little surprise that confidence was found to be a different construct to the other two, it was revealed that complexity and difficulty were also quite separate concepts. Moreover, while complexity was a predictor of performance, perceived difficulty was a superior predictor. In addition, while confidence was not a predictor of AA or AV performance it was for VV. An appraisal of confidence ratings and
performance accuracy together suggest that perhaps individuals are more certain about their performance when they can physically see something, rather than when they are required to hear something. It is possible that this may a result of the majority of sensory neurons in humans featuring in the visual system (Sivak, 1996) possibly making this the most relied upon sense, but further research is required to investigate this issue more fully.

Finally, in Study Four the possible structure of a hierarchical resource model was explored based on the understanding from Studies Two and Three that this model allows for some level of both modality-specific (which would facilitate competition-free cross-modal processing for a time) and modality-independent (at which point competition-free cross-modal processing would cease) resources which maintain performance under different operating conditions (specifically, different cognitive load conditions). Two options were considered for how such a two-tiered structure might function: interactive tiers in which exhausted modality-independent resources are topped up by modality-independent resources; and co-ordinated tiers in which modality-independent resources subsume the responsibilities of the modality-specific resources when processing requirements exceed the modality-specific capacity.

The results of this final study indicated that once cross-modal interference was evident (a sign that the modality-specific resources could no longer independently maintain performance and modality-independent resources were being consumed) additional low cognitive load tasks further impeded processing (as can be seen in Figure 7.4). This was interpreted to mean that the tiers work in an interactive relationship because, following the contribution of modality-independent resources, modality-specific resources are not released to handle other low level processing (as predicted by the co-ordinated hierarchical hypothesis) but instead remain involved in the primary processing of the initial dual-task. This hybrid interactive hierarchical resource hypothesis was offered as a potential model for
auditory and visual attentional processing that can explain how cognitive load in the current experiments, and in the previous research reviewed, moderates how tasks are resourced.

It is appropriate here to acknowledge an assumption made throughout all these studies; that AA and VV performance is generally equivalent. While the results reported in each study certainly reveal equivalence in performance measures – it is possible that the competition documented in both types of within-modality processing is not solely a result of attentional competition. The within-modal interference could also conceivably result from auditory masking and visual scanning. For instance, during AA trials the auditory stimuli were presented in RAP, a format which may result in early presented alphabetic letters obscuring the quality (or discriminability) of later presented letters, a process referred to as informational masking (Morrison et al., 2010). In a similar way the RSVP presentation of the visual stimuli, at two disparate presentation sites, may require visual scanning (movement of the head and/or eyes). The presence of either of these perceptual phenomena is likely to reduce target accuracy and increase target identification reaction times. Therefore, the level of interference in within-modal processing may in fact result from a combination of attentional and perceptual competition. This acknowledgement is made to encourage future researchers to discriminate between these sources of competition to overcome a criticism of the current, and certainly standard, method of experimenting in within-modal conditions. Despite this concession, in the current research substantial measures were undertaken in the preparation of the auditory and visual stimuli to limit any potential perceptual competition confound(s).

To reduce the likely impact of auditory masking in these experiments, pilot studies were first conducted to determine presentation rates which would enable perfect letter identification scores. Perfect discrimination therefore minimises the likelihood of auditory masking effects during experimentation. Meanwhile, to reduce the impact of visual scanning
the provision of a 16° separation distance between the two visual stimuli was adopted as it was considered a large enough distance so as to preserve the discreteness of the left and right hand side presentations, yet small enough so as to minimise the need to engage in eye movements to identify the presented letters (Morrison, et al., 2010).

Despite these measures being taken it is not possible to conclude that perceptual competition has been completely nullified. Nevertheless, should perceptual competition actually be present this would not diminish the value of these results. Firstly, at worst this would mean that either, or both, AA/VV performances are underestimated here. However, given the large effect sizes reported in these studies it would be unlikely that any cross-modal advantage overestimation would be to a degree so large that with the perfect control of perceptual interference the cross-modal advantage might cease to exist. Secondly, given that the results here have particular relevance in the applied field, and, that the co-occurrence of attentional and perceptual competition is a feature of all real world environments then the cross-modal advantages documented here are more likely to be underestimated as a result of the significant efforts taken to reduce perceptual competition. As such, while this auditory and visual equivalence assumption should be recognised here, it is the position of the candidate that perceptual competition has been minimised to an extent that its impact is unlikely to alter the conclusions made here.

8.3 INTERPRETING THE THESIS FINDINGS

8.3.1 Auditory and Visual Resource Modelling

The results presented here strongly support the proposition that cognitive load impacts how attentional resources are distributed during cross-modal processing, at least for the auditory and visual modalities. The continuum proposed (see Section 2.5) to exist between the pure multiple- and unitary-resource models, which is mediated by cognitive load and to a
lesser extent cognitive complexity, was supported across all studies here, studies that used methods that were developed to be model-neutral. The development of such model-neutral experimental methods was considered paramount from the outset given the thesis that the divide in the literature and support between the multiple- and unitary-resource models may have resulted from different methods which involved differing levels of imposed cognitive load.

The model-neutral methods adopted here demonstrated that by manipulating only cognitive load, the presence and magnitude of a cross-modal advantage in dual-task scenarios could be determined. In fact, the manipulations were so successful that outcomes under low levels of cognitive load were quite consistent with the purest predictions of multiple-resource models while under high levels of cognitive load outcomes were consistent with predictions of unitary-resource models. As such, the validity of model-continuum proposed as a heuristic to describe how the experimental methods of many past research reports contributed to support for contradictory models was supported.

The results here clearly depict a diminishing cross-modal advantage across the increasing levels of cognitive load. These results are consistent with the argument developed throughout this thesis to explain the seemingly inconsistent results in the literature regarding auditory and visual attentional resourcing as a function of cognitive load. However, it is of interest to note that in one study the opposite of this finding was documented – that is, as cognitive load increased so did the size of the cross-modal advantage (Wickens & Vidulich, 1983). Reasons for these conflicting results are not clear but it may be a result of different ‘code comparisons’ (see the Wickens & Vidulich Multiple Resources Cube) being undertaken between here (auditory versus visual) and in the Wickens and Vidulich study (spatial versus verbal). Clearly, a more detailed comparison between these two findings is warranted in future research to clarify this anomaly.
The Interactive Hierarchical (IHR) Model, supported in the final experiment here, provides a sophisticated way to re-evaluate the results of many past studies. It had often been considered that eventually either the multiple- or unitary-resource model would eclipse the other with past research supporting the discontinued model shown to be explained by alternative reasoning or questioned via methodological considerations. The IHR model, and indeed the model-continuum offered earlier, provide an alternative and attractive possibility – that both models produce accurate predictions, but for different levels of cognitive load. Such an approach is hardly novel with many recent researchers acknowledging how a hybrid model may just as effectively explain their results (Alais et al., 2006; Spence & Driver, 1996; Sinnet et al., 2006). For example:

> It is unlikely that attention is strictly limited by a single pool of resources, nor that the different sensory modalities have completely independent reservoirs. Rather, it appears that the attentional system maximizes its working capacity by dedicating attentional resources... depending upon the task at hand.

(Sinnett et al., 2006)

The ability of each of the multiple- and unitary-resource models to account for the full gamut of performance fluctuations in multi-task situations has been previously questioned, such that Spence and Driver (1996) proposed the separate-but-linked model of attentional resourcing. This separate-but-linked account was, until now, the best hybrid model on offer. However, at the time of this proposal they also conceded that a hybrid model consisting of modality-specific-plus-modality-independent resources could equally well account for their behavioural results. Spence and Driver reasoned that their approach was a
far less “radical” claim to the hierarchical proposal. It is argued here, in keeping with the principle of Occam’s Razor that the simplest explanation is often the best, that the IHR Model is more parsimonious than the separate-but-linked model for at least two reasons. Firstly, the IHR Model can account for research findings supporting modality-specific resources as well as others supporting modality-independent resources, for its tiers impose both forms of resource limitations, in contrast to the separate-but-linked model which does not provide predictions consistent with modality-independent limitations. Secondly, the findings here indicating that cognitive load impacts attentional processing cannot be reconciled within the separate-but-linked account – at least in its current form. For instance, in Study Four significant competition for resources between modalities was demonstrated as cognitive load increased – any such resource sharing of this kind is specifically precluded by the separate-but-linked model (Spence & Driver, 1996). The studies in this thesis allows progression beyond the resource model divide, the next challenge then is to develop an experiment that can decisively contrast the predictions of the IHR model with the separate-but-linked model.

8.3.2 Comparing Cognitive Load and Cognitive Complexity

Cognitive load and cognitive complexity were originally conceived as quite separate concepts yet this autonomy is often unclear observed in the literature. Sweller (1988) conceived cognitive load in a similar fashion to how Wickens (2007) discusses workload (although this theory is most often specifically directed toward memory as a cognitive resource) – as an experience of individuals as they engage in the activity of a task. Essentially cognitive load, or workload, has been seen to share an inverse relationship with the level of residual resources and in turn performance (i.e., ↑ load → ↓ resources → ↓ performance). In contrast Woods (1988), while still concerned with predicting performance
(akin to the concerns of Sweller and Wickens), believed such predictions could be derived from assessing the features of the task itself independent of the individual and from a somewhat similar inverse relationship between cognitive complexity and performance (i.e., \( \uparrow \text{complexity} \rightarrow \downarrow \text{performance} \)). Consequently, it is easy to conceive why the two terms are often used interchangeably.

Indeed, in Studies Two, Three and Four the relationship between the two constructs was shown to be quite strong (that is, \( \uparrow \text{complexity} \rightarrow \uparrow \text{load} \)) – with both being significant predictors of accuracy performance. However as shown in Study One, the assumption of a positive and linear relationship between cognitive complexity and cognitive load is unwarranted for while individuals were able to monitor changes in cognitive complexity these changes are not necessarily matched with changes in perceived difficulty. The importance of this divergence is made clear in Study Three for, while both were seen to predict performance accuracy, rated difficulty was a far more sensitive predictor of performance than was rated complexity. Failing to appreciate the non-linear relationship between the two has the potential to result in Type I or II errors and invalid conclusions.

For instance, in Study One had ratings of both not been taken then it may have been concluded that cognitive load does not impact resource allocation (given the absence of a cross-modal advantage across all manipulations of load). However, as both measures were taken it was evident that cognitive load had not been effectively manipulated, even though complexity had been, and therefore ratings of perceived load were included, productively, in the later studies. For this reason it is recommended that in future studies attempting to manipulate cognitive load subjective measures of both complexity and difficulty be taken so to ensure the anticipated manipulation of task complexity (imposed cognitive load) is being achieved and in turn whether this is having the desired impact on task difficulty (perceived cognitive load).
8.4 Future Research Directions

A model such as the IHR Model provides a solid foundation on which to base investigations of auditory and visual attention and the capabilities and limitations of the human attentional system. Support for both the multiple- and unitary-resource models has lead to a theoretical stalemate which has impacted many applied fields (e.g., workplace design, human-computer interfaces). On the other hand, the IHR Model has the potential to generate clear definitions of task complexity and perceived load and clearer performance predictions. However, further evidence supporting this model is first required.

To interpret all past research reports in terms of the IHR Model might be a desirable approach, however the feasibility of such a review is problematic. While in this thesis a general overview of the literature led to the proposed model-continuum and in turn the proposal of the IHR Model, it is unlikely that a categorical classification of all the literature could be carried out accurately – for one thing measures of task complexity and task difficulty were not included in all past studies and without these absolute classification is not possible. Moreover, as is evident in Chapter 2 the classification of research studies into those that support one model or another is not clear cut, as some features (e.g., task stimuli) may be considered ‘non-complex’ while other features (e.g., required response) may be considered ‘complex’. Determining how these individually contribute to the overall complexity of the task could be not much more than guesswork without further research.

What should be clear from past research is that systematic testing of models of attention is required to increase the probability that alternative explanations cannot provide better explanations of the results. For instance, Spence and Driver (1996) offered the IHR model as an alternative explanation to their separate-but-linked resource model. In this series of experiments the concerns of Spence and Driver (e.g., demonstrating the existence of a modality-independent resource) were addressed before a structure for such a hierarchical
model was considered. However, the full structure of the IHR model has not been explicated here. For instance, there is no consideration of how auditory and visual stimuli arising from a common source (that is, an audio-visual event) might be resourced (Burr & Alais, 2006), nor of what altering the spatial or temporal relations of a dual-task stimulus might have an impact upon (Spence, 2001). In the studies presented here the stimuli have been presented in reasonably symmetrical pattern and therefore, in some sense, a pattern of presentation and response may be predicted by the participant (particularly as their exposure to the task increases). In the real world stimuli are rarely encountered in such predictive patterns and as such the results here may not easily generalise to real world environments. For instance, Wickens and McCarley (2007) report that in auditory pre-emption, where the presentation of one stimulus actually interrupts the presentation of another stimulus, results in differential performance dependent upon the auditory or visual nature of the interrupting task. Essentially this means that the size of the cross-modal advantage may be subject to not only the nature of the tasks (auditory and visual) – but also to their temporal relationship (discrete or interrupting) (see Wickens & Colcombe, 2007). As such research studies examining spatial and temporal parameters may provide a method for contrasting the separate-but-linked and the IHR Models for the separate-but-linked model has strict performance predictions for when stimuli are co-located, while the IHR Model provides no differential hypotheses for when stimuli are co-located compared to when they are not. At a more global level, in this series of experiments no concern has been paid to how the auditory and visual attentional system might fit within a grander scheme involving tactile, olfactory and gustatory processing. Addressing these aspects will further test the generalisability and external validity of the model.

