Understanding Reading and Learning from Multi-representational Texts in a Second Language

Nurjanah Binti Mohd Jaafar

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy,

September 2019

ABSTRACT

This thesis explores how learners read and learn from text and pictures in a second language. Although research into multi-representational materials has developed substantially in recent years, there is a tendency to focus mainly on first language (L1) learners. Less is known about how the growing population of second language (L2) learners engage and learn from these materials. To address this, an investigation into the reading behaviour of English as a second language speakers (or L2 readers) as well as its outcomes was conducted.

Three studies were carried out. The first study explored the reading behaviour of the L2 readers and its relationships with individual characteristics of interest, as well as outcomes of learning. The eye movements of 50 L2 university students were collected as they read a multi-representational science text and then took an outcome test. In the second study, 42 native speaking English university students (or L1 readers) read the same textbook and took the same test. The third study compared the reading behaviour and outcomes of learning from this multi-representational material in L1 and L2 and so the two datasets from Studies 1 and 2 were combined and analysed.

The findings, in general, demonstrated that multi-representational reading and learning in the L2 was similar to reading L1. Like the L1 readers, the L2 readers paid much more attention to text than pictures, and their learning was positively linked with integrating the two types of representations. Study 3 found that the L1 and the L2 readers learned equally well from the material but the L2 readers took more time and adopted more effortful strategies when reading by displaying greater text-picture integration. Hence, although the role of text-picture integration in multi-representational learning has been emphasised and reiterated, it appears that it plays an even more crucial role in multi-representational learning in L2.

This thesis offers important insights into multi-representational learning in the L2 and demonstrates that the models of multi-representational learning that have been widely observed and promoted in L1 apply to L2 and should therefore similarly be encouraged. Understandably, reading and learning in L2 is a harder task compared to reading and learning in L1. This thesis offers some reassurance to the L2 learners that the extra time and effort exerted are not in vain and can lead them towards successful learning from multiple representations.

ACKNOWLEDGEMENTS

Alhamdulillah. All praises and glory belong to The Almighty.

I wish to express my love and gratitude to my parents; my mother, Zaharah and my father, Mohd Jaafar. Also to my five siblings (their spouses and children) and my in-laws, thank you for the constant prayers, love and support. My husband, Binyamin, thank you for your love, support and sacrifice. May you all be showered with continuous blessings.

My deepest appreciation and gratitude to my supervisors; Professor Shaaron Ainsworth and Dr. Walter van Heuven. Shaaron has taught, guided and encouraged me with utmost patience throughout this journey. For all the teasing, anti-teasing cards and strange analogies, I am very much indebted. I am very honoured to have been supervised by one of the best in the field.

Walter has taught me the simplest things like right-clicking on a MacBook to the complex of things like eye tracking tips and Python scripts, and I am most thankful. For his critical input and feedback, I could not have asked for a better second supervisor.

My appreciation also goes out to:

My kind housemates, friends and the Malaysian community in Nottingham who have always made me feel close to home despite the distance,
My colleagues in the LSRI and the School of Education, especially Ipek and Faith - thank you for the friendship, lunches, chats and advice,
My mentor, Professor Thang Siew Ming, for her encouragements and guidance, and
My research participants for their generous time and effort.

Finally, this PhD would not have been possible without the funding from the Ministry of Education, Malaysia and my employer, Universiti Kebangsaan Malaysia.

TABLE OF CONTENTS

ABSTRACT **2**

ACKNOWLEDGEMENTS **3**

TABLE OF CONTENTS **4**

LIST OF FIGURES **9**

LIST OF TABLES **10**

1. INTRODUCTION **11**

1.1. PERSONAL DRIVE BEHIND THE RESEARCH AND ITS OBJECTIVE ... 13

1.2. STRUCTURE OF THE THESIS 14

2. LITERATURE REVIEW **16**

2.1. INTRODUCTION 16

2.2. EXTERNAL REPRESENTATIONS AND LEARNING 16

2.3. READING MULTI-REPRESENTATIONAL TEXTS IN THE L1 22

2.4. READING IN A SECOND LANGUAGE 29

2.4.1. Reading in the L2 30

2.4.2. Predictors of L2 Reading 33

2.4.3. Eye Movements as Indicators of Reading Differences between the L1 and L2 35

2.4.4. Conclusion 38

2.5. LEARNING A SECOND LANGUAGE WITH MULTIPLE REPRESENTATIONS 38

2.6. LEARNING WITH MULTIPLE REPRESENTATIONS IN A SECOND LANGUAGE 43

2.7. RESEARCH QUESTIONS 49

3. METHODOLOGY **53**

3.1. INTRODUCTION 53

3.2. THE POST-POSITIVIST APPROACH 53

3.3. RESEARCH DESIGN 55

3.3.1. Non-Experimental Fixed Design 55
3.4. SAMPLING 57
3.5. THE MATERIAL 59
3.5.1. Exploring the Material 61
3.5.1.1. The text 61
3.5.1.2. The pictures 64
3.5.1.3. Structural and functional perspectives of the pictures 64
3.5.1.4. Readers' perspectives of the representations 69
DeFT: The functions of multiple representations 69
The verbal reports 70
3.5.2. Segmentation of Areas of Interest 80
3.6. METHODS OF DATA COLLECTION 82
3.6.1. Eye Tracking 82
3.6.1.1. Introduction 82
3.6.1.2. Relevance and limitations of eye tracking 84
3.6.1.3. Eye tracking measures 88
3.6.1.4. Apparatus 95
3.6.2. Tests 95
3.6.2.1. English vocabulary knowledge 95
3.6.2.2. English reading comprehension ability 96
3.6.2.3. Prior knowledge 96
3.6.2.4. Learning outcome and coding 97
3.6.3. Questionnaire 98
3.7. METHODS OF DATA ANALYSIS 98
3.7.1. Preparation for Data Analysis 98
3.7.2. Statistical Considerations 99
3.7.2.1. Handling violations of assumptions 100
3.8. ETHICAL PROCESSES FOLLOWED 102
4. STUDY 1: THE L2 STUDY 104
4.1. INTRODUCTION 104
4.1.1. The Nature of Reading Multi-Representational Texts 104
4.1.2. Text-picture Integration and Learning 107
4.1.3. Reading of Multi-representational Texts and Individual Characteristics 109

4.2. HYPOTHESES 110
4.3. METHOD 110
4.3.1. Participants 110
4.3.2. Materials 111
4.3.3. Eye Tracking Measures 111
4.3.4. Apparatus 112
4.3.5. Procedure 112
4.4. DATA ANALYSIS 112
4.5. RESULTS 113
4.5.1. Distribution of Attention across text and Picture Areas 113
4.5.2. Patterns of Eye Movements 113
4.5.3. Eye Movements and Individual Characteristics 120
4.5.4. Eye Movements and Learning Outcome 121
4.6. DISCUSSION 122
4.7. CONCLUSION 127
5. STUDY 2: THE L1 STUDY 129
5.1. INTRODUCTION 129
5.2. HYPOTHESES 129
5.3. METHOD 130
5.3.1. Participants 130
5.3.2. Materials 130
5.3.3. Eye Tracking Measures and Apparatus 130
5.3.4. Procedure 131
5.4. RESULTS 131
5.4.1. Distribution of Attention across Text and Picture Areas 131
5.4.2. Patterns of Eye Movements 131
5.4.3. Eye Movements and Individual Characteristics 139
5.4.4. Eye Movements and Learning Outcome 140
5.5. DISCUSSION 141
5.6. CONCLUSION 144
6. STUDY 3: THE COMPARISON STUDY 145
6.1. INTRODUCTION 145

6.2. METHOD 147
6.2.1. RESULTS 147
6.2.2. Comparing the L1 and L2 Readers 147
6.2.2.1. Differences in individual characteristics 147
6.2.2.2. Differences in eye movements 148
Absolute analysis 149
Proportional analysis 150
6.2.2.3. Differences in learning outcome 152
6.2.3. Patterns of Reading Behaviour 154
6.2.3.1. Absolute cluster analysis 154
The role of language group and individual characteristics 160
Eye movements and learning outcome 162
6.2.3.2. Proportional cluster analysis 163
The role of language group and individual characteristics 168
Eye movements and learning outcome 170
6.3. DISCUSSION 171
6.4. CONCLUSION 174
7. GENERAL DISCUSSION 176
7.1. INTRODUCTION 176
7.2. RESEARCH QUESTION 1 176
7.3. RESEARCH QUESTION 2 179
7.4. RESEARCH QUESTION 3 181
7.5. RESEARCH QUESTION 4 182
7.6. LIMITATIONS AND FUTURE WORK 186
7.7. IMPLICATIONS 189
7.8. CONTRIBUTIONS AND CONCLUDING REMARKS 190
REFERENCES 192
APPENDICES 203
Appendix 1: Research Stimulus 203
Appendix 2: Shortlisted Research Stimulus 1 204
Appendix 3: Shortlisted Research Stimulus 2 205

Appendix 4: Prior Knowledge Task 206

Appendix 5: Learning Outcome Task 209

Appendix 6: Scoring Rubrics for Learning Outcome Task..... 210

LIST OF FIGURES

Figure 1. Examples of the various text components within the material 63
Figure 2. Picture A and a corresponding text component 66
Figure 3. Picture G and three corresponding text components 66
Figure 4. Text describing the arrangement of solid particles and the corresponding Picture A that visualises the arrangement 73
Figure 5. Text referring to the rubbish mover machine, a photo-realistic picture of the machine, along with text captions over and beneath Picture D 74
Figure 6. Picture G that visualises the text describing the properties of gas particles on its right 75
Figure 7. Pictures A, B and C, each representing a different state of matter 76
Figure 8. Segmentation of the material into AOIs. Pink AOIs are text AOIs whereas green AOIs are picture AOIs 81
Figure 9. An example of the various eye tracking measures on the stimulus and labelling of the eye tracking measures 94
Figure 10. Dendrogram using Ward's linkage for Study 1 115
Figure 11. Means of eye tracking measures by cluster 116
Figure 12. Scatterplot of composite scores for first-pass fixation measures and integrative measures 117
Figure 13. Typical examples of a high integrator's scan path (top) and a low integrator's scan path (bottom) 119
Figure 14. Dendrogram using Ward's linkage for Study 2 132
Figure 15. Means of eye tracking measures by cluster 134
Figure 16. Scatterplot of composite scores for first-pass fixation measures and integrative measures 135
Figure 17. Typical examples of a high integrator's scan path (top), an intermediate integrator's scan path (middle) and a low integrator's scan path (bottom) 138
Figure 18. Dendrogram using Ward's linkage for the absolute cluster analysis 156
Figure 19. Dendrogram using Ward's linkage for the proportional cluster analysis 165

LIST OF TABLES

Table 1. Summary of the characteristics of the sentential representations.....64
Table 2. The classification of the graphic representations based on Lohse et al. (1994) and Levin et al.'s (1987) taxonomies 67
Table 3. The coding scheme to explore participants' perceptions of the functions of representations based on Ainsworth's functional taxonomy 71
Table 4. Eye tracking measures and their definition largely based on Mason, Tornatora, et al.'s (2013) study 91
Table 5. Mean proportion of fixation measures by AOI type 113
Table 6. Descriptive statistics of eye tracking measures by cluster 116
Table 7. Analysis of variance for individual eye tracking measures 118
Table 8. Descriptive statistics of individual characteristic tests by cluster 120
Table 9. Analysis of variance for individual characteristic test 120
Table 10. Correlations between eye tracking and individual characteristic measures 121
Table 11. Breakdown of cluster by language group 121
Table 12. Correlations between individual characteristic and learning outcome test 122
Table 13. Mean proportion of fixation measures by AOI type 131
Table 14. Descriptive statistics of eye tracking measures by cluster 133
Table 15. Analysis of variance for eye tracking measures 135
Table 16. Descriptive statistics of individual characteristic tests by cluster 139
Table 17. Analysis of variance for individual characteristic tests 139
Table 18. Correlations between eye tracking measures and individual characteristic 140
Table 19. Descriptive statistics of learning outcome test by clusters 140
Table 20. Correlations between individual characteristic and learning outcome test 141
Table 21. Descriptive statistics for individual characteristic variables by language group 147
Table 22. Descriptive statistics for eye tracking measures by language groups.....149
Table 23. Analysis of variance for individual eye tracking measures 150
Table 24. Descriptive statistics for eye tracking measures (proportions) by language groups 151
Table 25. Analysis of variance for individual eye tracking measures (proportions) 151
Table 26. Correlations between learning outcome test and individual characteristic and reading time.....153
Table 27. Descriptive statistics for eye tracking measures by cluster 155
Table 28. Analysis of variance for individual eye tracking measures 157
Table 29. Distribution of language groups by clusters 161
Table 30. Descriptive statistics of individual characteristic tests by clusters.....161
Table 31. Descriptive statistics of learning outcome test by clusters 162
Table 32. Descriptive statistics for proportions of eye tracking measures by clusters 163
Table 33. Analysis of variance for individual proportions of eye tracking measures 166
Table 34. Distribution of language group by clusters 169
Table 35. Descriptive statistics of individual characteristic tests by clusters.....169
Table 36. Descriptive statistics of learning outcome test by clusters 170

1. INTRODUCTION

Print-based materials such as textbooks are now rich with visual representations such as photos, diagrams and graphs to accompany the more common verbal (written) representations. In science textbooks for instance, the number of visual representations has nearly doubled compared to their number half a century ago (Lee, 2010) . Given this trend, it comes as no surprise that research into learning with multi-representational materials has multiplied in recent years. Some researchers, for instance, became interested in the tasks and processes that take place during the engagement with these materials and what sort of outcomes are produced (Ainsworth, 2006; Hegarty & Just, 1993; Mayer, 2014). Design parameters and principles are also proposed to guide material designers and educators alike (Ainsworth, 2006; Mayer, 2014). Extensive research has found two important yet slightly contradictory findings; learners are pre-dominantly text-oriented when engaging with these materials (Hannus & Hyönä, 1999; Johnson & Mayer, 2012; Schmidt-Weigand, Kohnert, & Glowalla, 2010), and integrating information from the different sources are crucial for the learners to learn successfully with these materials (Hannus & Hyönä, 1999; Johnson & Mayer, 2012; Mason, Tornatora, & Pluchino, 2015; Mason, Tornatora, & Pluchino, 2013).

Despite its growth and development, research to date has been mostly restricted to learners who engage with these materials in their first language (L1). What of those who learn them in their non-native language then? Do the two key findings mentioned above apply when these materials are read in one's second language, for instance? Unfortunately, not much research has been done to address these issues and this is what the present research is fundamentally concerned with, i.e. how are multi-representational materials read and learned in one's second language (L2)? This limited attention is rather perplexing given the increasing number of people studying in their non-native language. Take English, one of the world's lingua franca, for example. In the English-speaking United Kingdom and the United States of America, language minority students constitute a large number of the national population of school pupils due to trends in migration. In fact, around one in five pupils in the UK study through the medium of English as an additional language (Department for Education, 2019). One primary school in the UK even has 39 languages spoken by its pupils (Russell, 2016). In some non-English speaking countries on the other hand, it has become a global educational trend at universities to offer more courses in

English (Lee & Mayer, 2015). Taken together, more and more people are being educated in English, their non-native language. It is not surprising therefore to find the population of non-native English speakers now surpasses the number of its native speakers ("The english effect," 2013). As the number of non-native English speakers studying in English continues to grow, it is therefore timely to gain insights into their engagement with multi-representational texts and extend current research on multi-representational learning to this population.