To further investigate the IHR Model it is therefore most appropriate to conduct both additional laboratory-based experiments and also in invivo experiments, for as any applied
researcher would agree, testing such a model in a real world environment is essential to establishing its applied utility. Reconsidering the task used in this study (target letter identification from multiple RSVP- and RAP-presented distractor letters) it would be difficult to suggest that cognitive load manipulations cover the entire spectrum of difficulty that an individual is likely to encounter on a day-to-day basis. Essentially, it could be argued that the difficulty levels imposed in this series of experiments at best represent the lower end of the day-to-day difficulty spectrum, when compared for instance with air traffic control, event control, emergency service provision. Therefore, examining the IHR Model within more realistic settings with real-world stimuli – for example using a train control workplace environment with monitors displaying train movement and auditory alerts and instructions being delivered and responded to – could provide an effective method for evaluating the generalisability of the model.

8.5 IMPLICATIONS

Over and above the implications that the IHR Model may have for cognitive research there are many practical applications. The real world utility could be considered as, if not more, important than the theoretical contributions given that much workplace design, human-computer interfaces, government policy (e.g., safe driving practices) arise from the understanding of multi-tasking limitations. For instance, in Australia, the Roads and Traffic Authority (NSW) have shown 3 out of every 10 road incidents to be a direct consequence of mobile phone use (McEvoy, Stevenson & Woodward, 2006). Moreover, it has been shown that the use of a mobile phone while driving is equivalent to driving with a blood alcohol level of 0.05 (Strayer, Drews & Crouch, 2006). For these reasons in many Australian states, including New South Wales, only hands-free mobile use is permitted by statutory law. However, no limitations are placed over in-car driver-passenger communication which
essentially involves the same (auditory) cognitive processes as hand-free mobile communication, and arguably is more attention grabbing since in-car interaction presumably also often involves visual communicative elements. The research here suggests that hands-free mobile while driving should result in few driving performance breakdowns as long as the conversation or the driving are maintained at low cognitive load inducing levels, while in-car conversation may well result in greater resource expenditure due to processing the visual communicative signals being resourced by the same reservoir as the driving task itself. Would it therefore be wise to legislate regarding the number of passengers (that is, potential distractors) allowed, the types of conversations that may be engaged in, or even the ability to have conversations at all? Perhaps some level of training may improve driving performance in such conditions?

The results here (in Study 3) demonstrate that training (and motivation) can immunise the operator against the cross-modal advantage decrement that occurs under high load conditions (in Studies 2 and 4). In this regard it is of interest that in New South Wales legislation now prohibits provisional (P-plate) drivers from having more than one passenger in their car between the hours of 11pm and 5am. This law was instituted in response to the number of P-plate driving accidents resulting in death on the NSW roads, and was supported by research suggesting that more passengers increased ‘risky behaviour’ by the driver. However, a number of fatal P-plate accidents have occurred outside of these hours in the absence of ‘at risk’ behaviours with multiple passengers being present (Styles, Imberger & Catchpole, 2005). Could the answer be that, in addition to the risky behaviours encouraged by the presence of multiple passengers, attentional limits may also be reached by having so many sources of distractions especially when in the case of P-plate drivers individuals have to automatise their driving abilities? The fact that age is negatively correlated with road accident incidence may support this claim – for in general increasing age is related not only to
decreasing risk taking behaviour but also to increased driving experience. If this is the case, and certainly future research should be concerned with investigating the contribution of poorer attention management in P-plate drivers, then government policies should be directed to increasing the experience of young drivers in controlled (low attention requiring) in-car environments. Perhaps driving tests could be designed to include evaluation and training of attention management abilities, the imposition of zero-passenger limitations until certain attention management abilities are achieved.

Applications for these research outcomes can also be seen in the development of advanced technology human-computer interfaces. For instance, operational display systems on both flight control decks and in cockpits have undergone significant changes in the past decade with many designers favouring a system which divides information across various computer screens with additional information between cockpits and control decks being exchanged auditorily. A recent development in cockpit-control communication, the Controller Pilot Data Link System, has seen text-based messages exchanged between the two essentially making the communication less temporally based\(^1\). However, the present research would suggest that such a system may actually cause interference during conversations and attention to the other screens displaying plane trajectories, flight pattern changes, plane altitudes etc. As such, perhaps a system which still maintains oral/auditory communication but which then is transcribed into text-based visual information for later reflection if necessary, might be more advisable.

Of importance here is that cross-modal presentation was not found to be worse than within-modal processing at any time across these studies. As such, cross-modal presentation of information is probably most beneficial in the majority of situations. Therefore, many

\(^1\) Oral and auditory stimuli are generally considered to have a fine temporal structure (Rosen, 1992). That is, once the sound or word is spoken it can then only exist in memory – there is no external representation of it available to be re-considered or processed at a later time. This is not necessarily the case for visual stimuli which can transcend time.
workplaces, e.g., security operations, medical theatres, nuclear power plant operations, air traffic and piloting, rail control, high voltage electricity transmission station control could benefit by designing display systems which present information multi-modally. Moreover, training to facilitate response automatisation could enhance the cross-modal performance benefits gained. Additionally, with the understanding that cognitive load will impact cross-modal performance differently to within-modal processing, it may be possible to assess current workplaces and workstations to evaluate under what circumstances performance decrements are likely to occur and then to develop procedures which will come into effect when any such situations arise to obviate, or at least minimise, performance decrements.

8.6 Conclusions

The human attentional system is a multi-tiered system incorporating both modality-specific and modality-independent resources. As such neither the classic multiple-resource nor unitary-resource models adequately describe attentional processing. Rather a hierarchical model, such as the IHR model proposed here, better represents the resource limitations of the cognitive system. Such a hierarchical model can be used to demonstrate why the level of cognitive load experienced by an individual will moderate whether modality-specific or modality-specific-plus-modality-independent resources are required to maintain performance.

Predictions of performance capabilities within a hierarchical structure require assessment of the contributions of extraneous, intrinsic and germane cognitive load. The findings here suggest that extraneous cognitive load can be minimised by presenting information cross-modally, while germane cognitive load can be minimised in part by increased training and/or motivation. Intrinsic cognitive load cannot be manipulated, as it naturally evolves as a function of the task requirements. Estimating the contributions of each, extraneous, intrinsic and germane load, and therefore performance capabilities requires
evaluation of the task features and the operator-task relationship using both objective and subjective ratings measures. Objective measures, like task complexity, are certainly more straightforward and less time consuming to conduct, however accuracy in predicting individual differences in performance is compromised leading, at best, to a one-size-fits-all outcome. In contrast, subjective measures, like task difficulty, require significantly greater investment of time in experiments or in applied settings but lead to far superior performance predictions. It may well be that, depending upon the nature of the environment, the task, and the operators, both objective and subjective measures are worthwhile measures in future studies.

The aim of this thesis was to devise model-neutral methods and disambiguate the resource-model divide. The model-continuum proposed as a result of the literature review of this divide was well supported in a series of four experiments and does indeed provide a vehicle by which to clarify and reevaluate the divide. Moreover, the IHR model which arose from the results of the empirical studies derived from the model-continuum provides a level of description of auditory and visual attention which, until now, has been absent in hierarchical model proposals. Further testing of this model is desirable.
REFERENCES


<table>
<thead>
<tr>
<th>Alphabetic Letter</th>
<th>English (Australian) Pronunciation*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a</td>
</tr>
<tr>
<td>B</td>
<td>bee</td>
</tr>
<tr>
<td>C</td>
<td>cee</td>
</tr>
<tr>
<td>D</td>
<td>dee</td>
</tr>
<tr>
<td>E</td>
<td>e</td>
</tr>
<tr>
<td>F</td>
<td>ef</td>
</tr>
<tr>
<td>G</td>
<td>gee</td>
</tr>
<tr>
<td>H</td>
<td>aitch</td>
</tr>
<tr>
<td>I</td>
<td>i</td>
</tr>
<tr>
<td>J</td>
<td>jay</td>
</tr>
<tr>
<td>K</td>
<td>kay</td>
</tr>
<tr>
<td>L</td>
<td>el</td>
</tr>
<tr>
<td>M</td>
<td>em</td>
</tr>
<tr>
<td>N</td>
<td>en</td>
</tr>
<tr>
<td>O</td>
<td>o</td>
</tr>
<tr>
<td>P</td>
<td>pee</td>
</tr>
<tr>
<td>Q</td>
<td>cue</td>
</tr>
<tr>
<td>R</td>
<td>ar</td>
</tr>
<tr>
<td>S</td>
<td>ess</td>
</tr>
<tr>
<td>T</td>
<td>tee</td>
</tr>
<tr>
<td>U</td>
<td>u</td>
</tr>
<tr>
<td>V</td>
<td>vee</td>
</tr>
<tr>
<td>W</td>
<td>double-u</td>
</tr>
<tr>
<td>X</td>
<td>ex</td>
</tr>
<tr>
<td>Y</td>
<td>wy</td>
</tr>
<tr>
<td>Z</td>
<td>zed</td>
</tr>
</tbody>
</table>

* taken from http://www.omniglot.com/writing/english.htm
Appendix B

- Part Two -

Ratings Questionnaire

Participant Number
Trial Number: 1

How difficult did you feel this trial was?

1  2  3  4  5
It was easy  It was hard

How confident are you in your performance on this trial?

1  2  3  4  5
Not confident  Very confident

How many distractor letters were in this trial?

1  2  3  4  5
Few  Lots
**Trial Number: 2**

<table>
<thead>
<tr>
<th>How difficult did you feel this trial was?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>It was easy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How confident are you in your performance on this trial?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Not confident</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How many distractor letters were in this trial?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>Few</td>
</tr>
</tbody>
</table>
Trial Number: 

How difficult did you feel this trial was?

1  2  3  4  5
It was easy  It was hard

How confident are you in your performance on this trial?

1  2  3  4  5
Not confident  Very confident

How many distractor letters were in this trial?

1  2  3  4  5
Few  Lots
Appendix C

Pilot Test #1: Alphabetic Letter Presentation Duration

Pilot Test Overview

In the literature there appeared to be no standard presentation duration for alphabetic letters used in the target identification dual-task paradigms. For instance, in one study (Arnell & Jolicouer, 1999) the auditory letters was compressed\(^1\) to a duration of 90ms, however, to the candidate many letters when compressed to 90ms appeared distorted which was of concern as a potential confound. Similarly, compressions of 80ms (Arnell & Larson, 2002) and 56ms (Soto-Faraco, et al., 2005) were considered inappropriate. Longer presentation times of between 100 and 150ms (Arnell & Jolicoeur, 2002; Jolicoeur, 1999; Potter, et al., 1998) were considered more appropriate. As such, this pilot test was used to determine the appropriate duration of the auditorily presented letters to be incorporated into the rapid auditory presentation (RAP) sequences.

In this pilot test participants were presented with seven sets of the 26 alphabetic letters, with each set varying in the duration of the individual letters. The durations ranged from 56ms (the minimum length of those identified in the literature) through to 300ms which was the average duration of letters recorded prior to any compression. The following durations were examined: 56ms; 80ms; 115ms; 150ms; 175ms; 190ms and 300ms.

Method

Participants

In this pilot study ten participants (6 females, and 4 males, \(M\) age of 26.1 years) took part. All reported normal hearing. No incentives were awarded for participation.

\(^1\) Compression is used to ensure all stimuli are presented to the participants for equal durations which removes any potential confounds of changing presentation duration.
**Apparatus**

The experiment was conducted on an LG (P1 Express Dual) Laptop with a 17” widescreen display. The stimuli were presented via the Microsoft PowerPoint program which allowed participants to click on a sound icon, and then provide a name for the letter on top of the sound icon. This program was selected due to participant familiarity with it. The auditory letter stimuli were presented monaurally via headphones (Senneheiser 650), attached to the laptop headphone jack.