Moreover, simply assuming that L2 reading and learning with multi-representational materials occur in a similar manner as it is with L1 reading and learning should not be assumed given evidence that text, one of the components of multi-representational materials of interest to the present research, is read differently in the L1 and the L2. Studies on bilingual text processing for instance, found that bilinguals read less fluently in their L2 versus in their L1 (Cop, Drieghe, & Duyck, 2015). Such findings suggest that it could be possible that multi-representational reading in L1 and L2 differs as well. Hence, this research looks at the intersection between two distinct research fields, i.e. multi-representational learning and L2 reading, to examine the ways in which people read and learn with multi-representational materials in their L2. It is worth noting here that learning to read in the L2 or L2 learning in general is not the focus of the research, although these will be briefly touched upon in subsequent chapters. To address issues of interest, a research approach involving the use of eye tracking and a methodology driven by post-positivism were employed. Eye movement data of L2 English readers were collected and examined and their relationships with individual characteristics of interest as well as learning outcomes were scrutinised as people read authentic texts. These were then compared with the L1 English readers' data. Two separate studies were conducted, each dealing with a different language group. A third study compared and contrasted the two groups more directly.

Before introducing the research objective and summarising the contents of the thesis, the L1 and the L2 terms used throughout this thesis will first be explained. The term first language (L1) in this thesis, such as English as an L1 is based on the definition by Graddol (1997), where he defines L1 English speakers as "those for whom English is a first – and often only – language" (p. 10). Graddol adds that these speakers "live, for the most part, in countries in which the dominant culture is based around English". L1 in this thesis is used interchangeably with native tongue/ language. Second language (L2) English speakers on the other hand, are those for whom English is a second or an additional language (Graddol, 1997). This term also covers the term

English as an additional language (EAL) which considers English as one's third or fourth language (Hall, 2018).

Another issue worth noting here is how 'reading', 'comprehension' and 'learning' are approached in this thesis. The three indeed are difficult to distinguish. Although reading can exclusively refer to the act of decoding and comprehending *written text* (Rueda, 2011), in this thesis, reading can also be taken to mean the act of decoding picture(s). In subsequent sections, readers will encounter the phrase 'multi-representational reading' which refers to the decoding of both text and pictures. Comprehension, on the other hand, refers to the act of constructing meaning of a text, for instance, and the situation it describes (Kintsch & Van Dijk, 1978). Learning is believed to encompass both reading and comprehension, the act of decoding input and constructing a mental representation to arrive at meaning. To put it more explicitly, learning in the context of the thesis' main studies refers to the understanding of a text constructed via multi-representational reading and was taken away from readers and then measured later. Despite these terminologies being detailed as such, it is fully acknowledged that they are complex constructs and to distinguish one from the other is a difficult and complicated feat. These terminologies are further explained, wherever appropriate, in relevant sections.

1.1. PERSONAL DRIVE BEHIND THE RESEARCH AND ITS OBJECTIVE

This research originated from my experience as a TESOL (Teaching English to Speakers of Other Languages) practitioner in a Malaysian university. Five years ago, when I first started teaching first-year undergraduates the English language, I had a notion that a language classroom has to be lively, interactive and engaging. A number of teaching and learning aids containing visual and/ or audio input were used as an attempt to realise my notion of an ideal classroom. My students seemed to enjoy the lessons but not everyone learned as much as I had hoped. Little did I know at the time that enjoyment does not necessarily translate into learning. I became intrigued by how my students actually paid attention to the various input presented to them and how this is reflected in their performance. Hence, when I was asked to prepare a doctorate research proposal to the study leave committee a year into my teaching profession, I decided to focus on this intriguing issue. I chose to focus on what I believed at the time was a much simpler multi-representational material; print-based materials. The

committee approved the proposal and my study leave application. Since then, I became committed to my overall research objective: to understand the ways in which the L2 English readers read and learn from a multi-representational text in English. More specifically, I strived to understand how they decode both text and pictures (i.e. reading), and construct a mental representation out of these representations to arrive at meaning (i.e. learning).

1.2. STRUCTURE OF THE THESIS

Chapter 2: Literature Review. The literature review comprises five sections. The first section concerns *external representations and learning*. It considers the different ways in which external representations are categorised before specifying how this is done in the present research. Additionally, it considers arguments from the literature on why more than one external representation is helpful for learning and the sort of tasks learners need to engage in to benefit from multi-representational materials. The second section reviews studies on *the processing of multi-representational materials* and introduces one canonical study that is used to guide the present research. The literature on *reading in a second language* and studies that compare and contrast L1/ L2 reading are reviewed in the third section, whereas the fourth touches briefly upon *learning a second language with multiple representations*. Finally, the last section of the literature review considers the body of research that compares and contrasts learning with multiple representations in the L1 and the L2. Before this chapter is concluded, the research questions that this thesis aims to answer are specified.

Chapter 3: Methodology. The first section of the methodology chapter describes the post-positivist approach and the non-experimental fixed design employed for this research. The next section presents a detailed account of the sampling and the reading material used. Then, the methods of data collection are outlined. In particular, the eye tracking methodology and relevant measures are introduced. The next section describes the methods of data analysis. Finally, the chapter discusses how ethical considerations are addressed in the research.

Chapter 4: Study 1, i.e. the L2 study. In order to explore how multi-representational materials are read and learned in the L2, an eye-tracking study involving the participation of 50 L2 readers was conducted. In addition to uncovering the predominant reading approach employed by the participants, eye tracking analysis

revealed differences in their reading patterns. Additionally, the relationships between reading patterns and individual characteristics, as well as learning outcome are outlined.

Chapter 0: Study 2, i.e. the L1 study. This study is an initial step towards comparing and contrasting the ways in which multi-representational materials are read and learned in the L1 and the L2. 42 native English speakers or the L1 readers read the same material as the L2 readers in the previous study and with the exception of a couple of tasks (only relevant for L2 samples), went through the same procedures. Identical to the L2 study, the results provided information about the predominant reading approach employed by the participants, their different reading patterns and how these patterns are related to individual characteristics, as well as learning outcome.

Chapter 6: Study 3, i.e. the comparison study. In order to directly compare and contrast the ways in which multi-representational materials are read and learned in the L1 and the L2, this study combines and analyses the two datasets from the two previous studies. It begins by comparing the two groups in terms of individual characteristics and another variable, reading time. Next, two sets of analyses were conducted to compare their reading patterns and learning performances. Finally, the study examines differences in reading patterns irrespective of the participants' language group. Similar to the previous studies, their relationships with individual characteristics and learning are then examined.

Chapter 7: General Discussion. The concluding chapter begins with a discussion of the findings of each of the three studies. Next, it considers the research limitations as well as implications for further research. To conclude the thesis, its main contributions to the field are outlined.

2. LITERATURE REVIEW

2.1. INTRODUCTION

This chapter is divided into five sections. The first two concern the literature around learning with multiple external representations (MERs) in the L1. Specifically, Section 2.2 starts by defining and classifying external representations before detailing the use of MERs in learning. Section 2.3 reviews the body of research that examines the reading and learning processes that take place with MERs when people study MERs in their first language. From these studies, predictions about the reading and learning with MERs, relevant to the L1 samples, are derived. Attention is then turned to the L2. Section 2.4 reviews the literature on reading in an L2 to understand L2 reading and comprehension in more detail and how they differ from L1 reading and comprehension and Section 2.5 addresses learning an L2 with MERs to review whether MERs are helpful to acquire an L2. Studies that introduce MERs to facilitate the acquisition of L2 skills and competencies are reviewed to provide more insights on how MERs can be beneficial to the L2 readers. The final section, Section 2.6 reviews the small body of research that examines learning with MERs in the L2 and compares it with multi-representational learning in the L1. In addition to providing insights into multi-representational learning in the L2, more importantly, these works motivated the need for the present research. The chapter concludes by outlining the research questions that this thesis aims to answer.

2.2. EXTERNAL REPRESENTATIONS AND LEARNING

This section reviews the literature on the use of multiple external representations in learning, i.e. what might be understood as the activity of enhancing one's understanding, ideally for the long term. Before delving further into that topic, this section will first define and categorise external representations so that later comparisons can be made across various multi-representational systems. A representation can take numerous forms, be it an essay, a drawing or a song. These are examples of external representations that we can perceive with our senses. Representations can also be internal and not perceptual, something that is in the mind, like thoughts (Hegarty, 2005). Regardless of its concreteness, a representation conveys information about something or as Palmer (1978) puts it, "something that stands for something else," (p. 262).

This thesis pays particular attention to external representations, which come in a variety of forms and serve a variety of functions in current instructional materials. Similarly, there are a number of ways to describe and classify them. Some will be briefly mentioned here and the particular approach adopted in this research will be outlined. Ainsworth (2014) argues that external representation can be usefully described in terms of their functional or structural properties. Functional taxonomies such as that of Levin, Anglin, and Carney (1987) describe representations based on their intended use and purpose and therefore the content of the representations is of great interest. Structural taxonomies such as that of Lohse, Biolsi, Walker, and Rueter (1994) on the other hand group representations according to their physical structure and therefore place a greater emphasis on the form of the representations such as their modality, dynamism or abstraction. Although researchers do not have a single agreed-upon structural taxonomy, categorising representations based on their presentation mode (i.e. their modality) is arguably the most popular. This thesis also adopts this particular structural approach to classify the external representations. Since it is one of the most conventional approaches, classifying representations in this way allows the research to compare across more studies.

Using the modality approach, a given representation can either be verbal (e.g. a sentence, a narration) or non-verbal (e.g. a map, a piece of music) (Mayer, 2005b). The non-verbal representations of interest in this research are those that are closely linked to the visual modality. This research therefore categorises a given representation as either verbal or visual. Verbal representations are commonly referred to as text, words (Mayer, 2005b) and descriptive representations (Schnotz, 2002) whilst visual representations are those that have been called visualisations (Lohse et al., 1994; Scheiter, Wiebe, & Holsanova, 2009), diagrams (Larkin & Simon, 1987), and pictures (Levin et al., 1987; Mayer, 2005b) among others. To encompass the different terminologies, the thesis will refer to verbal representations as ***texts*** and visual representations as ***pictures***. Having said that, it is fully acknowledged that neither text nor pictures are unitary constructs for their structure, function or content could differ (Ainsworth, 2018; Scheiter et al., 2009). The structural and functional properties of the text and pictures used in the empirical studies will be described in greater detail in the following chapter.

Text is made up of arbitrary symbols and is related to its referent by a convention such as a natural language (Schnotz, 2002). A picture and its referent on the other

hand, are related by shared structural features. A picture also preserves the geometric and topological details of its referent (Larkin & Simon, 1987; Schnotz, 2002). Following Eitel and Scheiter (2015) and Mayer (2005b) broad classification, this thesis will consider any “visual display where space is meaningful” (Eitel & Scheiter, 2015, p. 155) as a picture, be it static or dynamic and any “information in a verbal code” as a text, be it written or spoken.

Having categorised external representations into two broad categories, i.e. texts and pictures, it becomes necessary to understand why there is a need for multiple external representations rather than a single one. Before delving further into that topic, it is crucial to first understand reading and (text) comprehension processes. After all, the multi-representational materials used in the research presented in this thesis comprises of a substantial amount of text, in addition to pictures.

In this thesis, comprehension is considered as the act of constructing meaning and text comprehension involves constructing a mental representation of the text and the situation it describes. Once this representation is constructed, various computations with regard to the text such as recall, summarising and question-answering can be performed. This concept of comprehension is derived from one of the influential models of text comprehension, that of Kintsch and Van Dijk (1978), which argues for three levels of representation during comprehension. These are the surface structure, the text base and the situation model. The surface representation is the verbatim, linguistic layer that preserves the wording and syntax of the text. The text base, on the other hand, is the conceptual or abstract layer that represents the semantic meaning of the text, and no longer tied verbatim to its surface form. At this level, linguistic information is extracted into larger linguistic units to construct relevant meaning in the form of propositions. These first two levels represent what the text is, “the words and relations among the words explicitly stated within the text,” (O'Brien & Cook, 2015, p. 217) whilst the third, the situation model, goes beyond and represents the verbally described situation or what the text is about. At this level, readers’ background knowledge and inferencing support and expand the text base representation, and the information content of the text become integrated into a much larger structure (van Dijk & Kintsch, 1983).

The construction of multiple levels of representation involves multiple component processes and among them are word recognition, parsing, semantic-syntactic representations, text representation and understanding (Tokowicz & Perfetti, 2005).

The processes that take place at the surface and text base representational levels (such as word recognition, parsing and semantic interpretation) are what some authors term low-level processes (Grabe, 2009; Tokowicz & Perfetti, 2005). For competent native speakers of a language, these processes are automatic and do not involve attentional control in their execution (Segalowitz & Hulstijn, 2005). They are therefore rapid, effortless and unconscious. The situation model involves high-level processes, such as synthesising the textual information with background knowledge which are more closely linked to comprehension strategies and are more introspective and attention-based (Grabe, 2009). Some of the key component processes will be briefly introduced here.

Visual word recognition, according to Dijkstra (2005), refers to the process of retrieving information about a word in a language user's mental lexicon based on a presentation of a visual stimulus. This lexicon is likened to a database in which all of the words that a language user knows in a particular language are stored. The information contained in this database includes its sound (phonology), spelling (orthography), and meaning (concepts/ semantics) among others (Dijkstra, 2005; Dijkstra & van Heuven, 2018). The processes underlying visual word recognition are typically studied through the influences of orthographic, phonological, or semantic similarity such as orthographic neighbours (i.e. words that differ from a target word in only one letter position).

In addition to recognising individual words, a reader must also deal with sentences. Sentence processing necessitates examination of the grammatical structure of sentences and the placing of each word in this structure (Staub, 2015). Apart from this syntactic parsing, semantic interpretation is also necessary during sentence processing. During semantic interpretation, the meaning of each word and phrase is combined to construct the overall meaning of the sentence. This then determines "a sentence's thematic roles (who did what to whom)" (Macdonald & Hsiao, 2018, p.171).

Now that a general albeit relatively simple overview of reading and (text) comprehension has been presented, this section will now move to considering why there is a need for multiple external representations rather than a single one, as learning (i.e. the act of decoding input and constructing a mental representation out of it to arrive at, ideally lasting, meaning) with MERs is the focus of this research. Much earlier research has shown that learning with MERs is (often) better than with

one representation alone. However, with the current learning environments becoming increasingly multi-representational, the single versus multiple debate has now become rather obsolete. A more practical and important issue is how to learn effectively in these environments. As mentioned in Chapter 1, this thesis is particularly concerned with the ways in which learners read and learn with MERs, particularly with conventional textbooks comprising written texts and static pictures (rather than e-textbooks). Though the literature abounds with benefits from learning with such materials, it is argued herein that it is important to go beyond the structural properties of the representations and examine their functional properties.