**Stimuli**

The 26 alphabetic letters were recorded using the TextAloud (v.2.0) text-to-speech program available for free download. Without compression nearly all letters ranged in duration between 290ms and 312ms – with the exception of the letter “W” which was 320ms in length.

These 26 letters were then compressed to form seven sets representing the seven duration lengths of interest.

**Procedure**

Participants were seated in front on the laptop computer at a comfortable distance (determined by them) to use the keyboard and mouse and instructed put on the headphones before starting the experiment. To familiarise participants with the procedure they were stepped through the screen layout and mouse click procedures on a practice screen where the auditory stimuli were three letter words (to avoid any practice effects occurring). When the participants were comfortable with the procedure the experimental screens were loaded.

Each set was presented separately from the others – with set exposure order being randomised across participants. Within each set each of the 26 letters were represented by
loud speaker icons which participants could click on to hear each sound. Participants could press on each icon only once (with the exception of the first icon engaged which they could press twice to allow some familiarisation of the presentation duration) before they were required to name the sound by clicking on the hash and using the laptop keyboard to designate the letter they believe matched the sound. Participants were instructed that there were 26 icons and that each of the 26 letters of the alphabet were represented. See Figure 1 for a screen cap of the computer screen as seen by the participants.

![Screen cap of each set of 26 alphabetic letters. Participants double clicked on the sound icon to hear the auditory stimulus, and then clicked on the # to type their identification of the sound.](image)

*Figure 1.* Screen cap of each set of 26 alphabetic letters. Participants double clicked on the sound icon to hear the auditory stimulus, and then clicked on the # to type their identification of the sound.

When each of the 26 sounds had been named by participants they were instructed to click on the [proceed to next page] button. The same screen of sound icons and #’s would
appear for the next duration set. The sound icon screen distribution placement was varied for each sound duration set.

This process was not timed and participants could take short breaks between the processing each set if desired.

Results

Participant accuracy on letter identification is shown in Table 1. It can be seen that while some participants were able to accurately identify all letters at some of the shortest compression durations (that is, 80ms), a minimum of 150ms were required for all participants to record 100% accuracy.

Table 1.

*Minimum Presentation Duration required for 100% Accuracy for each Participant*

<table>
<thead>
<tr>
<th>Participant Number</th>
<th>Minimum Presentation Length for 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150</td>
</tr>
<tr>
<td>2</td>
<td>115</td>
</tr>
<tr>
<td>3</td>
<td>115</td>
</tr>
<tr>
<td>4</td>
<td>150</td>
</tr>
<tr>
<td>5</td>
<td>150</td>
</tr>
<tr>
<td>6</td>
<td>115</td>
</tr>
<tr>
<td>7</td>
<td>115</td>
</tr>
<tr>
<td>8</td>
<td>150</td>
</tr>
<tr>
<td>9</td>
<td>150</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
</tr>
</tbody>
</table>

Feedback from participants did suggest that although they were able to easily identify the letter “W” from the other letters, its presentation seems rather distorted for anything shorter than the 190ms duration.
Discussion

As performance in the thesis studies is expected to reflect attentional capacities it was felt that the stimuli themselves should not impede the target identification dual-task experimental design being used in the main experiments. For this reason setting a high standard of 100% accuracy across all participants in this pilot study for letter labelling was considered appropriate.

As such, after the recording of “W” was removed – given the apparent distortions associated with it, presumably due to the considerable compression required due to its extended original duration, the 25 alphabetic letters of 150ms were determined the most appropriate to use in the subsequent thesis experiments.
Appendix D

CD of Auditory Recording of 25 Letters Compressed to 150ms
Appendix E

**Pilot Test #2: Target-Distractor Similarity Ratings**

**Pilot Test Overview**

Cognitive load can be manipulated via three avenues; changing how the task is presented (extraneous load); changing the complexity of the task (intrinsic load); and, changing the skills and/or knowledge of the operator engaging in the task. The first of these is evidently at the heart of the entire thesis, examining whether cross-modal presentation facilitates better performance than within-modal presentation as a first step it was considered most straightforward to manipulate intrinsic cognitive load by selecting a single task in which its integrated features could be altered to make the task easier or harder. As identified, the dual-task target identification paradigm was selected wherein participants were required to view streams of auditorially and visually presented letters (in RAP and RSVP letter sequences) and to identifying the pre-selected target letter from the milieu of distractor letters in the sequences. As such, it was considered appropriate to manipulate cognitive load by making the identification of this target more or less difficult.

One way to make a target stimulus more difficult to identify is to mask it amongst the distractors letters (Gelfand, 2004). From an auditory standpoint this involves making the surrounding stimuli more similar to it to promote informational masking - which is when item identification is obstructed because something similar just prior to its presentation essentially confuses perception or comprehension (Brungart, 2006). A similar form of masking is seen in visual processing, so that if similar stimuli are presented in close temporal succession often the second of these is cognitively overlooked, as in repetition blindness – for example in the sentence “when she spilled the ink ink was all over” may people will not perceive there to be a double presentation of “ink”, or in some cases will read the second as “it” (Kanwisher, 1987). Therefore, common to both the auditory and visual modalities is that when a target is
preceded by a stimulus sharing similar auditory or visual features the target is more likely to be missed. Subsequently, it was proposed that the more similar the surrounding distractor letters were to the target the more difficult differentiation of that target would be, and consequently, the most dissimilar the distractors are the easily target identification should be. Certainly the relationship between target and distractor stimuli have been previously examined and found to impact target comprehensibility (Blough, 1985).

The purpose of this pilot study was to identify the appropriate target letter to use in the subsequent experiments; a letter which is highly similar to some alphabetic letters and highly dissimilar to others.

Method

Participants

In this pilot study fifteen participants (9 females, and 6 males, $M$ age of 27.6 years) took part. All reported normal or corrected hearing and sight. No incentives were awarded for participation.

Apparatus and Stimuli

The experiment was conducted on two LG (P1 Express Dual) Laptops with a 17” widescreen displays. The stimuli on both were presented via the Microsoft PowerPoint program which allowed participants click on sound icons to hear the auditory stimuli and view the visual stimuli on separate screens. The auditory stimuli were those selected at the conclusion of Pilot Study 1 (all 25 letters of the alphabet, excluding the letter “W”) and were presented via headphones (Sennheiser 650), attached to the laptop headphone jack of Computer 1. The visual stimuli included all 26 alphabetic letters which were presented in 24 point black Times New Roman font on an off-white background screen on Computer 2. The
two computers were set up side by side to allow participants to easily switch between auditory and visual tasks as often as necessary.

Procedure

Participants were seated in front on the laptop computers at a comfortable distance (determined by them) to use the keyboard and mouse and instructed put on the headphones before starting the experiment. To familiarise participants with the procedure they were stepped through the auditory and visual screen layouts on two practice screens where the visual and auditory stimuli were three letter words (to avoid any practice effects occurring). When the participants were comfortable with the procedure the experimental screens were loaded.

Participants were required to firstly look at all the visual stimuli and then listen to all of the auditory stimuli to familiarise themselves with all the possible individual letters of interest. They were then able to click on any of the sound icons (on Computer 1) to rehear the individual letter stimuli at their leisure. They were able to look at the visual and listen to the auditory stimuli as often as they wished across the entire experiment.

Participants were then instructed that they needed to select a target letter which was the same for both the auditory and visual stimuli – and that could be rated as highly similar or highly dissimilar to other alphabetic letters in the same modality. Having the same target in both modalities was important to avoid confusion in the thesis experiments for participants. However allowing for different distractors to be associated with this target in different modalities was important for while visually “D” and “O” might be similar, auditorilly they are quite dissimilar (“dee” and “oh”).

Once a target letter was identified by the participant they were asked to drag it to the central box on both Computer 1 (auditory) and Computer 2 (visual) – See Figures 1 and 2 for
the screen caps of the auditory and visual displays. They were then asked to select for each target a set of five highly similar and five highly dissimilar letters (which would form the distractors in latter experiments) by dragging each individual auditory or visual icon into the ratings boxes on either side of the identified target.

Figure 1. Screen cap of the auditory display for identifying the target letter and rating the most highly similar and dissimilar letters to the target within the auditory modality.

Figure 2. Screen cap of the visual display for identifying the target letter and rating the most highly similar and dissimilar letters to the target within the visual modality.
Participants were instructed they could revise their target letter if they decided during this rating process it was not be best choice of target, and they were allowed to switch between the auditory and visual rating processes as often as required. Once the targets and distractors were selected participants were asked to verify that they were happy with the overall selection before the data were saved.

Results

Target Selection. As can be seen in Table 1 despite there being no collaboration between participants the letter “B” was selected as the target letter of choice by 7 out of 15 participants (47%), of the five other targets selected by participants no others were selected by more than two participants.

Table 1. 
Target Letter Selection and Participant Frequency in Selection of the Target Letters

<table>
<thead>
<tr>
<th>Target Letter</th>
<th>Frequency</th>
<th>Percentage of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>7/15</td>
<td>46.7%</td>
</tr>
<tr>
<td>A</td>
<td>2/15</td>
<td>13.3%</td>
</tr>
<tr>
<td>G</td>
<td>2/15</td>
<td>13.3%</td>
</tr>
<tr>
<td>C</td>
<td>1/15</td>
<td>6.7%</td>
</tr>
<tr>
<td>T</td>
<td>1/15</td>
<td>6.7%</td>
</tr>
<tr>
<td>H</td>
<td>1/15</td>
<td>6.7%</td>
</tr>
<tr>
<td>P</td>
<td>1/15</td>
<td>6.7%</td>
</tr>
</tbody>
</table>
Distractor Selection. Data were retained in this analysis for those participants who selected “B” as their target letter. The results indicated a high level of agreement between participants for those letters rated as being highly similar to the target letter in both the auditory and visual modalities. For the auditory stimuli a total of six letters were rated as highly similar to the target across all seven participants, while for the visual modality seven letters were rated as highly similar to the target. Those with the highest frequency ratings were “V, E, D, C and G” for the auditory stimuli and “F, E, D, P and R” for the visual stimuli (see Table 2 for a full description of the similar letters identified, and the frequency of this identification).

The level of agreement across both auditory and visual stimuli for dissimilarity ratings was somewhat less than for similarity ratings, with nine letters being rated as highly dissimilar to the auditory “B” and eight letters as being highly dissimilar to the visual “B”. Those with the highest frequencies were: auditory – “U, X, Z, I and L”; visual – “Y, J, T, N and M” (see Table 2 for a full description of the dissimilar letters identified, and the frequency of this identification).

Discussion

The results clearly identified that “B” was a firm favourite of participants for the role of a target letter given the ability to identify highly similar and highly dissimilar letter counterparts in both the auditory and visual modalities. In terms of the similar and dissimilar letters, those with the five highest frequency ratings in each (dis)similarity group, for each modality, are those of interest for the following experiments as they can be used to manipulate the level of target masking and in turn the complexity of the task being engaged by participants.
<table>
<thead>
<tr>
<th>Classification</th>
<th>Letter</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Highly Similar Ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Auditory</strong></td>
<td>V</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>6/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>6/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>5/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>1/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>1/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>U</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>5/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Z</td>
<td>5/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>4/7</td>
<td></td>
</tr>
<tr>
<td><strong>Visual</strong></td>
<td>L</td>
<td>4/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>3/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>3/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>3/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q</td>
<td>1/7</td>
<td></td>
</tr>
<tr>
<td><strong>Highly Dissimilar Ratings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Auditory</strong></td>
<td>Y</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>6/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>5/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>4/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>3/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>2/7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>1/7</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix F

### Studies One and Two: Table of Planned Contrasts

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Contrast Type</th>
<th>Condition Contrasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Condition 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
</tr>
<tr>
<td><strong>Modality Combination</strong></td>
<td>Linear (Within vs. Cross)</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td>Quadratic (Within comparison)</td>
<td>-1</td>
</tr>
<tr>
<td><strong>Cognitive Load</strong></td>
<td>Linear</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>0</td>
</tr>
</tbody>
</table>
Putting the ‘party’ back into the Cocktail Party:
Localising conversant voices improves comprehension in telecommunications

Natalie M. V. Hughes  
MARCS Auditory Laboratories and the School of Psychology  
University of Western Sydney

Denis Burnham  
MARCS Auditory Laboratories  
University of Western Sydney

Simon Carlile  
Bosch Institute  
University of Sydney

Abstract

Normally hearing listeners are adept at segregating and understanding speech in the presence of concurrent sounds. Much research on this “cocktail party problem” has focussed on speech detection and intelligibility using single words or closed sentences under tightly controlled experimental conditions (see Bronkhorst 2000 for review). The perception of differences in the locations of targets and maskers plays a key role in this process. In the experiments reported here, the role of spatialisation on the comprehension and use of the information in a three party conversation is examined. In Experiment 1 the ability of groups of three subjects conversing using headsets to solve a problem is measured in terms of the accuracy of the solution and the time to completion. Both measures demonstrated significantly improved performance when subjects heard a separate talker in each ear compared to when the two talkers were presented diotically. In a second experiment subjects listened to a recording of the three-party conversation over a single loudspeaker (single-channel) or three spatially separated loudspeakers (multiple-channel) and answered a questionnaire aimed at probing their comprehension of content-general and conversant-specific information. In the spatially separated condition, comprehension scores were higher for content-general information compared to the single channel condition. Together these results indicate that multiple-channel presentation improves the comprehension and use of speech information in audio-only displays when compared to single channel presentation.
1. **INTRODUCTION**

Everyday life presents a particularly complex acoustic environment (Alain & Izenberg, 2003), yet humans are also particularly adept at listening to and comprehending selected auditory objects and events, including single talkers in multiple speech streams (Bec & Micheyl, 2008; Kidd, Arbogast, Mason, & Gallun, 2005a). When there are a large number of talkers (Arons, 1992), or in intensely noisy environments (Bronkhorst, 2000) the ability to distinguish one speech signal from another becomes more difficult, but nevertheless normally hearing individuals accomplish this, often in a seemingly effortless manner (Bec & Micheyl, 2008).