There are several reasons why external representations are used in combination. Levin et al. (1987) for instance identify five reasons why pictures are used as text adjuncts; to decorate, represent, organise, interpret or transform the text content. Decorational pictures are visual representations that have little to no relationship to the text content and merely serve to decorate the page. The remaining four on the other hand, have more pedagogical functions. Representational pictures for instance make the text more concrete to the readers as they depict part or the entire text content. To make texts more coherent, organisational pictures arrange text contents into a useful, structural framework. Interpretational pictures help readers understand text content by clarifying difficult texts whilst transformative pictures facilitate recall of text content as they contain mnemonic elements. This taxonomy therefore identifies the purpose of having a picture as an adjunct to the more dominant text. Multi-representational materials however are not limited to a chunk of text that is accompanied by a solitary picture. Domain-specific materials for instance, may comprise multiple visual representations, hence Levin et al.'s taxonomy may be less applicable to such materials.

Ainsworth's (1999, 2006) functional taxonomy on the other hand is more encompassing. It specifies how MERs are beneficial for learning and problem-solving according to three main functions; by complementing, constraining and constructing. Two or more representations are complementary when they each convey different information, or when they both convey the same information but present it differently. With complementary representations, the learners benefit from the unique properties of each representation. A constraining representation on the other hand functions to constrain the interpretation of another that is either ambiguous or less familiar. In both instances, the more specific or the more familiar representation provides support for learners to understand and not to misinterpret the other that they know less about.

The final constructing function helps learners construct a deeper understanding of the subject matter. This can materialise when learners integrate information from the different representations. Here, multiple representations are necessary as such achievement would be difficult by exploiting only one representation. A more comprehensive discussion of this taxonomy is provided in the subsequent chapter.

The functional taxonomies help justify why instructional materials containing combinations of pedagogically-functional representations are helpful and valuable for learning. However, the literature also suggests that benefits from MERs cannot be materialised if the learners do not understand how the representations and how they relate to each other (Ainsworth, 2006; van der Meij & de Jong, 2006). The concept of integrating information from these different sources has been emphasised and returned to frequently in the literature, signalling how crucial integration is when learning with MERs. Different terminologies have been used when discussing this concept, e.g. relating/ coordinating representations, translating between representations, building referential links/ correspondences between representations and conjoint processing of representations. Other authors use a different label like coherence formation (Seufert, 2003) or text-picture integration (Schnotz, Wagner, Ullrich, Horz, & McElvany, 2017). Although the exact terminology differs, most authors agree that integration involves understanding the individual representations whereby learners make sense of their forms, operators, elements and relations between the elements and also how the corresponding elements between representations are found and related (Ainsworth, 2006; Seufert, 2003). Although the term ‘correspond’ implies similarities, integration also involves identifying discrepancies, in addition to, regularities between representations (Ainsworth, 2006; van der Meij & de Jong, 2006). Thus, it could be that learners understand each representation independent of one another and then relate them or instead learners may come to understand how a representation works by relating it to a better understood one. However, irrespective of the exact mechanism, it is argued that by integrating the MERs, learners can construct a coherent mental representation of the learning material, thereby fostering in them a deeper understanding of the subject matter (Ainsworth, 2006; Mason, Tornatora, et al., 2013; Mayer, 2014).

Although integrating multiple representations is often needed for learning with MERs, the literature has indicated that the task is not exactly easy or automatic and is arguably the most challenging task in learning with multi-representational materials (Ainsworth, 2014; Schnotz et al., 2017). Indeed, learning with multiple representations

can be overwhelming since it means that learners face additional and complex cognitive tasks compared to learning with one representation. According to Ainsworth (2006), when presented with a novel representation, learners are already dealing with complex cognitive tasks as they need to understand how the representation encodes and presents information and how it is related to the domain it represents. If need be, they have to choose an appropriate representation or construct their own. These tasks continue to grow as the number of representations increases. As the primary focus of this thesis is the L2 readers, it seems sensible to assume that these cognitive tasks on top of reading and processing materials in an L2 could easily be overwhelming for readers. Nevertheless, sometimes multi-representational environments are designed to help readers acquire an L2 (such as pictures books) and clearly many L2 readers must read multi-representational materials that have not been designed with their needs in place. Hence, this issue is returned to in Section 2.4 respectively.

2.3. READING MULTI-REPRESENTATIONAL TEXTS IN THE L1

One of the main concerns in multi-representational learning highlighted in the previous section is that although integrating MERs can be crucial to learn successfully from multi-representational materials, this task is cognitively demanding. This section further elaborates on this issue by reviewing research that examined how learners actually processed and learned with these materials. Note that the studies reviewed here concern those who are reading and learning in their L1. Evidence from these studies is subsequently used to derive predictions about the reading and learning with MERs, especially concerning the L1 samples. With regard to the L2 readers', the following three sections will provide analysis and insights that led to the formulation of relevant hypotheses.

Research on the processing of multi-representational materials examines the cognitive processes that learners undergo when interacting with these materials to understand the underlying reasons behind successful and unsuccessful learning. This focus on processes warrants the use of methodologies more commonly associated with cognitive psychology such as eye tracking and brain imaging. Since the present research is particularly interested in eye movements, research that implemented the eye tracking methodology will hence be reviewed.

Eye tracking research has consistently found that the processing of multi-representational materials is text-dominant. Learners predominantly attend to text areas and sometimes fail to devote a sufficient amount of attention to inspecting picture areas. This rather ineffective use of pictures was demonstrated in Hannus and Hyönä (1999) study on young Finnish learners who generally spent only six percent (around 16 to 20 seconds) of total study time on the pictures of authentic Biology textbook passages. In another study and across two experiments, native German speakers who learned with an animation accompanied with written text spent more time attending to text than on the animation (Schmidt-Weigand et al., 2010). In Experiment 1 for instance, they spent 112.4 seconds reading the text compared to only 65.7 seconds on the animation. In Experiment 2 which was self-paced, their tendency to process the text more intensively led to longer learning times. In this subsequent experiment, the text was attended to three times more than the animation (131.20 versus 41.90 seconds). Similar findings have been found in many other studies (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Johnson & Mayer, 2012; Ozcelik, Arslan-Ari, & Cagiltay, 2010; Rayner, Rotello, Stewart, Keir, & Duffy, 2001; Schwonke, Berthold, & Renkl, 2009).

The text-dominant processing found in the above-mentioned literature suggests that there is a mismatch between what is ideal and important in learning, i.e. integration, and what is actually going on during learning. This is worrying as the different representations need to be processed adequately and integrated in order for the learners to exploit advantages of MERs and to subsequently comprehend the materials successfully (Scheiter, Schüler, & Eitel, 2017). Indeed, this text-dominant tendency has been identified as one of the obstacles that may limit the efficacy of multi-representational materials (Scheiter et al., 2017). Several reasons have been proposed as to why this tendency is prevalent. The first concerns the redundancy of information conveyed by the MERs. When different representations offer learners redundant information, they may choose to focus on a single representation rather than all (Schwonke et al., 2009) and as is often the case, the visual information is usually ignored or processed less sufficiently than the verbal information (Arndt, Schüler, & Scheiter, 2015). This tendency is evident from the studies reviewed above. They show that readers are more likely to focus only on one representation or a subset of the available representations, and it is argued that this representation is usually the one that is more familiar to them (Arndt et al., 2015).

Apart from that, the processing of visual as opposed to verbal information can be more rapid and easier thereby giving readers an illusion of fully understanding the pictures (Shroeder et al., 2011; Scheiter et al., 2017). This can in turn lead to predominant processing of texts since verbal information is processed less rapidly. It could also be possible that readers process pictures only superficially because they do not have the relevant strategies to effectively process graphics (Hochpöchler et al., 2013). Acknowledging this tendency and the fact that text-picture integration is not exactly trivial, a number of principles and techniques such as spatial and temporal contiguity (Mayer, 2005a), signalling (Mayer, 2005a) and implementation intention (Stalbovs, Scheiter, & Gerjets, 2015) have been proposed to support a more strategic processing of multi-representational materials. Though these supports are beyond the scope of the present research, they signal the fundamental role of integrating information from the MERs in learning.

Accordingly, research has found that the lack of integration across MERs is associated with poor learning outcomes. Mason and colleagues in a series of eye tracking studies with Italian school children have consistently found positive correlations between integrative processing and learning outcomes. In one of these studies, the children were found to display three levels of integrative processing when reading an illustrated text in their native tongue; low, intermediate or high (Mason, Tornatora, et al., 2013). Those who displayed the lowest level of text-picture integration performed worse than the intermediate and/ or the high integrators in several learning tasks. The high integrators' reading behaviour on the other hand, was associated with the highest performance. These findings were replicated in a more recent study (Mason, Tornatora, et al., 2015) and more importantly, the degree of text-picture integration was found to predict learning outcomes. The more frequently and the longer readers integrated text and pictures during re-reading, the better they performed in the learning tasks. This positive relationship between integration and learning outcomes has also been observed in other studies (Hannus & Hyönä, 1999; Jian & Wu, 2015; Johnson & Mayer, 2012; O'Keefe, Letourneau, Homer, Schwartz, & Plass, 2014).

Despite the seemingly universal tendency to process multi-representational materials in a text-dominant fashion, differences in the processing patterns and strategies can still be observed. In Mason's studies for instance, the learners were grouped according to how much they integrated (or did not integrate) the MERs. Hence, readers can be low/ intermediate/ high integrators (Mason, Tornatora, et al., 2013) or

weak/ strong integrators (Mason, Tornatora, et al., 2015). The learners can also be grouped according to their referencing strategies; text-diagram, diagram-first, text-first or text-only (Jian & Wu, 2015). Unlike Mason, Tornatora, et al. who used cluster analyses to group readers, Jian and Wu's grouping method was more subjective. Two raters judged the participants' eye movements and grouped them based on whether they made many transitions between the representations (text-diagram), inspected the diagram or read a couple of sentences before "the first proper noun… and went back to the text section" (diagram-first), read almost the entire text before inspecting the diagram (text-first) or read only the text without inspecting the diagram at all (text-only). Despite these different grouping methods, when the different groups' learning performances were compared, the importance of integrating MERs for learning is reiterated. The low/ weak integrators (Mason, Tornatora, et al. 2013) and the text-only readers (Jian & Wu, 2015) who predominantly attended to the text or neglected to attend to the pictures, were found to perform worse than those who attended to both representations more extensively (the high/ strong integrators and the text-diagram, diagram-first, text-first readers).

Mason, Tornatora, et al.'s (2013) seminal study is largely modelled to guide the present examination into the processing of multi-representational materials and its association with learning. The study found a significant link between these patterns and learning and in a more recent study (Mason, Tornatora, et al., 2015), these findings were replicated. Considering these successful applications, the present research therefore used the study and its measures as a model to analyse the reading behaviour of different populations of readers (i.e. readers of English as an L1 and an L2) when learning from a similar (i.e. scientific text and picture) but different material, i.e. an authentic material with multiple graphic representations. Although Mason, Tornatora, et al.'s (2013) findings have been reviewed earlier, the overall study will be described in greater detail here as it informs specific methods used in the studies conducted for this thesis. Similar to the present research, Mason, Tornatora, et al.'s (2013) research centred on the processing of a multi-representational material and its link with learning. They did however focus on younger participants, i.e. Grade 4 pupils. The authors used fine-grained measures of total fixation duration, i.e. the measures of first-pass and second-pass reading, proposing that the former reflects the children's initial processing of the material whilst the latter reflects more intentional and strategic processing. The integrative processing was analysed using the second-pass reading measures or the look-from fixation times which total up how long the children spent on this processing. Another measure, the integrative transition, reflects the number of

integrative attempts. All these measures are subsequently adopted in the present research and further discussion on them is provided in the next chapter.

Mason, Tornatora, et al. (2013) were particularly interested in the varying reading patterns that the children might display during the learning episode, predicting that more or less strategic patterns would emerge. They postulated that a more strategic reading pattern would be characterised by a longer inspection of the picture during first-pass reading, and greater text-picture integrative processing, as displayed by more frequent integrative transitions and longer look-from fixation times between the text and picture. Additionally, a longer look-from fixation time between different text areas could be a characteristic of a strategic reading pattern. The children were asked to read an illustrated passage in their native language, Italian, consisting of a short text and a structure diagram on the topic of air and its characteristics. Before this reading stage (Session 2) commenced, the students undertook prior knowledge, reading comprehension ability and spatial ability tests (Session 1) to examine each of these individual characteristics' link with reading behaviour. Immediately after the reading stage in Session 2, the immediate post-tests were administered. In Session 3, a month later, the same tests were administered. As reviewed above, three reading patterns (low, intermediate and high) emerged and they were differentiated by the text-picture integrative processing. Two other measures that differentiated these patterns were the first-pass fixation time on picture and look-from text to text fixation time. Two of the individual characteristics; reading comprehension ability and prior knowledge were found to significantly correlate with the eye tracking measures. In particular, the higher the reading comprehension ability, the shorter the first-pass fixation time on text was. As for prior knowledge, the higher it was, the longer and the more frequent text-picture integrative processing was. After all individual characteristics were controlled for, the relationship between processing and learning was examined and as predicted, the high integrators who displayed the most strategic reading pattern performed best.

The influence of individual characteristics has been explored in previous work to address whether or not multi-representational learning works for everyone in the same way. The three examined by Mason, Tornatora, et al. (2013); prior knowledge, reading comprehension ability and spatial ability, arguably play the most important role in multi-representational learning and therefore will be discussed in greater detail. Prior knowledge in this research refers to pre-existing or previously acquired knowledge within a particular domain. This knowledge serves as a resource during

text comprehension as it can compensate for information that is not explicitly conveyed in text (Kintsch, 1998; Ozuru, Dempsey, & McNamara, 2009), and provide readers with a framework in which to organise and integrate newly learned information (Chi et al., 1981 in Butcher, 2005). Empirical support for the contributing role of prior knowledge in text comprehension is suggested by Ozuru et al. (2009). They found prior knowledge as one of the predictors of performance on different types of comprehension questions, i.e. text-based, local-bridging and global-bridging. Moreover, they found that when the comprehension task necessitated more extensive integration of information, the contribution of prior knowledge increased. Mason, Tornatora, et al.'s (2015) study also found a similar role of prior knowledge during the re-processing of an illustrated science text in terms of predicting performance on text-based factual knowledge. Additionally, its predictive role on verbal recall performance was only significant when combined with other individual characteristics (namely reading comprehension skill and self-concept). With regard to its role in the processing of multi-representational materials, Mason, Tornatora, et al. (2013) found that readers with low prior knowledge processed an illustrated science text 'less strategically' by engaging in a less integrative processing within and across representations. They made fewer or no attempts at integrating text and picture and spent less time integrating them compared to higher prior knowledge readers. It could be the case that readers with higher prior knowledge used their pre-existing knowledge to integrate these representations (Scheiter, Schüler, Gerjets, Huk, & Hesse, 2014). Taken together, prior knowledge has been shown to be a significant predictor not only of learning outcomes but also processing of multi-representational materials. It will hence be considered as one of the individual characteristics of interest in the thesis and its relationship with the processing of an authentic multi-representational material will be examined.