When speech signals are concurrently presented there are periods of overlap between the various signals, which act to reduce detection and intelligibility of the target signal (Bec & Micheyl, 2008; Hawley, Litovsky, & Culling, 2004). Energetic masking results when the target speech signal is *acoustically* obscured by a competing sound that occupies the same time and frequency domains (Brungart & Simpson, 2002). On the other hand, informational masking results from similarities between the target and distracter streams and disrupts the association of different words with specific talkers (e.g., similar sounding voices or discussion of the same subject matter) (Arbogast, Mason, & Kidd, 2005). Differences in characteristics of the talkers can be used to aid segregation as can differences in their spatial location (Hawley et al 1999). Localising the origins of differences in the perceived locations of multiple sound sources has been found to increase intelligibility in many experimental situations, including selective listening environments (see Bronkhorst, 2000), speech shadowing contexts (Hawley, et al., 1999), group interactions (Yost, et al., 1996) and telecommunications settings (Kang & Kim, 1996) by reducing the effects of energetic and/or informational masking (Kidd, Mason, & Gallun, 2005b). Differences in spatial location can reduce energetic masking through “better ear” listening resulting from a
decrease in the level of the masker talker produced by ear shadowing. Perceived differences in spatial location aid in informational unmasking by mediating selective spatial attention.

Many psychoacoustic studies of masking use intelligibility as a measure of segregation ability; participants are asked to shadow (repeat back) target speech from among many distracter speech signals (Cherry, 1953; Cowie, Douglas-Cowie, & Kerr, 1982; Stifelman, 1994). While, intelligibility is a good measure of signal detection in such circumstances, it is necessarily indicative of comprehension and utility of the information and it may not be particularly relevant in more complex settings in which comprehension is the measure of success, e.g., team decision making settings (Hinsz, Tindale, & Vollrath, 1997).

Accordingly, conversational and interactive speech was used in this study rather than multiple unrelated speech utterances (e.g., alphabetic letters, numerals, call signs) as used in many previous studies (Brungart, et al., 2001; Hawley, et al., 2004; Yost, et al., 1996). The effect of transmission of two or three voices via a single audio speaker (single-channel), was compared to talkers at different locations (multiple-channel), and was measured in two ways in two experiments: (1) a group problem solving situation in which accuracy and time to completion were the dependent variables, and (2) a study in which individuals’ understanding of the group problem solving discussions in Experiment 1 was measured by comprehension questions.

The specific aims of these experiments were to examine the impact of differences in the locations of different talkers in (1) the use of information in an interactive auditory task and (2) the comprehension of the speech information in a passive listening task.
2. EXPERIMENT 1: Information transfer in single- and multiple-channel group teleconferencing

The purpose of Experiment 1 was to evaluate the value of differences in the perceived location of talkers, in multi-member teams’ completion of “who-dunnit” tasks in a teleconference context as measured by accuracy and time to completion. On the basis of findings that localisation increases intelligibility (see Bronkhorst, 2000), and that intelligibility corresponds with comprehension (Schwartz, Berthommier, & Savariaux, 2004), it was predicted that accuracy would be higher and time to completion lower in multiple channel than in single channel transmission of the team members’ voices.

Teams of three conversants completed two problem solving tasks in audio-only teleconferences, one in a single channel system where all the talkers were presented diotically over headphones and the other, a multiple channel system where two talkers were presented in monotonically, one to each ear. In addition to accuracy and time to completion, participants were asked, on completion of both tasks, to compare the two conditions and indicate a preference for one and provide the reason for their preference.

2.1 Method

2.1.1 Design

A within-subjects design was used, with each team participating in both a single- and a multiple-channel teleconference. Three dependent variables were used, accuracy, time to completion (toc), and participants’ preferences.

2.1.2 Participants

Twenty one male undergraduate drama students (18 - 29 years of age ($M=19.32$, $SD=2.13$) were assigned to seven teams of three conversants. From an
original sample of twenty seven, any participants who exhibited strong individual voice characteristics (e.g., accents, lisps), were excluded from the study so as to provide a uniform set of voice characteristics across the participant groups. All participants gave their informed voluntary consent, and were reimbursed for their time and travel expenses.

2.1.3 Apparatus

For any one team, each of the three participants was seated in one of three acoustically isolated rooms (1.3m x .9m x 1.9m) with a set of headphones (Sennheiser HD 650) and a microphone headset (DPA 4060). Microphone inputs were fed through an amplifier (RME OctaMic) to a desktop computer (IMS Intel Pentium 4) running Hammerfeld DSP Mixing Software (2002, version 2.432). This arrangement enabled each participant to hear the voices of their team members over headphones in real-time.

In the single-channel condition participants heard the voices of the other two team members through both the left and the right headphones (diotic display) (see Figure 1a), i.e., the same combined message was played to each ear. In the multiple-channel condition the voice of one team member was played through the left headphone, and the voice of the second team member through the right headphone (dichotic display) (see Figure 1b), i.e., discrete messages were played to each ear. Each of the three participants’ voice outputs were captured through the Hammerfeld DSP program and recorded as individual temporally-synchronised sound files using Cool Edit Pro (2003, version 2.1). These sound file recordings were retained as stimulus materials for Experiment 2.
2.1.4 Stimulus Materials and Dependent Variables

Each team of participants completed two tasks, one under single-channel conditions, and the other under multiple-channel conditions. One task was the ‘Solstice Shenanigan’s Mystery’ (SSM) (Johnson & Johnson, 2000). Thirty-three ‘clue cards’ were evenly distributed between the three team members. Accurate task completion required team members to communicate their clues, and to solve the mystery. The other task, the Winter Wonderland Mystery (WWM), was created by modifying suspect names, motives, objects and the context of the Solstice Shenanigan’s Mystery. Completion of each task required the teams to answer four questions regarding the crime they were solving: (1) who was the thief, (2) what was stolen, (3) when was it stolen, and; (4) how it was removed from the premises. Pilot testing of the two tasks with an additional four three-participant groups, conversing in face-to-face interactions showed that mean completion time ($M_{SSM}=1208$ secs, $SD_{SSM}=37$ secs; $M_{WWM}=1220$ secs, $SD_{WWM}=36$ secs) and accuracy (100% in both SSM and WWM) were comparable between the tasks. Both task order (SSM and WWM) and presentation method (single- and multiple-channel) were counter-balanced between teams.

Following each task, participants completed two supplementary questions; a 2-alternative forced-choice question asking them to select which display mode (single- or multiple-channel) they had preferred to interact within, and an open-ended question asking participants to explain why they preferred that particular display mode.

2.1.5 Procedure

For each session, the three team members were placed in separate rooms, fitted with a microphone headset (positioned at the recommended distance of 1.0cm from the mouth) and headphones, provided with an instruction sheet, and asked to
begin their team discussion at will. The first task was completed once the team provided answers to each of the four task questions. Following a 10-minute break, the second task began under the same experimental conditions (same team members, same rooms), except that the other display mode to that used for the first task (single- or multiple-channel) was used.

Upon completion of the two tasks, participants were asked about their display mode preference. Participants were then debriefed about the purpose and expected outcomes of the experiment. Total participation took approximately 45mins. Repeated measures t-tests were used to compare team accuracy and time-to-completion data under the single- and multiple-channel modes. A thematic analysis was performed to evaluate the participant preference data.

2.2 Results

2.2.1 Team accuracy and time-to-completion performance

Multiple-channel (MC) interaction resulted in better accuracy than did the single-channel (SC) audio-display mode ($M_{MC}^{MC}=96.43\%, \ SE_{MC}^{MC}=3.57; \ M_{SC}^{SC}=78.57\%, \ SE_{SC}^{SC}=3.57; \ t(6)=3.87, \ p=.008$) (see Figure 2). Additionally, teams completed their tasks more quickly under multiple- than single-channel audio display modes ($M_{MC}^{MC}=1270secs, \ SE_{MC}^{MC}=16.76; \ M_{SC}^{SC}=1561secs, \ SE_{SC}^{SC}=28.07; \ t(6)=-7.29, \ p>.001$) (see Figure 3).

*** Figure 2 about here ***

*** Figure 3 about here ***
Overall, it can be seen that in conversationally-rich multiple-conversant teleconference interactions a multiple-channel display mode facilitates both more accurate and more timely problem solving that does the single-channel display model counterpart.

2.2.2 Team member preferences for and views on interaction modes

Nineteen of the 21 participants (90.5%) preferred the multiple-channel audio-display mode to the single-channel mode. Thematic analyses (see Aronson, 1994) of participant responses to the open-ended preference question revealed two main themes in participants’ preferences; ‘Voice Separation’ and ‘Realism’. (1) The Voice Separation theme was constructed to represent responses by participants with reference to the separation of the talkers into each ear. Of those participants who preferred the multiple-choice display mode, 18 made comments which contributed to this theme. Comments indicated that the multiple-channel mode allowed participants to subjectively dedicate an ear to each individual voice input and to reduce the need to engage attentional resources, whereas in the single-channel mode a high degree of concentration and a large investment of attention were required to follow the gist of the conversation. For example, “I identified each person by the location of their voice, and then it was easier to separate each person’s thoughts in my head”. The single participant who preferred the multiple-channel mode but did not make reference to this theme provided comments contributing to the second theme identified here. (2) The Realism theme was constructed to characterise those participant comments relating to the realistic nature of interaction facilitated under the multiple-channel mode, and the unrealistic nature of the interaction under the single-channel mode. Of those participants who preferred the multiple-channel display mode all 19 made reference to this theme. Participant comments included “[the multiple-channel
presentation] was just like an everyday real-life conversation… it felt more natural”.

For the two participants indicating preference for the single-channel display mode, one participant’s reasoning was that it felt like “how you normally chat on the telephone”, the other participant reasoned that it “was easier not to have to alter where I was paying attention to”.

2.3 Discussion

The results are consistent with the prediction that information transmission would be enhanced when teams interacted under a multiple-channel audio-display mode than a single-channel audio-display mode. Under the assumption that information must be exchanged between team members in order for teams to arrive at accurate and timely solutions, then it can be concluded that the multiple-channel audio-display mode allows more accurate and more timely information exchange than the single-channel mode.

These results are consistent with previous research (Arbogast, et al., 2005; Brungart, et al., 2001; Cherry, 1953; Yost, et al., 1996) and extend these findings to situations in which participants interact in a near real-world teleconferencing context. Moreover, the current findings also indicate that users readily identify the advantages of the multiple-channel display – an evaluation not previously made. The reasons provided for the multiple-channel audio-display mode preference offer important insight into the locus of the phenomenon; the multiple-channel mode was seen to provide both a high level of realism and facilitated auditory localisation of the voices of the individual conversants. Each is addressed further below.

Firstly, participants indicated that the multiple-channel mode provided a high level of ‘realism’ which they likened to face-to-face interaction. This comparison is noteworthy given that face-to-face interaction is considered the benchmark in
communication (Kang & Kim, 1996). It would be interesting in future research to contrast single- and multiple-channel audio-only communication with face-to-face interactions to evaluate to what degree the multiple-channel display mode approximates performance in face-to-face interaction.

Secondly, participants felt they were able to identify the specifics of each conversants’ contributions; they reported that each headphone audio speaker came to represent a single group member, helping them to attribute multiple speech streams to separate spatial origins.

In Experiment 2 these two possibilities are examined further by investigating whether the multiple- over single-channel advantage is the result of the participants’ sharing of general content and conversant specific information, or both.