Reading comprehension involves a wide range of cognitive activities such as word-decoding, understanding and recalling ideas, identifying important information and synthesising them, and inference generation that are activated at the lower and higher levels of information processing (McNamara & Magliano, 2009; Sinatra & Broughton, 2011). Reading comprehension ability helps readers to accomplish these activities and consequently form an integrated understanding of the text content (Ozuru et al., 2009). This ability is particularly important in the present research as the reading material requires participants to read and process a substantial amount of written text. Additionally, this ability is believed to facilitate the comprehension of images indirectly. To illustrate, better readers can comprehend text segments better and this in turn

guides their attention to the corresponding picture(s) (Hegarty & Just, 1993; Hochpöchler et al., 2013; Scheiter et al., 2014). It is therefore argued that possessing good comprehension skills is necessary for readers to take full advantage of complex visualisations (Scheiter et al., 2014). Thus, reading comprehension could explain what readers come to understand from both text and pictures. This ability is also predictive of text comprehension. Ozuru et al. (2009) found reading skills to predict performance on text-based and local-bridging comprehension questions. Surprisingly, the contribution of reading skill decreased (rather than increased) when the comprehension task necessitated more extensive integration of information. Similarly, Mason, Tornatora, et al. (2015) found this ability to predict performance on different types of learning outcomes. More specifically, during re-processing of an illustrated science text, reading comprehension ability predicted performances on graphic recall, factual knowledge and transfer of knowledge. Its combination with prior knowledge and self-concept was also predictive of verbal recall. This ability could also explain how readers read multi-representational texts. Unlike prior knowledge which is linked to text-picture integration during the re-reading of text, this ability is linked to first encounter or first-pass reading of the text. Rayner, Chace, Slattery, and Ashby (2006) believe that comprehension difficulties in text reading are reflected in longer first-pass fixation time on the text. Indeed, Mason, Tornatora, et al. (2013) found a negative correlation between the two. Less skilled readers spent more time reading the text elements of the illustrative text when they first read it compared to the more skilled readers. Given the significant role of reading comprehension ability in the processing and comprehension of both text and pictures, it will accordingly be considered as one of the individual characteristics of interest in this thesis.

The final individual characteristic, spatial ability, concerns how people deal with spatially-presented materials and is related to their “abilities in searching the visual field, apprehending the forms, shapes, and positions of objects as visually perceived, forming mental representations of those forms, shapes, and positions, and manipulating such representations "mentally."” (Carrol, 1993, p. 304). This ability has been argued to play a substantial role in fostering learning with pictures and in his meta-analysis, Höffler (2010) found a medium-sized effect (r = .34), suggesting the superiority of high over low spatial ability learners. Mason, Tornatora, et al. (2013) however did not find any significant correlations between this ability and any of their eye tracking measures. The authors’ explanation of this finding is due to the nature of their material which was relatively simple and less demanding in terms of mental manipulations. Although Höffler and Leutner’s (2011) study, for instance,

demonstrated the importance of this ability in learning from non-dynamic visualisations, the nature of the material (about surfactants and their role during dirt removal process) was much more complex and demanding. Even the authors acknowledged that the "topic might be far more difficult to be mentally animated – firstly, it is not observable in every-day environments; secondly, the inference of motion … of abstract surfactants can be expected to be harder to draw …" (p. 214). Hence, although spatial ability is indeed important to consider when dealing with visual representations, the spatial relations conveyed in the material to be read by the present research participants are relatively simple and not demanding. Inferences about distance and movements conveyed in the passage are expected to be much easier to draw than those of Höffler and Leutner's. Accordingly, this thesis will not consider spatial ability as an individual characteristic of interest.

From the literature reviewed in this section, a number of predictions regarding the L1 readers' reading behaviour and learning with the authentic multi-representational material used in the present research can already be formulated. These however will be outlined in the final section of this chapter, whereby any predicted differences with L2 readers can be outlined. With regard to the L2 readers' reading behaviour and learning the following three sections will provide more specific literature whose analysis led towards the formulation of relevant hypotheses.

2.4. READING IN A SECOND LANGUAGE

This section outlines how reading and comprehension take place in an L2 and how they differ from reading in an L1. This thesis particularly focuses on reading and comprehension in English as an L1 and an L2. Section 2.2 has presented an overview of reading and comprehension in general, as well as their constituent components. How these constituent components proceed during L2 reading are then detailed followed by predictors of L2 reading. Finally, some eye tracking studies that identify the differences in the ways L1 and L2 reading take place are reviewed. Overall, the research reviewed in this section helps to derive initial predictions with regard to the ways in which L1 and L2 readers differ when processing text or verbal representations.

2.4.1. Reading in the L2

Section 2.2 has presented a general overview of reading and comprehension. The components briefly introduced there provide a framework to study how L2 reading and comprehension take place. Although L2 reading does not necessarily include additional components, there are aspects that distinguish it from L1 reading. This section discusses some of these aspects.

When reading in an L2, readers are limited in terms of their linguistic knowledge (e.g. vocabulary and grammar), exposure to L2 reading materials, and social and cultural knowledge related to the L2 community (Grabe, 2009). Additionally, having a prior literacy or linguistic experience in the L1 necessarily implicates L2 reading development and experience by transferring its effects on the L2 reading or by imposing a certain worldview or background knowledge on the reader. Prior literacy should not be seen as necessarily detrimental to L2 reading. In fact, L2 learners can draw upon their prior literacy to facilitate their learning of the L2 (Bernhardt, 2005; Koda, 2004; Verhoeven, 2011). Interlingual or cross-linguistic interactions are one of the key issues examined in bilingual and second language research. For the former, such issues provide insight into how the two languages of a bilingual are stored and processed. Research in these domains examines interlingual interactions at various representational levels and what has become clear through extensive research is that the two languages of a bilingual do not operate in isolation but that while reading L2 words, L1 words can be activated and influence the recognition processes in the L2 (Jared, 2015). Similarly, while reading L1 words, L2 words can be activated and influence L1 word recognition. The following paragraphs will particularly focus on the interlingual interactions involving the influence of L1 on L2 processing and review some evidence at the word and sentence levels.

At the word level, examination of L1 influence during L2 word recognition typically involves presentations of interlingually ambiguous words or words that have some overlap between a bilingual's two languages. Such words can be cognates (translation equivalents that share orthographic and/ or phonological forms), false friends which include interlingual homographs (words that share orthographic form but not meaning) and interlingual homophones (words that share phonological form but not meaning), or neighbours (for a review, see Dijkstra and van Heuven (2018)). These ambiguous words are used as differences in performance on such words versus control words that exist only in the target language may indicate the activation

of representations in the non-target language or a parallel activation at a lexical level (Jared, 2015). Indeed, interlingual interactions have been observed as both facilitatory and inhibitory effects emerge when bilinguals process these words. Cognate facilitation effects, for instance, are observed among bilinguals whereby they respond faster and more accurately to cognates compared to matched non-cognates (Dijkstra, Grainger, & van Heuven, 1999; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011). Ambiguous words may also have inhibitory effects such as what has been observed by Dijkstra et al. (1999) in the case of interlingual homophones, which were responded to more slowly than control words were. In the case of orthographic neighbours, van Heuven, Dijkstra, and Grainger (1998) for instance found both facilitatory and inhibitory effects of within and between-language neighbours. Interlingual interactions have also emerged when these words are presented in sentence contexts rather than in isolation as done in the previously mentioned studies. Cognate facilitation effect, for instance emerged during bilingual sentence processing whereby cognates, as opposed to control words, in sentences were read faster. These were observed in both early and later stages of word recognition (Duyck et al., 2007; Van Assche et al., 2011). Importantly, when monolinguals were presented with the same stimuli, such effects were absent (Dijkstra et al., 1999; Van Assche et al., 2011; van Heuven et al., 1998), indicating that the effects should be attributed to the unique characteristics of the bilingual processing. Thus, these findings indicate that during word recognition, the presentation of a word in one of the languages activates the representations in the bilingual's two languages (rather than just the target language's representations).

L1 influence on L2 processing can also take place at a sentence level. In contrast to research on bilingual word recognition, research on bilingual sentence processing focuses on how semantic or syntactic representations are constructed. Evidence comes from studies using two different stimuli; sentences with morphosyntactic violations and sentences with ambiguous morphosyntactic structures (Jared, 2015). In the former, L1 and L2 sentences can either share similar morphosyntactic constructions, different morphosyntactic constructions, or one language can have a unique morphosyntactic construction not available in the other. A review of research on morphosyntactic processing by Tolentino and Tokowicz (2011) found that interlingually different or unique constructions are associated with differential L2 processing whereas interlingually similar constructions show no differences between the L1 and the L2. Overall, knowledge of L1 morphosyntactic constructions can influence a bilingual's ability to identify L2 morphosyntactic violations. In the second

type of stimulus, correct yet ambiguous sentences are used, enabling examinations into how parsing preferences in one language implicates the comprehension of another. Beginning bilinguals, for instance, have been observed to show an L1 parsing preference (i.e. by displaying shorter gaze durations on disambiguating verbs or subjects) when disambiguating ambiguous sentences in their L2. On the other hand, more proficient bilinguals display a preference adopted by natives of that language when performing the same task. Thus, it seems that proficiency level can influence bilingual sentence processing. The paragraphs above review some of the ways in which having a prior literacy or linguistic experience in the L1 can affect L2 text processing. Having said that, it is necessary to point out that one's knowledge in the L2 may also have implications for L1 reading. As previously mentioned, the two languages operate simultaneously, not in isolation. An account of how these two sources are considered to operate during L2 reading is detailed below.

In her compensatory model of L2 reading, Bernhardt (2005) argues that the knowledge sources (i.e. L1 literacy and L2 knowledge) that an L2 reader has should be seen as operating in a synchronised, interactive and synergetic fashion during the reading process. These sources are not additive, rather they take over or assist those which are absent or inadequate. Having a high level of L1 literacy, for instance, means that it is more available to assist a beginning L2 reader with impoverished L2 knowledge in her L2 reading processes. To illustrate, the reader could use her knowledge of connections between oral language and print in her L1, to understand connected discourse in her L2 (Bernhardt, 1994). Bernhardt's model (2005) which synthesises research and thoughts in the area stresses the importance of L1 literacy such as knowledge of alphabets, vocabulary and text structure in addition to L2 language knowledge. The model identifies both factors as predicting half of the variance in L2 reading, while the remaining half is attributed to unexplained variance that could range from affective factors such as motivation, interest and engagement to domain knowledge and comprehension strategies. McNeil's (2012) extended compensatory model offers strategic knowledge and background knowledge to account for some of the unexplained variances. Strategic knowledge is essential in reading comprehension as it allows readers to identify which of their knowledge sources is inadequate and in need of support from other sources. It also allows readers to identify and evaluate alternative sources to compensate for the inadequacies. This ability therefore entails readers' conscious cognitive and metacognitive reading comprehension strategies. This model considers background knowledge as content knowledge that includes readers' domain, topical or cultural

knowledge. Although it facilitates reading comprehension at both low and high levels of processing, it is more involved in the latter, more strategic processing in which it is integrated with the text-based form to construct the situation model.

2.4.2. Predictors of L2 Reading

Based on the previous subsection, several factors that predict L2 reading can be identified. These are L1 literacy, L2 knowledge or proficiency, domain or prior knowledge and comprehension strategies. The research in this thesis primarily focuses on L2 knowledge or proficiency and prior knowledge as L2 reading predictors and so these two are discussed in turn in the following paragraphs.

Knowledge of the L2 constitutes an array of components. Bernhardt's model above, for instance, conceptualises it as comprising of vocabulary knowledge, grammatical form, L1-L2 linguistic distance and the impact of cognates, among others. Vocabulary knowledge will be discussed in this section. Vocabulary is often seen as a building block towards acquiring proficiency in a target language (Ellis, 2005; Koda, 2004). It plays a pivotal role in many aspects of the language such as reading comprehension, communicating and understanding of the language as a whole (de Groot & van Hell, 2005; Ellis, 2005; Koda, 2004). With regard to reading comprehension, vocabulary knowledge compared to other factors has consistently been shown to be a strong predictor (Grabe, 2009; Koda, 2004). In order to understand a reading material, one must know a certain amount of text vocabulary. According to Nation (2006), for unassisted comprehension in English, a vocabulary of 8000 word families provides 98 percent of the coverage to deal with written texts. A word family is made up of a base word (e.g. work) and its derived and inflected forms (e.g. works, worked, and working) (Bauer & Nation, 1993). If the ideal coverage is 95 percent, a vocabulary of 3000 word families is needed (Nation, 2006). In addition to vocabulary size (or quantity), a reader also needs to possess a variety of knowledge aspects regarding a word (i.e. quality of vocabulary knowledge) (Schmitt, 2008). These include the item's forms (such as its spoken and written forms), meanings, grammatical functions and collocations (Nation, 2001). Nativelike vocabulary skills are required to read and comprehend texts fluently and successfully. These skills entail fast, automatic access to words in the mental lexicon so that adequate attentional capacity can be spent on higher-level processes (de Groot & van Hell, 2005). Without these skills, much of the attentional capacity will be spent on understanding the individual words and their

meaning, leaving inadequate resources for other processes (de Groot & van Hell, 2005; Pérez, Hansen, & Bajo, 2018).

The quantity and quality of vocabulary knowledge in L1 and L2 learners may differ, given proficiency differences between the two. In terms of the quantity, as to be expected, L1 versus L2 learners' vocabulary estimates are larger (Droop & Verhoeven, 2003; Grabe, 2014). In terms of the quality, it has been noted that the L2 learners compared to their L1 counterparts have fewer associative links between words leading to much weaker semantic networks of their L2 (Droop & Verhoeven, 2003). Both limitations may have serious consequences on L2 reading. Late unbalanced bilinguals (i.e. bilinguals who acquire their L2 in late childhood or in a classroom setting without having attained a nativelike proficiency (Cop et al., 2015; Vega-Mendoza, West, Sorace, & Bak, 2015)) have also been observed to display slower lexical processing when reading in their L2 versus in their L1 (Cop et al., 2015). This was gathered from their eye movement patterns; more and longer fixations, shorter saccades or shifts between fixations and less skipping of words, resulting in longer reading time. All this is indicative of a less fluent reading, which the authors likened to children who are still developing reading skills in their L1.

Koda (2004) speculates that the relationship between vocabulary knowledge and reading comprehension is more complex than what the 'vocabulary-supports-comprehension' view suggests. During their development, vocabulary knowledge and reading comprehension may be mutually interdependent and have a two-way reciprocal relationship. It is indeed true that word knowledge is crucial to understanding texts, but it is also true that with a more global understanding of the text and the context in which a word appears, its more precise meaning can be determined.