3. EXPERIMENT 2: Utility of single verses multi-channel audio displays in passive listening.

In Experiment 1 the impact of talker location on the ability of subjects to use multi-talker speech interactively to solve a problem was examined. Subjects were required to listen passively to a multi-talker interaction and their comprehension of that conversation was measured. Such a passive listening task is different to the interactive task in Experiment 1 in that spatial attention and talker selectivity is not related to the subjects’ communicational intent but must be derived from the presented material. In addition, post hoc comprehension of an interaction by a listener is quite a different measure of performance compared to the task outcome performance measures employed in Experiment 1. This measure probes comprehension of a complex interaction between talkers rather than the capacity to interactively use the information in a problem solving context.

Participants were asked to listen to a recording of a team interaction (from
Experiment 1 and selected due to the high level of conversation between team members). The first half of the conversation was presented via one audio-display mode (single- or multiple-channel) and the second half using the other display mode. Performance was compared under both presentation conditions based on their responses to a Comprehension Questionnaire, designed to target both general-content and conversant-specific information.

3.1 Method

3.1.1 Design

A mixed-methods design was used. The first within-subjects independent variable was audio-display type, with two levels: (i) single-channel and (ii) multiple-channel, and the second was the type of comprehension question, with two levels: (i) conversant-specific and (ii) content-general. A between-subjects independent variable, order of channel display exposure (single- → multiple-channel; or multiple- → single-channel), was introduced to evaluate carry-over or practice effects. The level of information exchange in each of the audio interfaces was measured by participant accuracy scores on the Comprehension Questionnaire, with the added preference questions (as in Experiment 1) to evaluate participants’ preferences for the two audio presentation displays.

3.1.2 Participants

Forty-eight participants (12 males, 36 females; 17 to 36 years, $M=18.49$, $SD=1.94$) were recruited from a pool of first year Psychology students at the University of Western Sydney. Each gave their informed voluntary consent prior to participating and each received course credit for their participation.
3.1.3 Apparatus

Participants were tested individually in a booth (1.3m x .9m x 1.9m) with three audio-speakers (Genelec Analogue 50/60Hz) orientated in front of the participant at -30°, 0° and +30° from the sagittal plane at a distance of 55cm from the participant. Each audio-speaker was connected via a sound card (Hammerfall DSP Multiface II) to a desktop computer. In the single-channel audio display all three sound files of one conversation (see Experiment 1 - 2.1.4 Stimulus Materials and Dependent Variables) were simultaneously presented through the single audio-speaker at 0° (see Figure 4a). In the multiple-channel audio display each of the three sound files were played simultaneously each to one of the three separate audio-speakers (see Figure 4b).

3.1.4 Stimulus Materials and Dependent Variables

Each of the fourteen recordings (7 groups x 2 interactions) analysed in Experiment 1 were reviewed, and a single recording was selected for use in Experiment 2. Selection was based on relatively high levels of conversational overlap (thus maximising informational and energetic masking), relatively small number of no-conversation periods and overall recording quality; a complex, yet natural, interaction. The recording was divided into two halves (first half = 10min, 12 sec and second half = 10min, 5sec) and these were used in the first and second audio-display conditions respectively (as these two parts were sequential, the order of recording presentations was not counterbalanced, but the audio-display conditions were).

Participants were required to complete a 5-option multiple-choice Comprehension Questionnaire following each half of the recording. Test items were designed to examine the ability of participants to recall content-general aspects of the conversation (e.g., “the group decided Mrs Klutz committed the crime…”) and conversant-specific statements made by the team members (e.g., “Ben felt that Mr Purloin committed the crime, however Adrian disagreed”). Pilot testing of 6
individuals on 36 proposed items revealed there were 20 items (10 questions for each half of the recording) with acceptable variability (correct responses in 30% to 70% of cases). A subsequent paired samples t-test indicated that the answers to the comprehension questions for the two halves of the recording were equitable with no significant performance difference found between the two question sets ($t(11)=1.78, p=.10$). Upon completion of the questionnaire following the second presentation, participants completed the two additional open ended preference questions (as in Experiment 1).

### 3.1.5 Procedure

Participants were tested individually. Participants listened to recording 1 (single- or multiple-channel, depending on order group) and then completed the Comprehension Questionnaire. After a five-minute break, participants listened to recording 2 (multiple- or single-channel), and then completed the Comprehension Questionnaire again, this time with the additional preference item added at the end.

A repeated measures t-test was used to compare the individuals’ comprehension scores in the single- verses multiple-channel modes. A thematic analysis was performed to evaluate participants’ preference data.

### 3.2 Results

#### 3.2.1 Comprehension Scores

Comprehension scores are graphically presented in Figure 5 for the two presentation order groups, in the single-channel and multiple-channel audio interfaces, for both content-general and conversant-specific questions. A $2 \times (2 \times 2)$ mixed methods analysis of variance showed that mean comprehension scores (out of 10) in the multiple-channel condition were significantly higher ($M=62\%, SE=2.3$) than those
in the single-channel ($M=46\%, SE=2.3$) condition ($F(1,46)=47.77, p<.001$, partial $\eta^2 = 0.51$), suggesting that the multiple-channel condition better facilitated comprehension and retention of multi-talker interactions than the single-channel condition. Channel type explains over 50% of the variance in participant performance and represents a large effect (Tabachnick & Fidell, 2007). There was also a significant main effect of general verses specific score type, $F(1,46)=46.35, p<.001$, partial $\eta^2 = 0.50$, indicating that mean scores for content-general comprehension ($M=62.3, SE=2.6$) were significantly higher than those for conversant-specific comprehension ($M=45.8, SE=2.2$), an understandable outcome given answers to the general questions required less attention to detail. The two-way interaction between display mode and question type was not significant indicating that overall neither the multiple- nor single-channel mode selectively enhanced information exchange specifically for a particular type of question - content-general or conversant-specific.

Turning to the order factor there was no main effect of channel presentation order, $F(1, 46)=.172, p=.68$, partial $\eta^2 = .004$, indicating no overall practice effect, but there was a significant 3-way interaction for order, channel type, and question type, $F(1,46)=8.97, p=.004$, partial $\eta^2 = 0.16$ (see Figure 5). Post hoc comparisons with alpha set at .006 (0.5/8 to control for familywise error; Tabachnick & Fidell, 1007) revealed that in the single → multiple presentation order there were significant increases from the first to second presentation for specific scores, $t(23)=4.80, p<0.000$, ($M_{\text{first}}=33.3, SE_{\text{first}}=3.9; M_{\text{second}}=60.5, SE_{\text{second}}=4.2$), but not general scores ($t(23)=1.88, p=.073$), whereas in the multiple → single presentation order there were reductions in performance scores from the first to the second presentation and these were significant for general, $t(23)=5.15, p<0.000$, ($M_{\text{first}}=72.5, SE_{\text{first}}=4.3; M_{\text{second}}=50.8, SD_{\text{second}}=4.8$), but not for specific scores ($t(23)=0.46, p=.232$).

These results indicate that participants generally performed better in the
multiple-channel than the single-channel condition, and were also better at content-general than conversant-specific questions. In addition, participants’ performance on conversant-specific questions was particularly poor when the single-channel condition was engaged with first and particularly improved when the single-channel condition followed the multiple-channel condition, suggesting that the multiple-channel condition potentiated subsequent information extraction in the more difficult single-channel condition for the more difficult conversant-specific questions.

3.2.2 Individual preferences

The multiple-channel display mode was preferred over the single-channel display mode by 44 of the 48 participants (92%). As for Experiment 1, thematic analyses of participants’ reasons identified two themes related to this preference, Voice Separation and Realism: (1) Voice separation: 92% of participants commented that the spatial separation of the audio-speakers made the three voices easier to discriminate. Comments included, “when two people were talking at once, [multiple channel] made it easier to decipher which person was saying what, instead of it all being mixed up [as in single-channel conditions]”. (2) Realism related to participant judgements of the ecological validity of the two interfaces including remarks about familiarity and comfort. Comments included, “a voice was allocated to each speaker and because of their positioning, it felt as though the people themselves were in the room with me, sitting in a circle”.

While the single-channel display mode was preferred by 4/48 of the participants only one of these participants actually performed at a superior level under the single- compared to multiple-channel condition. This suggests that participant preference ratings do not necessarily indicate the utility of such interfaces when measured more objectively.
3.3 Discussion

These results indicate that passive listening with a multi-channel audio display facilitates comprehension of multi-party conversations compared to a single channel display. These results with passive listeners reinforce the finding in Experiment 1 that participants involved in a teleconference extract more information from and prefer a multiple-channel display compared with a single-channel display. Moreover, these results extend those of Experiment 1 by showing that multiple-channel presentation is superior to single-channel not only with two information sources (Experiment 1) but also with three information sources (Experiment 2).

Participants generally performed better on the content-general than the conversant-specific questions. There was no two-way interaction of this with the multiple->single-channel effect so participants generally performed better on (i) content-general and (ii) conversant-specific questions in the multiple- than the single-channel condition. These results are consistent with the notion that there is greater release from both (i) energetic masking and (ii) informational masking in the multiple- than the single-channel condition, giving little indication of the specific locus of the multiple->single-channel advantage. However the three-way interaction of channel, question type, and order throws some light on this: the interaction shows that the multiple-channel presentation always results in better performance on content-general questions, whereas multiple-channel presentation only facilitates better performance on conversant-specific questions when the multiple-channel display occurs as the second of the two conditions. So it may be said that when participants are presented with the multiple-channel condition first there is release from both energetic and informational masking so they can extract both general and specific information (and this even potentiates later single-channel information
extraction, but when presented with a single-channel condition first, informational masking remains relatively secure so it is relatively more difficult to extract specific information. Further research investigating this issue more analytically with purpose-specific manipulations would be welcome.

4. General Discussion

The findings support the view that the localized presentation of conversational material, facilitated in this experiment by a multiple-channel presentation mode, significantly improves the level of information exchange between multiple conversants over that of a non-localized (single-channel) presentation.

Experiment 1 reflects a real-world real-time interactive encounter between three remote participants, and shows better and more efficient performance, along with participant preference for a multiple-channel mode, in which each of the two other speakers’ voices are spatially separated, over a single-channel interaction mode. In Experiment 2, these results were confirmed in a more passive single listener design, and extended to show that multiple-channel presentation also provides greater information transfer than single-channel presentation with three localized sound sources. Importantly, the participants’ preference data and the reasons they give for their preferences (‘Voice Separation’ and ‘Realism’) are consistent across the two experiments, suggesting that the performance measures taken in Experiment 2 would generalise to more interactive real-world environments (as in Experiment 1).

The results underline an important, as yet understudied, distinction for telecommunication research; single- and multiple-channel systems forms of audio-teleconferencing give rise to quite different outcomes, so it is neither justified nor helpful to group them under the single category of ‘audio teleconferencing’. A failure to acknowledge the difference between these two presentation modes may have led to
the benefits of multiple-channel systems being overlooked in past research. With this in mind, revisiting many of the problems previously associated with audio-only communication (e.g., inability to navigate conversations, poor sociality) (Graetz, Boyle, Kimble, Thompson, & Garloch, 1998; Harmon, et al., 1995) may be appropriate future research directions, not only in teleconferencing research but also in other applied fields where there is a reliance on communication interfaces (e.g., aviation and event security).

In these studies it has been shown that the localisation advantage identified in previous psychoacoustic tests of intelligibility (Arbogast, et al., 2005; Cowie, et al., 1982; Hawley, et al., 2004; Kidd, et al., 2005a) transfers to more applied ‘real world’ domains, such as teleconferencing. The results also successfully demonstrate techniques for both simulating real world conversations and measuring performance by way of comprehension which are not common in psychoacoustic research of speech (e.g Yost, et al., 1996). Further, the use of both quantitative and qualitative measures has provided converging evidence that supports the claim that the multiple-channel presentation of talkers is beneficial in terms of facilitating information exchange when compared to a single-channel presentation. By using these techniques these studies were again able to capture, in an experimental setting, the cocktail party environment originally conceived of by Cherry (1953).

While there has been much psychoacoustic research examining the role of spatial separation in auditory perception there is little attention to its role in active or passive communication scenarios typical of real world interactions (Bronkhorst, 2000; Yost, et al., 1996). Past psychoacoustic studies have provided insight into how sound localization enables spatial release from masking (Bronkhorst, 2000; Hawley, et al., 2004). In this study the investigation of the contribution of spatial release from masking in more natural communication scenarios has provided an important
grounding and illustrated the critical importance of this facility in the solution of the real world cocktail party problem.

5. Acknowledgements

We extend our great appreciation to Dr Jörg Buchholz, now at Denmark Technical University, for constructing the auditory presentations in each of the two experiments presented here.
6. References


Figure 1. Team member (TM) audio presentation in Experiment 1: (a) Monotic headphone presentation where participants each hear the voices of both other team members (TM$^1$ and TM$^2$) through both the left and right headphones, and; (b) Dichotic headphone presentation where participants hear the voices of TM$^1$ through the left headphone and TM$^2$ through the right headphone.
Figure 2. Mean (+SE) percentage team accuracy performance ($N=7$) on the problem solving tasks in single-channel and multiple-channel display modes.
Figure 3. Mean (+SE) team time-to-completion performance (in secs) \((N=7)\) on the problem solving tasks in single-channel and multiple-channel display modes.
Figure 4. Audio displays used in Experiment 2: (a) single-channel display mode in which participants are presented with the voices of all three recorded talkers from the single audio-speaker at $0^\circ$ at a distance of 55cm. The other two audio-speakers (-30° and +30°) remained silent throughout the experimental condition; and, (b) multiple-channel display mode in which participants are presented with the recorded voices of the three talkers individually from separate audio-speakers at -30°, 0°, and +30° from the sagittal plane at a distance of 55cm.
Figure 5. Mean Comprehension accuracy scores (+SE), divided to demonstrate presentation order single-channel → multiple-channel (n=24) and multiple-channel → single-channel (n=24). Bars represent participant scores in the single-channel (n=48) and multiple-channel (n=48) conditions across general-content (n=48) and conversant-specific (n=48) comprehension question types.
Running Head: ATTENTION AND COGNITIVE LOAD

Cognitive Load Impacts Cross-Modal Dual-Task Processing: Supporting Both Unitary- and Multiple-Resource Attention Models
ABSTRACT

Objective: To investigate whether cognitive load impacts how attentional resources are allocated during dual-task processing. Background: Unitary-resource models posit a collective attentional resource responsible for maintaining information processing across all modalities. In contrast, multiple-resource models posit a system of modality independent resources. The two models predict similar within-modal performance but seemingly contradictory cross-modal performance. This study hypothesizes that these discordant outcomes result from different researchers adopting tasks that impose different levels of cognitive load. Method: Participants monitored two simultaneously-presented alphabetic letter arrays consisting of a target and multiple distractor letters. The dual-task design required the identification of targets in within-modal (two spoken or two text-based letter arrays) and cross-modal (one spoken and one text-based letter array) trials. Cognitive load was manipulated by altering the number of distractor letters in each array. Results: Modality combination and cognitive load main effects were found, indicating (1) superior performance in cross-modal conditions (in terms of reaction time and accuracy) and (2) decreasing performance resulting from increasing cognitive load. An interaction effect indicated that the cross-modal advantage evident under low levels of cognitive load (supporting multiple-resource predictions) is absent under high levels of cognitive load (supporting unitary-resource predictions). Conclusion: Neither the unitary- or multiple-resource models alone can account for dual-task processing. Application: System designs influenced by predictions of unitary- or multiple-resource models separately may compromise system integrity. This research indicates that an appraisal of an operator’s tasks, and the associated cognitive load, is necessary to design a best-fit system that can optimize operator performance.

Key words: Attention; Resources; Cognitive Load; Unitary-Resource; Multiple-Resource
INTRODUCTION

At any single moment in time, most, if not all, of an individual’s senses are confronted with a considerably large number of stimuli (Driver, 2004). As attentional resources are finite (Driver, 2001), a role of the human attentional system is to select only a small number of these stimuli for further processing to avoid the entire system becoming overloaded with excessive information, most often indicated by a performance decrement (Ernst & Bülthoff, 2004). However, in many environments, particularly complex workplaces (e.g., hospital emergency departments, flight control centers), any suppression of information by the attentional system, result in important information being overlooked, with potentially fatal outcomes (Yates, 2000).

To combat the potential for such workplace oversights, much effort has been focused towards designing advanced technology systems which reduce the potential for human error (Hersh, 1999). These systems operate to present users with relevant information in a manner which is consistent with the capabilities of their attentional system, in turn taking full advantage of the individual’s attentional resources to maintain an optimum level of performance in all task scenarios (Allen & Abate, 1999).

The designers of such workplace systems rely on theoretical models of the human attentional system to inform their engineering directions. However, despite intensive research being undertaken in the area of attention, more recently with an emphasis on multi-modal attention, there remains no consensus on a single model of attention (Soto-Faraco, Morein-Zamir, & Kingstone, 2005). At present, a debate ensues between advocates of two distinct resource models of cross-modal attention. These models make seemingly incompatible performance predictions in multi-task situations – in particular they propose contradictory performance outcomes in cross-modal processing situations.
In the case of the supramodal-, or unitary-resource model, often associated with Kahneman (1973), attention is regarded as a general pool of resources from which the necessary resources may be distributed for all sensory and cognitive processing (Proctor & Read, 2006). The predictions of unitary-resource models revolve around the assumption of a single resource limitation. That is, all tasks can be undertaken across any or all modalities until the system resources are exhausted. Therefore, once this point of overload is reached, any type of additional processing will be completed at the expense of another, concurrent task. More specifically, in the case of a dual-task scenario, performance on a within-modal (i.e., two auditory tasks (AA) or two visual tasks (VV)) dual task should be equivalent to a cross-modal (i.e., one auditory and one visual (AV)) dual task – in which auditory and visual tasks are of similar difficulty. This is because both auditory and visual processing require a share of the attentional resources supplied from the same resource pool. Therefore, resource overload will occur as a function of the additive requirements of all concurrent tasks, irrespective of the modalities in which they are being processed (Proctor & Read, 2006).

On the other side of the resource argument is the multiple-resource model (Wickens, 2002) which posits a number of modality-specific stores of attentional resources. The predictions of multiple-resource models stem from the assumption that each modality (e.g., auditory and visual) has separate pools of resources at their disposal. Therefore, while AA or VV dual-tasks will compete for resources (in much the same fashion proposed by the unitary-resource account) AV dual-tasks will be resourced by two discrete resource pools – one from an auditory and the other from a visual resource. Therefore in AV conditions competition for resources and any associated performance decrements should not occur.

Clearly the cross-modal predictions of the unitary- and multiple- resource models are in conflict with one another, but this has not prevented an extensive number of reports of empirical support for both the unitary-resource (Soto-Faraco et al., 2005; Arnell & Jenkins,
A review of the literature reveals an interesting point of discrimination between much of the research supporting the unitary-resource and the multiple-resource accounts. Specifically, research designs that employed tasks involving simple stimuli and simple detection tasks (e.g., auditory tone changes; visual chromatic shifts), and/or tasks which might be considered highly automated (e.g., driving while talking), appeared to support a multiple-resource account of auditory and visual processing (Alais et al., 2006; Wickens, 2002). In contrast, tasks involving more complex stimuli and identification paradigms (e.g., identify a target word; simple problem solving; voice shadowing) and other, more complex tasks where the participants have had little, if any, previous exposure, have provided support for a unitary-resource account (Soto-Faraco et al., Arnell & Jenkins, 2004).

It follows that the experimental design selections made by researchers may affect the difficulty of the task and in turn that the data support cross-modal performance of participants and accordingly the model to which the research offers. Therefore, it is worth speculating whether the human attentional structure adheres to the frameworks proposed by either unitary- and multiple-resource enthusiasts alone or whether it is possible that a combination of the two frameworks operate in a codependent fashion whereby the difficulty of the task may impact how attentional resources are pooled.

As attention is a type of cognitive resource, it appears sensible to consider a theoretical perspective which models the relationship between cognitive abilities, performance, and task complexity and/or difficulty. *Cognitive Load Theory* (CLT) (Sweller, 1988) posits that every task in which an individual engages will require some investment of his/her limited cognitive resources – in this case, attention. The amount of information...
required to be processed at any given time will result in a load or demand being placed on these resources. The size of this demand will depend upon the type of task and the abilities of the individual engaging in the task (and the motivation of the operator) (Brünken, Plass & Leutner, 2003; Woods, 1988). Essentially, CLT predicts that the more information required to be processed at any one time, the greater the demands will be on the available cognitive resources and that this can only be mediated by exposure or experience with the specific task (Shanteau, 1992).

Being able to determine how attentional resources are deployed under a range of multi-task circumstances is essential in developing an accurate model of human information processing. Moreover, evaluation of how the unitary- and multiple-resource models may work together, rather than in opposition, is a necessary step in trying to understand why such diverse findings are so regularly reported and why there is continued support, in separate experiments, for two seemingly incompatible models. Establishing a comprehensive model – that might explain the previous conflicting reports in this field is essential for workplace designers so that they may present information to users in a manner which makes maximum use of the limited attentional resources on offer and thus maximizes performance.

Accordingly, the aim here is to investigate whether both the unitary- and multiple-resource models can be supported in the same experiment by manipulating only the level of cognitive load imposed all within a single experiment. It is predicted that, under levels of low cognitive load, cross-modal processing will be significantly superior to within-modal processing indicating support for a multiple-resource model, whereas under levels of high cognitive load, cross-modal processing will be similar to within-modal processing, indicating support for a unitary resource model.
METHOD

The potential role of cognitive load in resourcing attention was investigated here by evaluating performance in auditory and visual modalities. These two modalities where used because it was reasoned that if any interactive processes were to occur between any two modalities, it would occur between the auditory and visual systems which are used in parallel during much perceptual processing (Lloyd, et al., 2006). Moreover, auditory-visual illusions documented in the literature (e.g., McGurk Effect (McGurk & Macdonald, 1976)) suggest that a close relationship exists between the two modalities.

A letter identification task was selected on the basis that it has been used extensively in previous research and does not appear to be exclusively associated with either of the resource models under investigation. Auditory letters were presented in rapid auditory presentation (RAP) sequences while the visual stimuli were presented in similar Rapid Serial Visual Presentation sequences (RSVP). A dual-task design was used, such that participants were required to monitor two concurrent letter sequences (one on their left hand side (LHS), the other on their right hand side (RHS)). These letter sequences could be both auditory (AA), both visual (VV), or, one auditory together with one visual (AV). The target letter to be identified remained the same in both modality sequences and appeared randomly within both concurrent letter sequences on each trial.

Consistent with previous research (Simon, 1997; Brünken, Steinbacher, Plass, & Leutner, 2002), cognitive load was manipulated by altering the number of items to be processed within a given time period. All trials were of 4,500ms duration however the number of distractor letters presented during this time frame was manipulated to create varying levels of cognitive load.
**Design**

A (3 x 3) within-subjects design was used in this experiment. The first Independent Variable (IV) was modality combination which consisted of three levels: (i) auditory/auditory (AA); (ii) visual/visual (VV); and (iii) auditory/visual (AV). The second IV was cognitive load, again, with three levels: (i) low cognitive load (LCL) - two distractor letters on both the LHS and RHS; (ii) medium cognitive load (MCL) – five distractor letters on both the LHS and RHS; and (iii) high cognitive load (HCL) – eleven distractor letters on both the LHS and RHS. The target letter was presented in all trials, that is, it was presented on both the LHS and RHS letter strings in every trial.

Two Dependent Variables (DV) were of interest – reaction time (RT) (from the time the target is presented to the corresponding key press target identification occurring within a valid time frame window) and accuracy (when two valid key presses were recorded for the trial). Valid target identification was deemed to be when a key press was recorded within a window of 150ms to 384.2ms following the presentation of each target. The lower limit of 150ms was adopted according to time frame window used by Spence and Driver (1996) where it was concluded that responses prior to 150ms were not representative of voluntary attentional processes. The upper limit in this experiment was determined in a single-task pilot study (males = 5, females = 6) utilizing the same stimuli but presenting only single letter sequences (for a total of 50 trials). A median (as mean scores tend to be skewed in such tasks (Spence & Driver, 1996; Ward, 1994)) for this group of participants was established \( Med = 360.0 \text{ms, } SD = 12.1 \text{ms} \) and an upper limit two standard deviations above this median was accepted. This upper limit is considerably shorter than the one adopted in the Spence and Driver study because accepting their upper limit of 1,500ms here would make it impossible to distinguish which of three different letters falling within the 150-1,500ms range represented the participants ‘target identification’.
These time windows were also used to evaluate identification accuracy: correct responses were those in which a valid key press was recorded in response to the presentation of both targets (one on the LHS, the other on the RHS) in each trial.

**Participants**

Twenty-two participants took part in this experiment, 14 males and 8 females, with a mean age of 27.8 (SD=6.2) for males and 25.3 (SD=3.8) for females. All participants provided informed consent and received $50 reimbursement for their travel and time.

**Stimulus Construction**

The auditory stimuli in this experiment consisted of strings of English alphabetic letters articulated in a Word 2003 based Text-To-Speech add on, using an English-Australian Male voice, and recorded using Cool Edit (2003, version 2.1). From pilot testing it was determined that the optimal target letter for use in the experiment would be the letter “B”, and the auditory distractor letters were determined to be U, O, Z, I and L as these were rated as being highly dissimilar to “B” (to minimize perceptual masking). Pronunciation of letters ranged in duration from 147ms to 190ms and were then compressed to a uniform 150ms. All compressed letters were pilot tested alongside the original length letters and it was found that compression did not impede the speed of identification.