The second predictor of L2 reading that will be focused on is prior knowledge. As previously mentioned, prior knowledge is essential during text comprehension, especially in the construction of a situation model. During text comprehension, this knowledge serves as a resource that helps readers build a coherent mental model of the text content. Prior knowledge is especially crucial when the information that is explicitly conveyed is inadequate as it can compensate for this lack of information (Kintsch, 1998; Ozuru et al., 2009). Koda (2004) notes that it is particularly useful when readers have insufficient linguistic knowledge in the L2. She argues that the construct predicts text comprehension better than L2 proficiency does and high level

of prior knowledge can offset low L2 proficiency. Indeed, many studies have noted the benefits that L2 readers gained from possessing this knowledge. Whether this construct is conceptualised as familiarity with culture represented in the text, discipline-related knowledge, or general topic knowledge, positive impacts on reading and comprehension performances have been observed (for a review, see McNeil (2011)). However, contrary to Koda's (2004) argument, McNeil's (2011) synthesis of relevant studies concludes that the influence of prior knowledge can be limited and confined by knowledge of the L2, especially in the case of readers with low and high L2 knowledge. Since the role of prior knowledge in predicting L2 reading has been contested, its role in processing multi-representational materials will be considered in this thesis (see also Section 2.3).

2.4.3. Eye Movements as Indicators of Reading Differences between the L1 and L2

The reading process and its efficiency can be examined using online processing measures such as eye tracking which records the eye gaze and movements during reading, providing a detailed account of its underlying cognitive processes (Rayner, 2009). This section reviews eye tracking studies that compared eye tracking data of native versus non-native speakers of a language, monolinguals versus bilinguals, as well as data of bilinguals reading in their L1 and L2 to formulate hypotheses with regard to how text is read in the L1 and the L2. A meta-analysis of eye tracking studies have established that during silent sentence reading, skilled readers of alphabetic writing systems like English make an average fixation duration at 225 to 250 milliseconds, an average saccade length at 7 to 9 letter spaces, and a regression rate of 10 to 15% (Rayner, 2009). Some deviations however may occur during the reading of a continuous text as illustrated by Cop et al.'s (2015) study. This study recorded that when reading an English novel, English monolinguals made an average fixation duration of 216 milliseconds, an average saccade length at 9.9 letter spaces, and a regression rate of 24.2%. Reading skill however, influences some aspects of eye movements such that slower readers like beginning, less skilled and dyslexic readers make longer fixations, shorter saccades and more regressions during reading (Ashby, Rayner, & Clifton, 2005; Cop et al., 2015; Rayner, 1998).

Despite propositions suggesting the differences in the ways L1 and L2 reading proceed, not much research has been conducted to examine how global text processing relates or implicates comprehension. Oller and Tullius's (1973) study

provides one of earlier accounts on global text processing between native and non-native speakers. The authors compared the eye movements of non-native and native English speakers when reading a standardised text for advanced university level students. Overall, native speakers' reading behaviour was more fluent as they made less and shorter fixations, and had faster reading rate compared to non-native speakers. The two groups however did not differ in the number of regressions or backward saccades to previously attended areas. The authors also correlated the non-natives' eye movements with their performance on a separate English proficiency test. A multiple regression analysis revealed moderate negative correlations between the test scores and several eye tracking measures, i.e. number of fixations and regressions. Moderate positive relationships were found for word span and reading rate. The relationship between the test scores and fixation duration was very low.

Cop et al.'s (2015) study described earlier provides bilinguals' sentence-level processing information. The bilinguals demonstrated differences when reading a complete novel in their two languages (Dutch as the L1, English as the L2). L2 sentence reading times were slower due to a higher number of fixations that were longer and closer together and a lower word-skipping rate. The authors likened this inefficiency to that of children who are still acquiring L1 reading skills. These findings as indicated by the authors suggest that reading in one's weaker language results in a disadvantage and it is plausible to assume that this disadvantage can escalate with a longer piece of text. This is because the bilinguals' individual fixations in L2 versus in L1 were longer. As stated earlier in the section, the authors attributed this inefficiency on slower lexical processing in the L2, given that it was acquired after the acquisition of L1. Even though the comprehension performance was not the main focus of the study, it is nevertheless interesting to note that the bilinguals performed similarly well on both the L1 and L2 multiple-choice items. The authors also collected English monolinguals' reading patterns, even though these were not compared with the bilinguals' English reading. Using the means and *SD*s provided, the effect sizes were calculated for each eye tracking measure. All effect sizes were very small. The differences in English reading patterns between the monolinguals and bilinguals were therefore trivial, even though the bilinguals were reading in their weaker language. The two groups' comprehension scores also did not indicate much difference.

Other studies have also revealed less fluent reading patterns when bilinguals were reading in their L2 compared to their L1. Whitford and Titone (2015) who examined sentence-level processing of L1-dominant English-French and French-English

bilinguals for instance found that the bilinguals exhibited significantly reduced reading fluency when they were reading in their L2. Bilinguals showed slower reading rates, shorter forward saccade length and higher number of regressions for L2 reading. However, L1 and L2 reading patterns were modulated by the bilinguals' current exposure to their L2, such that those with more current L2 exposure demonstrated increased reading fluency in L2 but decreased reading fluency in L1. The same authors also found similar results when examining paragraph-level processing of L1-dominant English-French and French-English bilinguals (Whitford & Titone, 2012). Again, the bilinguals exhibited significantly reduced reading fluency in their L2, this time indicated by longer gaze durations (i.e. first-pass fixation duration) and total reading time. This study however focused more on the effects of word frequency and how current L2 exposure modulates their impact on L1 and L2 reading. The reduced reading fluency in L2 indicates that the magnitude of word frequency effects was larger on L2 than L1 reading but this effect was modulated by current L2 exposure.

In another study that contrasted the reading patterns of L1 and L2 readers, Kang (2014) presented L1 and L2 speakers of English with a hypertext lesson in English containing 11 stimuli to examine the differences in the two groups' attentional distribution, speed at locating answers, learning time and comprehension. The author concluded from his findings that despite differences in English proficiency and learning time, the L1 and L2 readers distributed their attention across stimuli in a similar manner and performed equally well in a subsequent comprehension test. However, the author examined 'what' and 'how much' of the content were attended to without specifying the areas of interest (AOIs) for analysis. For each stimulus, two heat maps were generated (one for each group) and compared. A heat map is a visual representation of fixation distribution over a visual stimulus that displays the focus of visual attention by colours. The author concluded from his examination of these heat maps that "there were no *significant* differences found between L1 and L2 readers' attentional distributions or eye gaze patterns" (p. 5). Examination of heat maps however involves a considerable amount of inferences and interpretations and as pointed out by Conklin and Pellicer-Sánchez (2016) "do not in any way constitute an analysis of data" (p. 457). Hence, drawing statistical conclusions from them may not be ideal, especially given the availability of fixation count and duration as alternatives. Findings derived for other questions were concluded without provision of *p*-values or effect sizes. Since the author only provided the descriptive statistics for five out of eleven stimuli presented, effect size calculations cannot give definite answers. In terms of how much content was attended to, even though the author reported no

significant differences between the L1 and L2 readers, 3 out of the 5 stimuli had effect sizes (*d*) between 0.4 and 0.7. For the 3 stimuli, the L1 readers attended to the contents 10% more than the L2 readers did. In terms of reading time, the author made the conclusion from 5 out of 11 stimuli that the L2 readers spent 62% longer on reading than the L1 readers. Again, no *p*-value was provided and calculation of effect sizes does indicate large differences in reading time between the two. Despite methodological issues, there seems to be differences in the ways the L1 and the L2 readers processed the hypertext lesson.

2.4.4. Conclusion

This section discussed the differences in the ways that L1 and L2 reading comprehension take place. Reading comprehension in one's L1 is already a multi-layered construct to begin with, made up of several components and processes. When reading comprehension occurs in one's L2, two languages rather than one are involved and prior literacy in the L1 necessarily implicates L2 reading processes by facilitating or inhibiting the L2 reading processes. Such interlingual interactions can be observed at different levels in the language system. Finally, the L2 reader also has limited linguistic knowledge of the L2. Prior L1 literacy or reading ability in addition to L2 knowledge have been forwarded as major predictors in L2 reading. The differences in the two readings can be observed from eye tracking data with inferior processing patterns displayed during L2 versus L1 reading. However, no differences in comprehension skills were observed. Text processing differences in the L1 and the L2 suggests that multi-representational reading in the two languages could differ as well. The next section will look at how multiple representations are used as an instructional tool to facilitate L2 learners in acquiring the language and its various competencies.

2.5. LEARNING A SECOND LANGUAGE WITH MULTIPLE REPRESENTATIONS

The realm of second language acquisition (SLA) has widely used multi-representational materials as an instructional tool. Materials that include audio-visual input in particular are often promoted as contextually rich resources for language teaching and learning (Montero Perez, Van Den Noortgate, & Desmet, 2013). Plass and Jones (2005) via their integrated model of SLA with multimedia argue that multiple

representations can support SLA by providing language learners with comprehensible input, facilitating their interaction with the L2 in a meaningful manner, and eliciting comprehensible output from them. In each of these phases, the learners are believed to engage in cognitive processing such as selection/ apperception, organisation/ comprehension and integration/ intake, all of which can be supported with multiple representations. When the learners are presented with a verbal input, for instance, they first need to attend to relevant aspects of the target language and by combining the verbal input with another representation can help this process. An L2 aural input presented in a listening skill class for instance can be enhanced by accompanying it with dynamic visualisations.

In this literature, the use of visual representations (i.e. pictures and even text in a written form) to accompany core verbal representations, for instance, is often attributed to the picture superiority effect from memory research, proposing that pictures are recalled better than words (Bisson, Van Heuven, Conklin, & Tunney, 2014b; Carpenter & Olson, 2012). Theories that contribute to this effect are the dual coding theory (Paivio, 1991), and the sensory-semantic model (Nelson, Reed, & McEvoy, 1977). Dual coding, for instance, suggests that pictures can be encoded in the memory as verbal and non-verbal information, both of which can then serve as a retrieval cue. The sensory-semantic model argues that pictures are perceptually more distinctive and varied in mnemonic features than words are. These features make them more likely to activate semantic features and encourage more elaborative semantic processing (Nelson et al., 1977).

Empirical support for the contributing role of multiple representations in SLA and its various competencies has been found in various studies. The following review begins with the use of multiple representations in the teaching and learning of vocabulary. From thereon, the review focuses on the particular representations or technologies used to accompany the core verbal representations and what kind of role they play in L2 learning. Although this literature is not directly related to the present research, it is nevertheless necessary in order to provide initial insights into how people read and learn with multiple representations in their minority language.

Teaching and learning of vocabulary typically involve multiple representations. The words can be verbalised for learners to learn their phonological forms and often written down so that not only their orthographic word forms are learned but to also segment the auditory input into manageable chunks (Bisson et al., 2014b). Such

exposure however does not facilitate meaning acquisition easily and therefore another input is necessary to help learners associate meanings with the phonological and/or orthographic word forms. One way of doing this is by providing learners with L1 translations in a spoken or written form. Another means is by exposing learners to corresponding pictures (static or dynamic) so that meaning can be derived from the pictorial information.

Empirical work on the use of pictures in vocabulary acquisition however does not always find positive results. Lotto and De Groot (1998), for instance, found that foreign language words in Italian were recalled better by pairing them with their L1 translations (Italian-Dutch pairs) than with corresponding pictures (Italian-picture pairs). Similarly, Carpenter and Olson (2012) in three experiments found pairing foreign words in Swahili with pictures (Swahili-picture pairs) had no advantage on cued recall performance over pairing them with their L1 translations (Swahili-English pairs). The authors explain this null effect by the participants' overconfidence in their ability to remember the foreign words from the pictures and in doing so, they failed to establish a sufficient association between them. After a retrieval practice and a warning against being overconfident, this bias was successfully removed, and the picture superiority effect emerged.

More empirical support was provided in various other studies. Farley, Ramonda, and Liu (2012), for instance, found that in addition to exposing learners of Spanish to abstract words (both written and spoken) and their L1 translations, meaning recall can be enhanced by presenting them with corresponding pictures that are "metaphorical, emotive, or symbolic" (p. 450). Bisson and colleagues in a series of studies that examined the impact of pictures on the incidental acquisition of foreign words similarly found an advantage for foreign words that are accompanied by pictures. In Bisson, van Heuven, Conklin, and Tunney (2013), combining spoken and written Welsh words with corresponding pictures in an incidental learning phase resulted in them being more accurately translated in a subsequent explicit learning phase than words that the participants have not been exposed to. This finding suggests that during the incidental learning phase which involved a letter-search task, the participants processed the pictures and linked the Welsh word forms to their meaning from the pictorial information, although the pictures were not directly relevant to the letter-search task. In Bisson et al. (2014b), pairing spoken Welsh words and written L1 translations (Welsh-English pairs), and combining these pairs with corresponding pictures (Welsh-English-picture combination) emerged superior to simply exposing

participants to spoken Welsh words. Words presented in the first and second conditions were more accurately translated in a subsequent explicit learning phase than those in the third condition. In a delayed task, more words in the second condition were recalled than the other two suggesting an important role of pictures in incidental vocabulary acquisition. This was supported by the finding that time spent looking at the pictures significantly predicted recognition and recall performances after the one-week delay.

Similar facilitative role was found with the use of videos (for instance, Lin, 2010) although arguments were made against their capacity in providing opportunities for language learning (Montero Perez et al., 2013). The use of on-screen texts or subtitles that accompany the dynamic pictures and the L2 audio is one way of resolving this issue. Captions or same-language subtitles, for instance, have been found to facilitate both vocabulary acquisition and listening comprehension and in a meta-analysis by Montero Perez et al. (2013), the use of captioned videos emerged superior to non-captioned videos in these two aspects.

Multimedia glosses or annotations which provide readers with brief explanation of key vocabulary items in the form of text, picture and/ or audio were found to facilitate both vocabulary and reading comprehension. This is especially true in the case of glosses that combine verbal and visual representations (Plass, Chun, Mayer, & Leutner, 1998; Plass, Chun, Mayer, & Leutner, 2003). This reference technique allows readers to look up the meaning of words while reading (Golonka, Bowles, Frank, Richardson, & Freynik, 2014; Yun, 2011) and therefore encourages language learners to interact with the reading material (Plass & Jones, 2005). Researchers typically manipulate the types of glosses to see which one or which combination is best for language learning. For instance, by clicking on a glossed word, a learner can see its L1 translation (verbal gloss), its corresponding picture or video (visual gloss) and/or listen to its pronunciation in the target language (audio gloss).

Despite the brief review, it can be seen that multiple representations by and large have been found to facilitate the acquisition of the L2 and its various competencies. But how exactly, and do the language learners really make use of the additional representations? To answer these questions, some researchers look for clues in the learning process itself, instead of deriving conclusions solely from performance tasks. Such investigation, especially those that examine learners' eye movements are particularly relevant to this thesis as they can provide an insight into how people

learning in their L2 distribute their visual attention across different representations and how this is related to learning. As reviewed earlier, Bisson et al. (2014b) have observed that pictures were processed and pictorial information extracted in the incidental learning phase in addition to the verbal input. Additionally, the amount of time spent on the pictures significantly predicted recognition and recall performances of foreign vocabulary. The processing of pictures in addition to verbal input was also found on other multi-representational language learning materials such as picture storybooks with audio (Tragant Mestres & Pellicer-Sánchez, 2019), videos with subtitles (Bisson, Van Heuven, Conklin, & Tunney, 2014a; Tragant Mestres & Pellicer-Sánchez, 2019), and passages with glosses (Warren, Boers, Grimshaw, & Siyanova-Chanturia, 2018) although they were generally processed much to a less extent than the text.

A study by Verhallen and Bus (2011) on the processing of static pictures in picture storybooks further illustrates the potentials of using pictures to promote vocabulary acquisition and text comprehension. It involved young preliterate children learning Dutch as an L2 and tested the idea that fixations are "time locked to referential expressions in the text" (p. 480). Note that the relationship between the pictures and the story text was harmonious or representational. However, there were visual elements not mentioned in the text, and similarly, there were textual elements not depicted in the pictures. The children viewed the pictures on a computer screen and the text was presented orally. The children were exposed to one of five storybooks four times over a period of two weeks and their eye movements on the first and fourth exposures were compared. Analyses revealed that the text appears to limit which element of the pictures the children attended to. More and longer fixations were dedicated on visual elements that were highlighted by the text. After 3 encounters, the number of fixations on these elements dropped but the length increased, suggesting more information processing with each individual fixation. Interestingly, the number of fixations on non-highlighted elements remained stable with increasing exposure, suggesting that text does not, to an increasing degree, necessarily limit picture processing. The findings in general suggest that language as provided by the narration was not the only source of information young L2 children use to construct meaning. Although this study is less applicable to the present research given the non-availability of text display, attention should be drawn to the text-guided processing of pictures. Additionally, even though the participants were young, preliterate children, and have limited proficiency in the L2, they did manage to 'concretise' and attend to

relevant parts of the pictures, suggesting the potentials of using pictures to promote vocabulary learning and overall text comprehension.

Overall, the use of multiple representations in the realm of second language is largely facilitative. Multiple input, when attended to, can help L2 learners develop their competencies in the language and given these findings, it seems plausible to predict that the L2 readers in the present research would similarly benefit from the different types of representations in the reading material even though linguistic supports such as glosses and L1 input that are typically provided in language learning materials are not available. The next section will examine a more closely-related literature, one that deals with multi-representational learning in an L2.

2.6. LEARNING WITH MULTIPLE REPRESENTATIONS IN A SECOND LANGUAGE

This section reviews the relatively recent literature on the learning of content knowledge with multiple representations in an L2 to provide more insights into how multi-representational learning in the L2 takes place. Additionally, this section attempts to illustrate via empirical evidence, how multi-representational learning in the two languages (i.e. in the L1 and the L2) may differ, towards justifying the need for investigating experience with multi-representational materials in the L2. Learning with multiple representations here is considered as decoding of an L2 input and construction of a mental representation by integrating it with prior knowledge to arrive at meaning.

Some researchers like Mayer and colleagues have become increasingly interested in the L2 learning experience given the growing number of people acquiring knowledge in their non-native language. Such research stemmed from similar convictions driving the use of multiple representations in the L2 learning environment: learners with weak language skills can be supported in their learning with multiple representations. In Section 2.5, we have learned that visual information in the form of pictures can facilitate meaning acquisition of new words presented only in L2 verbal forms. Presentation of written text to accompany corresponding spoken text on the other hand can facilitate learning of orthographic word forms as well as segment the auditory input into manageable chunks. These ideas, along with the idea that people learning in their L2 may have underdeveloped linguistic skills to fully comprehend

verbal information, were put to the test by researchers such as Mayer and colleagues in their effort to examine whether certain multimedia design principles more commonly associated with multi-representational learning in the L1 are also applicable in the L2. These researchers typically manipulate how input is presented to provide the most conducive and optimum learning environment for L2 multi-representational learning. This section will briefly look at this body of research arguing that although these principles do not exactly map to the research conducted in this thesis given its focus on learning from a printed material with written text and static pictures, however, understanding some of the arguments put forward in these studies are nevertheless relevant and might explain why and how the learning with multi-representational materials in L1 and L2 may differ.

The review begins with the influential multimedia principle that states people learn better from text and pictures than from text alone (Mayer, 2005b). Mayer and colleagues first tested its applicability in a study that attempted to alleviate learners' difficulty in accessing meaning of difficult words that were narrated in the L2 by adding redundant and representational visual information in the form of a video to the verbal information (Mayer, Lee, & Peebles, 2014). In this study, undergraduates in an American university were presented with a 16-minute lesson about wildlife in Antarctica containing an L2 audio (audio group) or the audio accompanied by a video (audio+video group) containing pictures that depicted the animals and actions that were being narrated. Multimedia effect was indeed present as the audio+video group significantly outperformed the audio group on a comprehension test. Do note however that this effect was to be expected given that the instructional material was an excerpt of a documentary, naturally meant to be listened to and watched simultaneously. Depriving the audio group of the visual content could have consequently deprived the participants from a more effective material from which to fully comprehend the content. Other methodological concerns should be pointed out such as finding a floor effect on a summary task (i.e. means of 2.85 and 2.97 out of 114 main ideas). Additionally, the participants were a mix of international and local undergraduates with the local "who reported that they had initially learned a language other than English" (p. 656) making up a large majority of the sample. Despite a potential large variance in English proficiency, only self-reported rating of English listening ability was available from which to gauge the participants' English proficiency. Additionally, they used a British English narration when the participants would almost certainly be more familiar with American English.

Nevertheless, the findings were replicated in another study with two different populations studying in a non-native English environment; native-speaking Korean high school students and Korean university students (Experiment 2) in Korea (Lee & Mayer, 2015). The same Antarctica documentary was presented to these two groups. Unlike the first study (Mayer et al., 2014), however, a methodological check was conducted. Specifically, the audio was translated into Korean, the participants' L1 (Experiment 2B) to examine whether the facilitative effect of video would disappear. As the authors predicted, the facilitative effect of the video disappeared and it was argued that adding the redundant video was neither beneficial nor harmful when the audio was in the students' L1 as their linguistic skills were already well-developed. It therefore appears that the redundancy of the pictures affected learning in L1 and L2 differently and as suggested by the authors, it could be due to the differing proficiency in the instructional language.

Similar mismatch between L1 and L2 learning with multi-representational materials can be seen with regards to two of Mayer's multimedia design principles: the modality and the redundancy principles. The modality principle states that people learn better from images accompanied with spoken rather than written text, whereas the redundancy principle argues that people learn better from images accompanied by corresponding spoken words than from images accompanied by both corresponding spoken and written words. According to some researchers, the problem with having written text together with pictures is the possibility of overloading the visual channel due to the processing of two visual inputs. Both modality and redundancy principles are commonly associated with learning in L1 with several boundary conditions already outlined.

Mayer and colleagues tested the applicability of these principles in an experiment that attempted to alleviate learners' difficulty in listening skills and phonological processing in L2 by adding another verbal input, i.e. L2 subtitles that are completely redundant to its spoken form (Mayer et al., 2014). In the experiment, a group of international undergraduates were asked to watch an excerpt of a MythBusters episode with L2 subtitles (subtitle group) or without the subtitles (no-subtitle group). The subtitles were neither beneficial nor harmful to the learners' retention or transfer of knowledge as the subtitle group performed similarly as the no-subtitle group. Interestingly, analysis of the learners' self-report cognitive load measure revealed that the subtitle group did not report more difficulty with the lesson they received compared to the no-subtitle group yet reported less effort used for the lesson. Hence, it appears that having more

input to attend to is not much of a problem when learning was in the L2. This however did not translate to better learning outcome performance. Similar studies on these two principles have *not* found evidence on the modality and redundancy effects in the L2 multi-representational learning. They are either nullified (Kozan, Erçetin, & Richardson, 2015; Liu, Jang, & Roy-Campbell, 2018; Şendurur, Doğusoy, & Yondemir Çalişkan, 2018; Van Der Zee, Admiraal, Paas, Saab, & Giesbers, 2017) or reversed (Lee & Mayer, 2018; Lin, Lee, Wang, & Lin, 2016). In explaining the reversed effect, Lee and Mayer (2018) argued that, "these (modality and redundancy) principles may not apply when the words in the lesson are unfamiliar or the verbal material is complex for the learner," (p. 649).

Indeed, research in the L2 learning environment has suggested that some of the multimedia design principles are not as relevant for language learners as they need redundant, repetitive exposure to L2 input to acquire competency in the language (see Section 2.5). Applying these principles to those learning *content* knowledge in their L2 may also lead to similar conclusions, especially if these learners are still acquiring competency in the language. Overall, the null and reversed effects strengthen the idea that L1 and L2 multi-representational learning may not be identical. As suggested in the above-mentioned literature, L2 readers could process multi-representational materials differently than the L1 readers due to inferior language skills in their L2.

The above-mentioned studies are generally outcome-oriented and drew their conclusions from tests or surveys conducted post-learning. Little was done to look at how the representations were actually processed during learning and how these processes can be linked to the outcomes of learning. Several exceptions are reviewed next.

One study gave a particular attention on how graphic novels in English were read as a literary text by Malaysian primary school pupils (Yusof, Lazim, Salehuddin, & Shahimin, 2019). The authors were particularly concerned with whether the verbal and visual elements in the novel were effectively used and which of the two received more of the pupils' attention. To this end, the pupils were presented with extracts from the *Nancy Drew* and the *Hardy Boys* series on-screen and were asked to read them silently. Although their comprehension of the graphic novels was assessed, its relationship with processing measures was not addressed in this study. The analysis of the eye tracking data involved qualitative inspection using both gaze plots and heat

maps. As such, no statistical results were available. Although this technique may be appropriate to address their technical questions on reading paths, it is argued here that statistical tests may be more appropriate to address issue of attentional distribution between text and pictures. Regardless, their qualitative inspection of the pupils' eye movements revealed that the pupils by and large attended to the text more than they did the pictures.

A similar finding that multi-representational processing in the L2 was largely text-dominant was revealed in Pellicer-Sanchez et al.'s (2018) study that examined how non-native English speakers processed textual and pictorial representations in the presence or absence of audio. Unlike Yusof et al.'s (2019) study, Pellicer-Sanchez and colleagues examined the relationship between processing and comprehension measures. The eye movements of two L2 groups, i.e. English as a foreign language (EFL) children in Spain and English as a Second Language (ESL) university students in the UK, were collected and conveniently, the authors also collected data from native English speakers for comparison purposes. The reading materials for the two L2 groups were not the same, given their differing age groups and English vocabulary size. However, the L1 and L2 adults received identical reading material. The children read an illustrated story modified from an existing graded reader, whereas the adult groups read an illustrated mystery, which was written for the study. Both materials included a block of text and one picture on each page. Half of the story was presented in reading-only mode and the other half in reading while-listening mode. Analysis of dwell time percentages revealed that regardless of the reading condition, all participant groups spent a longer time on text than they did on the pictures. And as to be expected, provision of audio, led to significant differences in processing of the two representations as all groups spent significantly less time processing the text and significantly more time processing the images in this condition than in the reading-only condition. Interestingly, the L1 and L2 adults spent similar amounts of time on text and pictures regardless of the reading condition. Despite this finding, their comprehension performances differed with the L2 adults surprisingly outperforming their L1 counterparts. Additionally, regardless of the reading condition, all participant groups performed similarly on the text and picture-related comprehension questions. Interesting differences were found when the relationship between dwell time and comprehension scores was examined. More specifically, dwell time on text was positively related to text-related comprehension scores for the L2 adults but the relationship was negative for the L1 adults. Additionally, dwell time on pictures was positively related to text-related comprehension scores only for the L1 adults. Finally,

dwell time on pictures was related to picture-related comprehension scores for both the L1 and L2 adults. The L2 children's dwell time data however was not related in any way with their comprehension scores.

Pellicer-Sanchez et al.'s (2018) findings show that processing information is useful to predict the comprehension from a multi-representational material in the L1 and the L2 and more importantly, they show how the L1 and the L2 processes may be different. One issue to point out with regard to the eye tracking studies is the nature of the materials used. Both were narratives and the accompanying pictures mainly served to make concrete the characters, actions and scenes in the stories. In the case of Pellicer-Sanchez et al. (2018), the pictures "made it possible to create a narrative around them" (p. 14). It could be argued that the pictures in this type of materials may play much less of a role in learning than pictures used in more informative and domain-specific materials. As such, conclusions that were drawn from these studies could be more applicable to multi-representational materials that are narrative in nature. It is unfortunate that only a few studies have looked at the processes that learners go through during learning with multi-representational materials in the L2. This is actually surprising given the relatively recent interest given to the L2 community. Older studies with the L1 community have already addressed this issue, combined process and outcome measures and found associations between them.

The above-mentioned studies also make salient a number of points. It appears that when learning conceptual knowledge takes place in one's L2, certain principles that commonly apply to learning with multi-representational materials in L1 may disappear. The reasons being that the L2 speakers have underdeveloped L2 linguistic skills compared to their L1 counterparts and therefore need extra support to facilitate their comprehension of the multi-representational materials. This extra support however is not necessary in the case of the L1 speakers as their linguistic skills are already well-developed.

Another point is the scarce amount of studies that directly compare learning from multi-representational materials in L1 and L2. More research is needed to determine the differences in learning from multi-representational materials in these two conditions. Additionally, a common aspect in the reviewed studies is the interest in presenting content in dynamic and auditory forms such as via videos, subtitles, narrations or a combination of these media. Even though these multimedia environments are not at all foreign when learning in an L2, there is still a need to

examine more conservative multi-representational materials like printed or digital textbooks comprising of written text and static images. These materials are still widely used and relevant in academia despite the influx of more modern and advanced materials. Still on the issue of the instructional materials, many of the studies reviewed here did not use materials from existing academic courses, with the exception of (Van Der Zee et al., 2017; Wynder, 2018). Some used existing television documentaries or designed videos from which to gauge outcomes of learning. As pointed out by Mayer et al. (2014), there is a need to use materials from existing academic courses that are more conducive to learning and can subsequently increase ecological validity. Additionally, by using authentic instructional materials, more reliable conclusions can be drawn and hopefully they can be generalised to other similar materials.

To conclude, this section hopefully has made salient the need to conduct an investigation on multi-representational reading and learning in the L2. It has provided empirical support for proposing that multi-representational reading and learning occur differently in the L1 and the L2. The research questions formulated to guide this investigation is discussed in the following section.

2.7. RESEARCH QUESTIONS

This section introduces the four research questions that the thesis aims to answer in order to fulfil the research objective, i.e. to understand the ways in which the L2 readers read and learn from a multi-representational text in English. Note that reading here refers to the decoding of input (i.e. text and pictures), whereas learning (in the context of the main studies) refers to participants' understanding of the materials that retained and (partially) measured later through a post-test. The first three questions which provide insights into the L2 readers' engagement with multi-representational texts and extends current research on multi-representational learning are addressed in Chapter 4 (or *the L2 study*). The final research question is addressed in Chapter 6 (or *the Comparison study*).