These letters were then formatted into rapid auditory presentation (RAP) strings of 4,100ms in duration with either 3, 6, or 12 letters presented evenly spaced during this time frame to create strings of varying difficulty. The target letter was present in every string, with distractor letters randomly allocated to the remaining letter positions.

The visual stimuli used consisted of strings of English alphabetic letters, presented on a computer monitor in Times New Roman size 72 font (capitalized) for 150ms each. Again,
the target letter was “B” – primarily to reduce confusion in tracking different target letters. The visual distractor letters were Y, J, Q, N and X as, again, these were rated as being highly visually dissimilar to the target during pilot testing. These text letters were then formatted into rapid serial visual presentation (RSVP) strings. Consistent with the RAP strings the total string duration was 4,100ms, during which time, either 3, 6 or 12 text letters were presented. Again, the target letter was present in every string, with the visual-distractor letters filling the remaining letter positions.

From the base RAP and RSVP strings, ‘string pairing’s’ stimuli were prepared to facilitate dual monitoring (a visual depiction of the string pairing sequences can be seen in Figure 1). At the beginning of each trial a 400ms ‘readying’ sequence was introduced consisting of 125ms of silence, 150ms of visual hashes (####) and white noise, and then a further 125ms of silence before the onset of the string pairing’s. In the AA condition, two RAP strings were paired so that one string was presented on the participant’s LHS earphone, while the other was presented on the RHS earphone. Similarly, in the VV condition, two RSVP strings were paired, with one RSVP string presented on the LHS of the computer monitor and the other presented on the RHS of the computer monitor. For the AV condition, a RAP string was presented to one earphone, while an RSVP string was presented on the contralateral side of the computer monitor. All of these string pairings consisted of either two 3-letter strings (low cognitive load), two 6-letter strings (medium cognitive load) or two 12-letter strings string (high cognitive load). At no time throughout the experiment did two strings appear on the same presentation side and further, the strings within each pairing were temporally offset so that no individual letter stimulus overlapped with another during the trial.

*** Figure 1 about here ***
**Procedure**

Participants were seated before a 17inch wide-screen laptop computer (LG 1.83 GHz Genuine Intel® CPU) for the text-based presentation, connected to 2 auditory loud speakers (Sennheiser HD 650) mounted above the computer monitor, for the auditory spoken letter presentation. The horizontal displacement of the visual stimuli and auditory stimuli was 34mm a distance not expected to be noticeable to participants (Broadley & Kirkland, 1979).

Participants completed 324 test trials (plus 18 practice trials) during the experiment (AA=108, VV=108 and AV=108). Each set of 108 trails contained 36 trials under each level of cognitive load level. The 324 trials were randomly presented to the participant via the DMDX (Forster & Forster, 2003) auditory and visual stimuli presentation and timing response Windows program.

For each experimental trial, participants were first instructed (via the computer screen) as to whether the ensuing trial would be AA, VV or AV. Following a short period of visual and auditory silence, a sequence of hashes (####) appeared on the screen along with white noise via the headphones. Following a further period of auditory and visual silence the AA, VV or AV stimulus strings were presented. Participants were required to identify the target letter on the LHS (irrespective of whether auditory or visual) by pressing the left shift key and the target letter on the RHS by pressing the right shift key. Participants were instructed that the target letter could be present in either one or both strings and that they were required to press either one or both shift keys as appropriate. The left and right shift keys were colour-coded red and blue for ease of discrimination.

**RESULTS**

Two sets of analyses were conducted: one for RTs and one for accuracy. Trials included in the analyses were those where two valid responses were recorded. Any trial in
which one (or both) response were either not recorded, or did not fall within the valid time frame window were excluded from analysis. Trials where only one identification was required were used for performance exclusion purposes, so that if a participant performed below an accuracy of 80% they would be excluded from further analysis. One participant was excluded from the analysis based on this analysis.

For each dependent variable a repeated measures ANOVA was conducted to assess the mean differences in performance across modality conditions (AA, VV and AV) and load condition (LCL, MCL and HCL). Assumptions of normality, for both RTs and accuracy, were satisfactory although RT data did appear to be slightly positively skewed toward the upper limit of 384.2ms. However, given the robust nature of the experiment and that the valid time window was quite small the skewed nature of the data was not considered to violate the assumption of normality (in accordance with the recommendation of Tabachnick and Fidell, 2007). Additionally, Sphericity was not assumed for the main effects or the interaction effect analyses. Consequently, a Greenhouse-Geisser correction was implemented before interpretation of the results was made. Alpha was set at .05 for the analysis of main and interaction effects.

**Reaction Time**

As can be see in Figure 2, there appears to be no differences between the speed of processing over modality conditions. The analyses bear this out; there were no significant differences between cross- and within-modal conditions ($F_{AV\ vs\ AA+VV}(1,24)=3.12, p=.09$, partial $\eta^2=.12$) or between the two within-modal conditions ($F_{AA\ vs\ VV}(1,24)=.07, p=.80$, partial $\eta^2=.01$). Further, the results show that increasing levels of cognitive load does not alter the speed of processing ($F_{Lin}(1,24)=.07, p=.80$, partial $\eta^2=.01$; $F_{Quad}(1,24)=.43, p=.52$, partial $\eta^2=.02$). No interactions between modality combination and cognitive load were evident.
Taken together these results suggest that participants’ reaction times to the presentation of the target stimuli were similar across all experimental conditions in this experiment. This could indicate that, at least with the task used here, changes in stimulus presentation do not affect reaction time, or that the valid reaction time windows (that is, the truncated RTs) imposed in this experiment did not allow for a complete evaluation of the longer reaction times of participants.

Individual differences were also evident in the data, with 23% of participants responding on the majority of trials within the first 100ms, and where responses were not made inside this 100ms they in fact did not respond at all during the trial in 92% of cases. In contrast, a further 17% of participants consistently responded within the final 100ms and often times made responses just outside of the pre-established valid time frame.

### Figure 2 about here###

**Accuracy**

As can be seen in Figure 3, AV performance is consistently better than AA or VV performance; generally cross-modal presentation is associated with more accurate performance than within-modal presentation, while no differences in accuracy can be seen between the two within-modal conditions ($F_{AV vs. AA+VV}(1,24)=42.77, p<.001$, partial $\eta^2=.64$; $F_{AA vs. VV}(1,24)=3.63, p=.07$, partial $\eta^2=.13$). In terms of cognitive load there is a linear, but not a quadratic, decrease in accuracy performance across the increasing levels of cognitive load ($F_{Lin}(1,24)=146.57, p<.001$, partial $\eta^2=.86$; $F_{Quad}(1,24)=2.15, p=.16$, partial $\eta^2=.08$). As such, this indicates that generally as cognitive load increases accuracy performance steadily decreases.

### Figure 3 about here ###
An interaction between cross- versus within-modal and the linear cognitive load trend is also seen with cross-modal performance decreasing more rapidly across the increasing levels of cognitive load compared to the shallower decline in within-modal performance across the same levels of cognitive load ($F_{AV vs. AA+VV^*Lin}(1,24)=14.18$, $p=.001$, partial $\eta^2=.37$). This interaction shows that while performance for AA and VV trials steadily decreases across the load conditions, in the AV condition performance remains quite consistent between the LCL and MCL conditions, but quickly declines between the MCL and HCL conditions. In fact, the overall reduction in performance for cross-modal processing is substantially larger than the overall reduction in performance in the two within-modal conditions.

Given the individual differences apparent in the RT data it follows that for some participants who regularly made responses just outside of the valid response window accuracy scores were somewhat lower than other participants. There was no evidence to suggest that those reactions made just outside the valid window were more likely to belong to any of the modality combination or cognitive load conditions more than others. As such any individual differences seen were felt equally across all experimental conditions.

**DISCUSSION**

It was hypothesized that when participants operated in conditions of low cognitive load results consistent with the predictions of a multiple-resourcing account of the attention system would be yielded—specifically, that cross-modal performance would be superior to within-modal performance; whereas under high cognitive load on a similar task, participants would yield results reflecting the predictions of a unitary account of attentional resourcing—specifically, that cross-modal and within-modal performance would be equivalent. Both of these hypotheses were supported.
Specifically, the results demonstrate that under conditions of low cognitive load cross-modal accuracy was close to ceiling at over 94%, whilst at a same level of cognitive load within-modal performance was substantially lower (80.16-81.22%). Such a superior outcome for cross-modal performance is consistent with the expectations of a multiple-resourcing account of attention. However, under conditions of high cognitive load, while cross-modal performance remained superior to within-modal performance (a ‘cross-modal advantage’) the cross-modal advantage was no longer statistically significant – an outcome inconsistent with the predictions of a multiple-resourcing account (Wickens, 2002; Alais et al., 2006; Duncan et al., 1997). Instead, the results suggest that, under higher levels of cognitive load, percent correct performance is similar irrespective of whether two tasks are undertaken in the same or different modalities – an outcome consistent with the claims of unitary-resource supporters (Soto-Faraco et al., 2005; Arnell & Jenkins, 2004; Joliceour 1999).

While it was anticipated that a cross-modal advantage would be evidenced by increasing percent accuracy accompanied with decreasing reaction times this prediction was not supported. While heightened cross-modal accuracy was seen to be superior to within-modal accuracy there appeared to be no significant difference in RTs. This outcome was not considered to obviate the claim that a cross-modal advantage was demonstrated under conditions of lower cognitive load as due to the small reaction time window used to define valid responses it may well be that any valuable differences between cross- and within-modal processing are evident only when a full spectrum of RTs are evaluated – as in Spence and Driver (1996). Moreover, numerous attentional resource studies used accuracy scores, in the absence of RT data, to measure performance in cross- and within-modal dual task scenarios (Sinnet, Costa & Soto-Faraco, 2006; Potter, Chun, Banks & Muckenhoupt, 1998; Soto-Faraco et al., 2005)
In order to combat the potential biases associated with previous research methods, the present study made use of a task which could be easily manipulated to allow varying levels of cognitive load in within- and cross-modal conditions. These results suggest that neither the unitary- nor multiple-resources accounts alone can explain how attentional resources are allocated during information processing. This supports our claim that the experimental methods selected by past researchers may have inadvertently influenced their results and in turn the resource model the results supported. For instance, Alais et al. (2006) found support for a multiple-resource model of attention using visual grating patches and auditory pitch contrasts – stimulus conditions we believe are consistent with inducing low cognitive load. In contrast, Jolicour (1999), supported a unitary-resourcing account when he used tone identification and alphabetic letter sequences, together with speeded and recall tasks – methods we believe invoke relatively greater levels of cognitive load.

The results here suggest there may be a need to incorporate the features of both the unitary- and multiple-resource models into a more comprehensive and complex arrangement - maybe a hybrid or hierarchical framework. Such a suggestion is not novel in the literature (Spence & Driver, 1996; Burr & Alais, 2006; Schwartz, Berthommier, & Savariaux, 2004; Posner, 1990), however, the interactivity or relationship between the two stand alone models has not until now been investigated.

The development and testing of such a hierarchical model is presently being investigated by the authors. The eventual aim is to test such a model in simulated environments in which individuals interact with more meaningful tasks (e.g., train controlling or air traffic control). Moreover, as individual differences were demonstrated in this experiment (some participants performed significantly faster than other participants across all conditions), further studies are now being conducted to explore the degree to which domain experience and motivation may affect an individuals attentional capabilities. Importantly,
despite such differences in individual performance, the cross-modal advantages found here were present for nearly all participants. This finding is fundamental for system designers as it suggests that by and large there is significant value to presenting individuals with information cross-modally, for even though performance (1) may vary across individuals and across cognitive load levels and (2) the cross-modal advantage may diminish under heightened levels of cognitive load, there was certainly no evidence to suggest cross-modal presentation was in any way detrimental to performance.