Existing research has already revealed much about the reading of multi-representational texts in L1. Section 2.3 has shown that L1 reading of multi-representational texts is largely text-dominant. However, differences in the L1 processing patterns and strategies can still be observed and these differences can predict what readers come to understand. In contrast, not much is known about the

L2 reading with multi-representational texts. As seen in the eye tracking studies in Sections 2.5 and 2.6, there are indications that L2 reading of these texts is similarly text-dominant. However, the materials used in these studies were narratives and the pictures that accompany the text mainly served to make concrete the characters, actions and scenes in the stories. Pictures in this type of materials may play much less of a role in learning than visual representations used in more expository and domain-specific materials such as science textbooks. As such, there is a need to study the reading behaviour of L2 readers when processing these types of multi-representational materials. Hence, the first research question is addressed is:

1) How do readers of English as a second language (i.e. L2 readers) read authentic multi-representational material in English?

Given the limited research on this issue, RQ1 is exploratory in nature and examines the L2 reading behaviour by specifically looking at:

(i) The differences in the amount of attention allocated to text versus pictures, and

(ii) The differences in the extent of integration made between text and pictures

An important factor in multi-representational learning, or any learning for that matter, is whether it works for everyone. Both reading and learning with multi-representational materials are influenced by the characteristics of the learners. In Section 2.3, we have seen the important roles of prior knowledge and reading comprehension ability on the processing of text and pictures, as well as on comprehension and learning of these materials. Again, bear in mind that this may only be relevant to the L1 samples. Given that this thesis is concerned with readers of English as an L2, their proficiency in the language may play a substantial role on their processing and learning of the material. As such, two English proficiency variables; English reading comprehension ability and English vocabulary knowledge, a strong predictor of the former, are considered. More specifically, these three individual characteristics will be examined to see how they are related to the reading behaviour of the L2 readers. The second research question is thus:

2) Is the reading behaviour of the L2 readers when studying an authentic multi-representational text in English related to individual characteristics in their

prior knowledge, English reading comprehension ability and English vocabulary knowledge?

In Section 2.2, it was argued that integrating information available from MERs is crucial to realise the benefits from MERs and in Section 2.3, empirical evidence was found to support this proposition. Text-picture integration in particular has been shown to be systematically related to the outcomes of learning in a positive manner. However, this may only be true in the case of L1 reading. With regard to the L2 reading, the literature reviewed in Sections 2.5 and 2.6 showed that fixations on pictures are also predictive of outcomes of learning. These studies however did not particularly look at text-picture integration and the materials used were either narratives or meant for language teaching and learning, or both. Conclusions drawn from these studies may therefore be less applicable to the present thesis. The third research question:

3) Is the reading behaviour of the L2 readers when studying an authentic multi-representational text in English related to the outcomes of learning?

will use the eye tracking data gathered for RQ1 and relate this with learning outcome performance to see what kind of role text-picture integration plays in multi-representational learning in the L2.

The body of research reviewed in Section 2.4 has generally demonstrated that people read less fluently in their L2 than in their L1. This difference suggests that L1 and L2 multi-representational reading could differ as well, especially with regard to the processing of text areas. Further strengthening this proposition are the findings that certain multimedia design principles that work in L1 multi-representational learning do not seem to apply when the learning occurs in the L2 (see Section 2.6). It seems therefore imperative to compare and contrast the reading behaviour and the learning of multi-representational texts in the L1 and the L2. The *comparison study* in Chapter 0 will therefore focus on the question:

4) Are there differences in the reading behaviour and learning outcomes of the L1 and L2 readers when studying an authentic multi-representational text in English?

The reading behaviour and learning outcome data gathered from *the L2* and *the L1 studies* are thus combined and analysed. Since the L2 readers are predicted to have lower English proficiency than L1 readers, a primary concern is that they might also be disadvantaged in learning with this material. Hence, in addition to prior knowledge, the L1 and the L2 readers are compared in terms of their English reading comprehension ability before their reading behaviour and learning outcomes are compared and contrasted.

3. METHODOLOGY

3.1. INTRODUCTION

This chapter describes and justifies the philosophical and methodological approaches and considerations undertaken in this research. The chapter begins with a description of the post-positivist approach the research subscribed to followed by the explanation of its design, i.e. the non-experimental fixed design and the justification of using laboratory research more commonly associated with the field of psychology. A detailed account of the sampling, reading material, classification of its various representations along with participants' own account of the representations are also presented. Then, the methods of data collection and analysis are detailed and justified. In particular, the eye tracking methodology and relevant measures are introduced. Finally, the chapter explains how ethical considerations were addressed in the research. It should be noted that, in line with post-positivist tradition, details specific to each study (such as the participants) are discussed with each study.

3.2. THE POST-POSITIVIST APPROACH

This research subscribes to the post-positivist worldview which perceives reality as existing but impossible to attain. The pursuit of reality is seen as imperfect and probabilistic in nature partly due to the shortcomings or fallibility of the researchers themselves (Creswell, 2014b; Phillips & Burbules, 2000; Robson, 2011). Despite acknowledging that attainment of truth is not possible, and they themselves are fallible, the post-positivists strive to deliver knowledge that is highly credible and warranted by means of a rigorous inquiry. For the post-positivists, reality can be understood by seeking for general laws and theories. To find out what truth is, the scientific approach of investigation is adhered to and theories are tested through data collection (Creswell, 2014b; Robson & McCartan, 2016). Precise and unambiguous hypotheses and research questions are formulated, and variables that have been carefully specified are quantitatively measured to support, refine or abandon the theories. Should the investigation be conducted successfully, then there is a need to accept the evidence as the best and most credible for the time being. Additionally, someone can be more confident in their conclusion when related studies "point in the same direction" (Robson, 2011, p. 23). However, this conclusion is, similar to the researchers, fallible and upon discovery of more credible evidence would

consequently have to be refined or abandoned. The post-positivists' version of reality is therefore self-correcting, new knowledge results from detection and correction of errors to the old ones (Kantowitz, Roediger, & Elmes, 2009). Additionally, the post-positivists do not deny that their inquiry can be externally influenced but they are committed to always keep these potential biases in check and remain objective.

In alignment with the post-positivist stance, this research has strived to understand how two groups of readers read and learn from an authentic multi-representational text in English. To obtain an insight, the research adhered to the scientific approach, starting with theories derived from relevant literatures and putting them to test by formulating relevant hypotheses and research questions. Variables identified were then measured through quantitative analyses, resulting in some theories being supported, refined and even abandoned. It is believed the research has arrived at the best and most credible evidence, at least for the time being. Nonetheless, they are fallible and not absolute. Upon the presence of more credible evidence, these conclusions should therefore be refined or abandoned.

Undoubtedly, there are other alternatives of understanding this phenomenon of interest. The closest alternative would be the positivist approach which not only perceives that reality exists but is also absolute and indisputable (Phillips & Burbules, 2000). For the positivists, there is only one account of truth and it can be found and measured objectively. Like the post-positivists, the positivist inquiry is scientific. However, this inquiry necessitates the researchers to maintain an absolute distance with what is being researched so that the inquiry and the findings remain objective and value-free. Their methods of inquiry are also decontextualised for this purpose. Even though the present research has similarly strived towards objectivity, it is possible that the researcher's own values such as background knowledge and interpretation of findings have fed into the inquiry but as explained above, potential biases as much as possible were identified and eradicated. Furthermore, the findings and conclusions presented in this thesis are not absolute for they are circumstantial and can be falsified with more credible evidence. The positivist approach of inquiry is therefore unsuitable to approach the phenomenon of interest.

By accepting that an inquiry of knowledge is circumstantial and imperfect, the present research could have considered the constuctionist approach as well. The constructionists view knowledge as something that is constructed by individuals and therefore dependent on how they view and make sense of the world (Robson &

McCartan, 2016). This sort of inquiry however would have been very much value-laden, not only by the values of each individual participant but the researcher herself for meanings are co-constructed via interactions between the researcher and the participants. To remain objective and keep potential biases in check would then be difficult. This is not to deny that each participant's account of 'reality' is valuable but by solely depending on these multiple "believed realities" (Phillips & Burbules, 2000, p. 37), it would have been difficult to formulate methods of inquiry "that can support commonly adjudicated truth claims" (Phillips & Burbules, 2000, p. 37) and as a result, the research could be irreplicable. It would not suit the analytic approach within this thesis which uses statistical methods to group individuals and look for commonality across participants.

3.3. RESEARCH DESIGN

3.3.1. Non-Experimental Fixed Design

Post-positivism is commonly associated with the quantitative research paradigm. In this paradigm, hypotheses and research questions are first formulated to test theories and data are then collected to support or refute them (Creswell, 2014b). The type of quantitative design used in the present research was the non-experimental fixed design in which researchers observe variables of interest instead of manipulating or changing them deliberately (Robson & McCartan, 2016; Shadish, Cook, & Campbell, 2002). Since the present research was interested in understanding and when possible, explaining the L1 and L2 readers' reading behaviour, the research was primarily driven towards describing this behaviour and documenting its characteristics. The nature of non-experimental fixed design therefore lends itself well to these objectives.

Since the research involved understanding the reading behaviour of the L1 and L2 readers, the research design associated with the field of psychology was employed. Psychology is the science of behaviour which aims "to understand why people think and act as they do" (Kantowitz et al., 2009, p. 4). Explanations of behaviour derive from research that complies with scientific methodology, involving the use of well-developed techniques to collect data and generate theoretical explanations. One of the research techniques widely used in psychology involves experimentation carried out in laboratories (Kantowitz et al., 2009; Myers & Hansen, 2011). Just like other

research techniques in the field, laboratory research examines the processes that govern behaviour (Kantowitz et al., 2009). Another aim of laboratory research is to demonstrate "the conditions under which certain psychological processes" are observable (p. 22). The present research employed the use of laboratory research to collect data and generate theoretical explanations about L1 and L2 readers' reading behaviour.

Since this research was theory-driven and centred on behaviour, an experimental design could be an alternative research approach. Conducting such design would mean that the researcher has more control over the variables and can manipulate them more actively. The research could for instance tackle questions such as "Do readers read differently when reading single and multi-representational materials?" and "Do readers learn better with single or multi-representational material?". Although answers to such questions are indeed valuable, they are not particularly useful within the context of current learning environments. As mentioned earlier in the thesis, the materials that learners encounter these days are already multi-representational and so students are already frequently learning with them. Hence withdrawing a representational element and making stimulus *not* multi-representational would not reflect what really is going on and therefore would not be particularly useful. Having said that, potential noise in the observation and outcome was controlled as much as possible by conducting the research in a controlled laboratory setting. Understandably, the ways in which people read in a laboratory setting may not be representative of how they read in real-life situations. However as argued by Kantowitz et al. (2009), what is crucial is "to understand the underlying processes rather than the physical situations that produce these processes" (p. 21). The physical situations in the laboratory and the real world therefore do not need to be very similar as long as the same processes are taking place. Nonetheless it is acknowledged that there are limitations in conducting laboratory studies, thus attempts to increase external validity of the research were undertaken by using an authentic reading material as a stimulus. Additionally, it is acknowledged that findings of the research are circumstantial, restricted to the specific circumstances under which observations were made. Conducting non-experimental laboratory research is therefore argued to be the middle ground between a research that values control and manipulation over variables and one that is representative of how learners actually read in real-life situations.

In addition to not using the experimental approach, other approaches commonly associated with the flexible research design were similarly left out. Such approaches are case studies, ethnographic and grounded theory. Such approaches are not as theory-driven and starting data collection without a theoretical framework is generally accepted. They are flexible because they are adaptive as the research evolves (Robson & McCartan, 2016). These approaches however are not particularly useful for the present research for although the reading and learning of multi-representational texts in an L2 is not very well-researched and documented, the researcher has to make use of theories and findings from relevant and related literatures and used them to design the research. It was therefore not adaptive as the research progressed because the theories, hypotheses, and methods of data collection and analysis were already laid out before data collection was carried out. Moreover, flexible design approaches focus on participants' perspectives and aim to present multiple realities. This could jeopardise the need to put forward objective findings and conclusions.

3.4. SAMPLING

A few criteria were first identified for sampling. General criteria which applied to both the L1 and L2 readers pertain to their science education and reading difficulties. For the first criterion, their most recent contact with formal science education was considered. Specifically, only those whose science education is below the standards of A-level or its equivalence were considered. In cases where this could not be determined (e.g. they did not sit for national examinations in their home country), their age when they last learned a science subject in a formal context was accounted for. In this case, only those who last learned it when they were 17 or younger were considered. The decision to exclude high-knowledge readers was because they could already be at ceiling in terms of knowledge of the subject matter discussed in the learning material. They may therefore not learn anything when presented with the material. Secondly, readers with visual impairments or reading difficulties were excluded from the sampling. Their reading patterns could be very different from readers with correct or corrected-to-normal vision and consequently influence the outcomes of the study.

To sample the L2 readers, two further criteria were identified; their English proficiency and first language. The English language proficiency criterion refers to performance

in an English proficiency test, such as IELTS or TOEFL. To be eligible for sampling, the L2 readers should be at most, 'good' users of the language. The sample therefore only included those who scored a maximum band of 7.5 in IELTS or a maximum score of 109 in TOEFL iBT. In cases where they did not sit for either test, their performance in the most recent English proficiency test was considered. 'Very good' or 'expert' users of English were excluded from sampling as their gaze behaviour may mirror those of native English speakers and could subsequently influence the outcomes of the study. A minimum requirement was not set as the sample was a group of English as L2 university students in the United Kingdom who have already met their university's minimum English language entry requirements. For example, the minimum requirement set by the University of Nottingham for its undergraduate programmes is IELTS 6.0.

The reading direction of the L2 readers' L1 was also considered. Even though the languages vary, the reading direction must match that of the English language, i.e. left-to-right and top-to-bottom, as this may influence how readers look at text and pictures, which have been spatially organised on the textbook. In visuospatial tasks for instance, native left to right and right to left readers have been shown to have differing levels of leftward aesthetic bias (Friedrich & Elias, 2016). Additionally, the two groups displayed different spatial tendencies when performing a number-based visuospatial task (Kazandjian, Cavézian, Zivotofsky, & Chokron, 2010). Native readers of non-alphabetic scripts like Mandarin were also included in the sample. Traditionally, Mandarin was written in vertical top-to-bottom columns, and arranged from right-to-left (Bergen & Chan Lau, 2012; Furhman et al., 2011). Readers of the traditional Mandarin would typically begin reading on the top right and move downward then leftward (Bergen & Chan Lau, 2012). A switch to writing the Mandarin script in horizontal rows from left to right occurred within the last century, and in Mainland China, the switch became official in 1956 (Furhman et al., 2011).

Even though there is a concern that the Mandarin readers' eye movement patterns may be qualitatively different from readers of alphabetic scripts', it has been argued that their reading processes are fundamentally similar Li, Bicknell, Liu, Wei, and Rayner (2014). For alphabetic scripts like English, word properties such as frequency and predictability affect readers' eye movement patterns. In Chinese however, words are not spatially segmented and composed of visually complexed characters. Character properties therefore are proposed to play a more influential role in determining eye movements in reading Chinese. Li et al.'s (2014) study showed that

despite the larger role of character properties, the pattern of effects that word properties have on reading Chinese is very similar to the effects that word properties have on the reading of alphabetic scripts. Hence, since all L2 readers would read the material in the same language, i.e. English, the underlying processes that drive their eye movements should be even more similar as processing of English scripts is primarily word-based.