The present findings and the subsequent development of a comprehensive attentional resourcing model has significant implications for cognitive psychological theory and human factors research. Just as importantly, the present findings have important implications for many applied domains in which swift and accurate responses to auditory and visual stimuli are required (e.g., tactical response, flight control, rail control). The results indicate that a certain cross-modal advantage can be capitalized upon if a task imposes low cognitive load, nevertheless, a reliance on a limitless cross-modal advantage (an advantage purported by multiple-resource models) may be unrealistic, unproductive and at worst dangerous in some workplace environments.
REFERENCES


Figure 1. A depiction of the stimulus string sequences (all lasting for 4,500ms) across three levels of cognitive load: (a) Low, consisting of 1 target and 2 distractor letters on both the LHS and RHS (shown here as an AA trial) (b) Medium, consisting of 1 target and 5 distractor letters on both the LHS and RHS (shown here as a VV trial) and (c) High, consisting of 1 target and 11 distractor letters on both the LHS and RHS (shown here as an AV trial).
Figure 2. Mean participant RTs (+SE) across modality combination and cognitive load conditions.
Figure 3. Mean participant accuracy percent scores (+SE) across all modality combinations (AA, VV, and AV) and cognitive load levels (low, medium and high).
Appendix J

Study Three: Table of Participant Reaction Times to Single-Task Trials and Subsequent Individualised RT Windows

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean Reaction Time (ms)</th>
<th>SD (ms)</th>
<th>Established RT Window*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>231</td>
<td>27</td>
<td>177 – 285</td>
</tr>
<tr>
<td>2</td>
<td>248</td>
<td>26</td>
<td>196 – 300</td>
</tr>
<tr>
<td>3</td>
<td>189</td>
<td>18</td>
<td>153 – 225</td>
</tr>
<tr>
<td>4</td>
<td>155</td>
<td>13</td>
<td>150 – 221</td>
</tr>
<tr>
<td>5</td>
<td>300</td>
<td>16</td>
<td>268 – 332</td>
</tr>
<tr>
<td>6</td>
<td>255</td>
<td>19</td>
<td>217 – 293</td>
</tr>
<tr>
<td>7</td>
<td>241</td>
<td>36</td>
<td>169 – 313</td>
</tr>
<tr>
<td>8</td>
<td>301</td>
<td>35</td>
<td>231 – 371</td>
</tr>
<tr>
<td>9</td>
<td>168</td>
<td>12</td>
<td>144 – 192</td>
</tr>
<tr>
<td>10</td>
<td>172</td>
<td>16</td>
<td>140 – 204</td>
</tr>
<tr>
<td>11</td>
<td>171</td>
<td>17</td>
<td>137 – 205</td>
</tr>
<tr>
<td>12</td>
<td>210</td>
<td>35</td>
<td>140 – 280</td>
</tr>
<tr>
<td>13</td>
<td>178</td>
<td>12</td>
<td>154 – 202</td>
</tr>
<tr>
<td>14</td>
<td>289</td>
<td>22</td>
<td>245 – 333</td>
</tr>
<tr>
<td>15</td>
<td>244</td>
<td>26</td>
<td>192 – 296</td>
</tr>
<tr>
<td>16</td>
<td>255</td>
<td>20</td>
<td>215 – 295</td>
</tr>
<tr>
<td>17</td>
<td>288</td>
<td>39</td>
<td>210 – 366</td>
</tr>
<tr>
<td>18</td>
<td>390</td>
<td>22</td>
<td>346 – 434</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>19</td>
<td>350</td>
<td>26</td>
<td>298 – 402</td>
</tr>
<tr>
<td>20</td>
<td>244</td>
<td>25</td>
<td>194 – 294</td>
</tr>
<tr>
<td>21</td>
<td>310</td>
<td>19</td>
<td>272 – 348</td>
</tr>
<tr>
<td>22</td>
<td>158</td>
<td>12</td>
<td>134 – 182</td>
</tr>
<tr>
<td>23</td>
<td>187</td>
<td>34</td>
<td>119 – 255</td>
</tr>
<tr>
<td>24</td>
<td>222</td>
<td>33</td>
<td>156 – 288</td>
</tr>
<tr>
<td>25</td>
<td>203</td>
<td>35</td>
<td>133 – 273</td>
</tr>
<tr>
<td>26</td>
<td>265</td>
<td>35</td>
<td>195 – 335</td>
</tr>
<tr>
<td>27</td>
<td>301</td>
<td>20</td>
<td>261 – 341</td>
</tr>
<tr>
<td>28</td>
<td>355</td>
<td>19</td>
<td>317 – 393</td>
</tr>
<tr>
<td>29</td>
<td>300</td>
<td>39</td>
<td>222 – 378</td>
</tr>
<tr>
<td>30</td>
<td>315</td>
<td>25</td>
<td>265 – 365</td>
</tr>
<tr>
<td>31</td>
<td>336</td>
<td>25</td>
<td>286 – 386</td>
</tr>
<tr>
<td>32</td>
<td>290</td>
<td>24</td>
<td>242 – 338</td>
</tr>
<tr>
<td>33</td>
<td>303</td>
<td>26</td>
<td>251 – 355</td>
</tr>
<tr>
<td>34</td>
<td>188</td>
<td>18</td>
<td>152 – 224</td>
</tr>
<tr>
<td>35</td>
<td>201</td>
<td>27</td>
<td>147 – 255</td>
</tr>
<tr>
<td>36</td>
<td>401</td>
<td>33</td>
<td>335 – 467</td>
</tr>
<tr>
<td>37</td>
<td>350</td>
<td>24</td>
<td>302 – 398</td>
</tr>
<tr>
<td>38</td>
<td>324</td>
<td>18</td>
<td>288 – 360</td>
</tr>
<tr>
<td>39</td>
<td>325</td>
<td>18</td>
<td>289 – 361</td>
</tr>
<tr>
<td>40</td>
<td>389</td>
<td>20</td>
<td>349 – 429</td>
</tr>
<tr>
<td>41</td>
<td>371</td>
<td>24</td>
<td>323 – 419</td>
</tr>
<tr>
<td>42</td>
<td>364</td>
<td>23</td>
<td>318 – 410</td>
</tr>
<tr>
<td>43</td>
<td>366</td>
<td>22</td>
<td>322 – 410</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>44</td>
<td>311</td>
<td>19</td>
<td>273 – 349</td>
</tr>
<tr>
<td>45</td>
<td>332</td>
<td>28</td>
<td>276 – 388</td>
</tr>
<tr>
<td>46</td>
<td>189</td>
<td>27</td>
<td>135 – 243</td>
</tr>
<tr>
<td>47</td>
<td>177</td>
<td>24</td>
<td>129 – 225</td>
</tr>
<tr>
<td>48</td>
<td>201</td>
<td>26</td>
<td>149 – 253</td>
</tr>
<tr>
<td>49</td>
<td>389</td>
<td>25</td>
<td>339 – 439</td>
</tr>
</tbody>
</table>

* Where the lower bound was <150ms, 150ms was taken to be the lower limit of response times, due to the findings of Spence and Driver (1996) suggesting 150ms is the faster possible response time.
### Appendix J

**Study Three: Table of Planned Contrasts**

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Contrast Type</th>
<th>Condition 1</th>
<th>Condition 2</th>
<th>Condition 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modality Combination</td>
<td>Linear (Within vs. Cross)</td>
<td>AA</td>
<td>VV</td>
<td>AV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1</td>
<td>-1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Quadratic (Within comparison)</td>
<td>LCL</td>
<td>MCL</td>
<td>HCL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td>Linear</td>
<td>Motivated</td>
<td>Non-Motivated</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>Linear</td>
<td>Exposure</td>
<td>No-Exposure</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>
Appendix K

Ratings Questionnaire

Participant Number
**Trial Number: 1**

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>How difficult did you feel this trial was?</td>
<td>1-5</td>
<td>It was easy - It was hard</td>
</tr>
<tr>
<td>How confident are you in your performance on this trial?</td>
<td>1-5</td>
<td>Not confident - Very confident</td>
</tr>
<tr>
<td>How many distractor letters were in this trial?</td>
<td>1-5</td>
<td>Few - Lots</td>
</tr>
<tr>
<td>How motivated were you as you completed this trial?</td>
<td>1-5</td>
<td>Not at all - Very Motivated</td>
</tr>
</tbody>
</table>
Trial Number: 2

How difficult did you feel this trial was?

1  2  3  4  5
It was easy

How confident are you in your performance on this trial?

1  2  3  4  5
Not confident

How many distractor letters were in this trial?

1  2  3  4  5
Few

How motivated were you as you completed this trial?

1  2  3  4  5
Not at all

Motivated

Very

Motivated
Trial Number: …

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>How difficult did you feel this trial was?</td>
<td>1-5</td>
<td>It was easy - It was hard</td>
</tr>
<tr>
<td>How confident are you in your performance on this trial?</td>
<td>1-5</td>
<td>Not confident - Very confident</td>
</tr>
<tr>
<td>How many distractor letters were in this trial?</td>
<td>1-5</td>
<td>Few - Lots</td>
</tr>
<tr>
<td>How motivated were you as you completed this trial?</td>
<td>1-5</td>
<td>Not at all - Very Motivated</td>
</tr>
</tbody>
</table>
### Appendix L

Study Four: Table of Participant Reaction Times to Single-Task Trials and Subsequent Individualised RT Windows

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean Reaction Time (ms)</th>
<th>SD (ms)</th>
<th>Established RT Window*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>244</td>
<td>25</td>
<td>194 – 294</td>
</tr>
<tr>
<td>2</td>
<td>310</td>
<td>19</td>
<td>272 – 348</td>
</tr>
<tr>
<td>3</td>
<td>158</td>
<td>12</td>
<td>134 – 182</td>
</tr>
<tr>
<td>4</td>
<td>187</td>
<td>34</td>
<td>119 – 255</td>
</tr>
<tr>
<td>5</td>
<td>222</td>
<td>33</td>
<td>156 – 288</td>
</tr>
<tr>
<td>6</td>
<td>203</td>
<td>35</td>
<td>133 – 273</td>
</tr>
<tr>
<td>7</td>
<td>265</td>
<td>35</td>
<td>195 – 335</td>
</tr>
<tr>
<td>8</td>
<td>301</td>
<td>20</td>
<td>261 – 341</td>
</tr>
<tr>
<td>9</td>
<td>355</td>
<td>19</td>
<td>317 – 393</td>
</tr>
<tr>
<td>10</td>
<td>300</td>
<td>39</td>
<td>222 – 378</td>
</tr>
<tr>
<td>11</td>
<td>315</td>
<td>25</td>
<td>265 – 365</td>
</tr>
<tr>
<td>12</td>
<td>336</td>
<td>25</td>
<td>286 – 386</td>
</tr>
<tr>
<td>13</td>
<td>290</td>
<td>24</td>
<td>242 – 338</td>
</tr>
<tr>
<td>14</td>
<td>303</td>
<td>26</td>
<td>251 – 355</td>
</tr>
<tr>
<td>15</td>
<td>188</td>
<td>18</td>
<td>152 – 224</td>
</tr>
<tr>
<td>16</td>
<td>201</td>
<td>27</td>
<td>147 – 255</td>
</tr>
<tr>
<td>17</td>
<td>401</td>
<td>33</td>
<td>335 – 467</td>
</tr>
<tr>
<td>18</td>
<td>350</td>
<td>24</td>
<td>302 – 398</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>19</td>
<td>324</td>
<td>18</td>
<td>288</td>
</tr>
<tr>
<td>20</td>
<td>325</td>
<td>18</td>
<td>289</td>
</tr>
<tr>
<td>21</td>
<td>389</td>
<td>20</td>
<td>349</td>
</tr>
<tr>
<td>22</td>
<td>371</td>
<td>24</td>
<td>323</td>
</tr>
<tr>
<td>23</td>
<td>364</td>
<td>23</td>
<td>318</td>
</tr>
<tr>
<td>24</td>
<td>366</td>
<td>22</td>
<td>322</td>
</tr>
<tr>
<td>25</td>
<td>311</td>
<td>19</td>
<td>273</td>
</tr>
<tr>
<td>26</td>
<td>332</td>
<td>28</td>
<td>276</td>
</tr>
<tr>
<td>27</td>
<td>244</td>
<td>25</td>
<td>194</td>
</tr>
</tbody>
</table>

* Where the lower bound was <150ms, 150ms was taken to be the lower limit of response times, due to the findings of Spence and Driver (1996) suggesting 150ms is the faster possible response time.
## Appendix M

### Study Four: Table of Planned Contrasts

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Contrast Type</th>
<th>Condition Contrasts A</th>
<th>V</th>
<th>AA</th>
<th>VV</th>
<th>AV</th>
<th>AVa</th>
<th>AVv</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single vs. Dual</td>
<td>3</td>
<td>3</td>
<td>-2</td>
<td>-2</td>
<td>-2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Dual</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Dual vs. Tri</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Auditory vs. Visual</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Cross vs. Within</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>AA vs. VV</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>AVa vs. AVv</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

### Cognitive Load

<table>
<thead>
<tr>
<th>Cognitive Load</th>
<th>LCL</th>
<th>LMCL</th>
<th>MCL</th>
<th>MHCL</th>
<th>HCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Quadratic</td>
<td>2</td>
<td>-1</td>
<td>-2</td>
<td>-1</td>
<td>2</td>
</tr>
<tr>
<td>Cubic</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>Quartic</td>
<td>1</td>
<td>-4</td>
<td>6</td>
<td>-4</td>
<td>1</td>
</tr>
</tbody>
</table>