3.5. THE MATERIAL

This section introduces the reading material used as the stimulus in the research, how and why it was selected and provides a description of the various representations contained in the material.

Three materials, all taken from the same science textbook for Year 7 (Levesley, Johnson, & Gray, 2008), were shortlisted as the research stimulus. A science material was chosen as the stimulus because it is more common to find a variety of external representations in science than non-science materials. A vast number of representations is available in its various domains. Additionally, and more importantly, despite the variety in numbers and forms of representations used in science materials, each is commonly used for a functional purpose and to add value to learning. Their inclusion for instance supports the learning of abstract and unfamiliar scientific concepts, contributes to a deeper understanding of complex scientific domains and scaffolds problem-solving (Ainsworth, 2006, 2014; Lee, 2010; Rau, 2017; Stieff, Hegarty, & Deslongchamps, 2011). Finding an authentic stimulus on a science-related topic containing a variety of functional representations was therefore easier and hopefully participants when reading it would be more likely to feel that they are reading in an authentic environment.

Each of the three materials was a two-page spread discussing different topics. “Pressure problems” discussed gas pressure and its effects, whereas “Using electricity” was about the use of current and electricity in everyday life. “Bits and pieces” on the other hand discussed the properties of solids, liquids and gases. They were shortlisted because they were representationally rich, consisting of chunks of text and several types of graphic representations. All are therefore reflective of the complex learning materials currently used both inside and outside of the classroom. It is common to come across more than one graphic representation per science

textbook page, for instance (Cromley, Bergey, et al., 2013). These representations have also increased in variety (Slough, McTigue, Kim, & Jennings, 2010). As pointed out by Hannus and Hyönä (1999), many materials used in the literature have been rather simple, consisting of text portions accompanied by only one picture. Authentic textbook materials however are much more complex, and learners therefore encounter complex cognitive tasks when presented with such materials (Ainsworth, 2006; Hannus & Hyönä, 1999). Having participants read and learn from simplified materials may therefore not reflect the complexity of the learning tasks and environments that students immediately face. Furthermore, learning performance examined from interactions with these materials may also not be applicable to the real-life learning environments. The present research therefore used an authentic passage of a science textbook to better reflect the complexity of processing and learning from such materials.

Additionally, due to the constraints of a lab study, the stimulus material had to be readable and accessible to readers within a short amount of time, yet still likely lead to learning. The selected materials met these requirements as the participants would be non-science students and would neither have studied the subject matter substantially nor recently. Finally, the absence of practical or hands-on activities (e.g. completing a table or drawing a diagram) also contributed to the selection as such an activity would potentially jeopardise the quality of the eye tracking data.

Once these materials were shortlisted, they were pre-piloted with several students. Participants were asked to read all three materials and their eye movements were recorded. After each reading, they were asked some comprehension questions. After the last material was read and answered, the participants were asked which material they liked most and why, which one they found easiest and most difficult and why. They were also asked about whether or not the accompanying pictures were useful for their learning of the materials. Analysis of their eye movement data, comprehension tasks and verbal reports provided initial ideas for which of the three should be selected for the research. The majority of the participants felt that “Bits and pieces” was easiest in terms of language and its pictures were the most helpful. The other material was judged to be rather technical, ‘sciencey’ and its pictures were not helpful. To finalise the material selection, the three materials were then presented at a weekly meeting with several other educational researchers which generated a lengthy discussion about each material’s strengths and weaknesses. The input from this discussion supported the findings from the pre-pilot study and the “Bits and

pieces" passage was finalised as the stimulus (see Appendix 1: Research Stimulus). This selection was made because there was no content error identified in the material, unlike one of the remaining two. Secondly, most of the attendees of the weekly meeting preferred the material over the other two and although affective factors were not particularly relevant to this research, it is nevertheless useful to have a likeable stimulus. Thirdly, and most importantly, the picture components of "Bits and pieces" are good examples of how pictures should be used in a science text. None of the pictures are decorative, all are functional and adding value to learning. Overall, "Bits and pieces" comprised of 615 words in English and seven pictures. It was scanned from the textbook and exported as a tagged image file format (TIFF). The two remaining materials were included in Appendix 2: Shortlisted Research Stimulus 1 and
Appendix 3: Shortlisted Research Stimulus 2.

3.5.1. Exploring the Material

This section provides a description of the various representations contained in "Bits and pieces". The first part will focus on the sentential representations or the text, its components and characteristics. The second part will describe the graphic representations or the pictures, describing them from the structural and functional perspectives. Additionally, an account of how participants themselves perceived the various representations, is also presented.

3.5.1.1. The text

The text is made up of five components, broadly categorised as contents, questions, outcomes, captions and headings (see Figure 1). The contents refer to parts of the text that discuss the main ideas of the material, i.e. the properties of solid, liquid and gas, making up the main body of the material. The questions are eight open-ended items, located at various positions on the two-page spread. All answers to these questions are explicitly stated in the material, either within just one sentence or requiring readers to integrate information across larger distances, such as across multiple sentences or multiple representations (i.e. text and picture). The outcomes refer to a small section at the end of the material that summarises what readers should have gained from the lesson. Captions are either labelling of pictures located directly below them, or more lengthy annotations located within the pictures themselves. Captions however have not been segmented as text in the eye tracking analysis and

are analysed as part of the pictures. Finally, headings are the title of the material and the two headings that divide the text content into parts. The ‘How Science Works’ logo is also categorised as a heading. The categorisation of sentential representations is only for descriptive purposes and to help consider how well the findings will generalise to other materials. In the eye tracking analysis, these different text components excluding captions, are not differentiated and constitute just a single unitary representation, i.e. the text.

Table 1 summarises the characteristics of the text in terms of its word and sentence counts, average word frequency and readability. All items in the table with the exception of the average word frequency were obtained from Microsoft Word’s readability statistics. The text was typed into the word processor before these statistics were obtained. The Flesch reading ease and Flesch-Kincaid grade level tests are assessments of the overall text’s readability. They measure length, in that their ratings are based on the average number of syllables per word and words per sentence ("Get your document's readability and level statistics," 2018). Hence for a piece of writing, the longer its words and sentences are, the more difficult it is to read (Flesch, 1948). For the average word frequency, the copied text was first submitted to an online text analyser (i.e. Online-Utility.org) to generate its wordlist. This list was then submitted to another online text analyser (i.e. http://meshugga.ugent.be/open-lexicons/interfaces/subtlex-uk/) to obtain the word frequencies which are based on a large subtitle corpus of British television programmes (i.e. SUBTLEX-UK). Zipf scale was used to measure the word frequencies (van Heuven, Mandera, Keuleers, & Brysbaert, 2014). On the whole, the values in Table 1 indicate that “Bits and pieces” is fairly easy to read, consists of high frequency words and is age-appropriate for its intended audience; the Year 7 pupils.

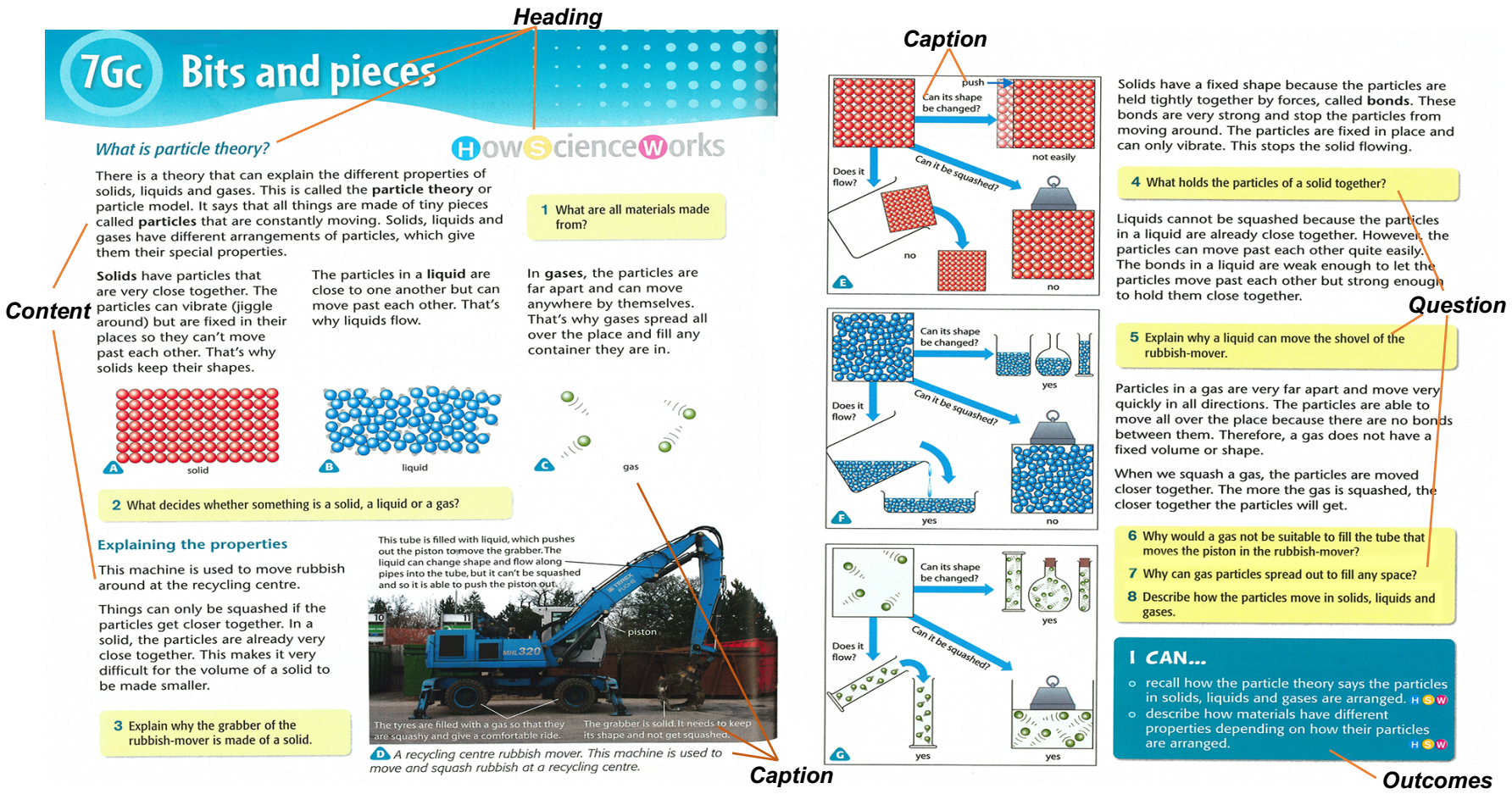


Figure 1. Examples of the various text components within the material

Table 1. Summary of the characteristics of the sentential representations

Number of words	615
Number of sentences	55
Number of words per sentence	10.7
Average word frequency (Zipf scale[1])	5.32
Flesch reading ease score[2]	74.3
Flesch-Kincaid grade level score[3]	5.5

3.5.1.2. The pictures

This sub-section describes the graphic representations from two perspectives. Ainsworth (2014) argues that representations can be described in terms of their structural or functional aspects. Following this argument, the graphic representations are classified using the structural and functional taxonomies already established in the literature. The other perspective considers what readers themselves have to say about the representations. Again, just like the text, the categorisation of these graphic representations serves a descriptive purpose. It also serves to consider how representative the material used in the studies is and consequently how likely the findings are to generalise to other materials. In the eye tracking analysis, the seven pictures, will not be differentiated by type and constitute just a single unitary representation, i.e. the pictures. The research did not focus on how readers distribute their visual attention across the different types or functions of these graphic representations, although as argued in the final chapter this could be a useful extension in future work.

3.5.1.3. Structural and functional perspectives of the pictures

This subsection will first categorise the seven pictures using Lohse, Biolsi, Walker, and Rueter's (1994) structural taxonomy. Structural taxonomies generally focus on the physical structure of the graphic representations and therefore the form of the representations is of utmost importance. Using Lohse et al.'s (1994) taxonomy, three

[1] The scale goes from 1 (very low frequency words) to 7 (very high frequency words). The score of 5.32 indicates that the material as a whole consists of high frequency words (applies to scores of 4 and higher).

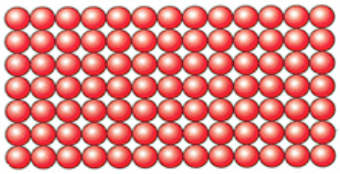
[2] The scale goes from 0 (practically unreadable) to 100 (easy for any literate person). The score of 74.3 indicates that the material is fairly easy (applies to scores of 70-80) to read.

[3] Text is rated on a U.S. school grade level. The score of 5.5 means that the UK Years 6 and 7 pupils can understand the material.

categories can be identified; structure diagrams, process diagrams and a photo-realistic picture. Refer to

Table 2 for an overview of the graphic representations. The first three pictures (i.e. pictures A, B and C) are structure diagrams (SDs). Each of these diagrams illustrates a static description of a physical object (Lohse et al., 1994). Picture A for instance is a static description of the solid particles. Spatial data in SDs convey the true coordinates of objects in the diagrams. All components of objects therefore appear in their real physical locations. Picture D is the only photo-realistic picture (Photo) in the material, and it visualises a realistic image of a rubbish-mover. The Photo is realistic as it has a one-to-one correspondence with the real-life machinery and preserves the real-life space between objects in the picture. The last three pictures (i.e. pictures E, F and G) are process diagrams (PDs). Such diagrams visualise the interrelationships among physical objects and relevant processes involved. The spatial data in PDs therefore reflect dynamic or temporal processes or relationships among objects. In Picture F for instance, one of the processes depicted is the flow of liquid particles from one container to another. The use of arrows is common in such diagrams as they suggest dynamism and progression.

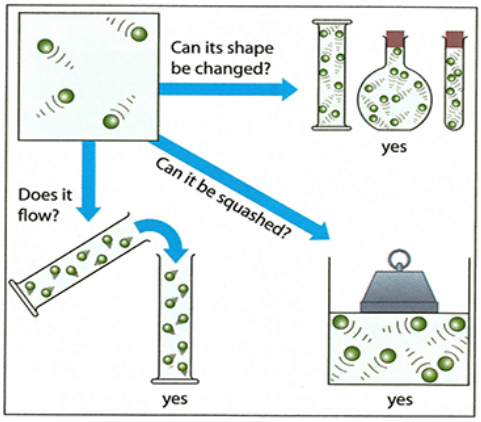
The functional taxonomies on the other hand, classify graphic representations in terms of their intended use and purpose and therefore focus on the content of the representations (Lohse et al., 1994). Levin, Anglin, and Carney's (1987) functional taxonomy of pictures in prose is used here to identify the function of each representation. The seven pictures in the material were all relevant to and supplement the text content, serving either a representational or organisational function. Pictures that serve a representational function are those that are considered to depict the entire or part of the text content, making it more concrete to the readers. Pictures A, B, C and D are therefore representational as they either express all or part of the text that they correspond to. Picture A (Figure 2) below for example, represents the entire content of its corresponding text, displayed on its right. The text describes the very close distance between solid particles and this is visualised by the structure diagram. Note that the text however only expresses part of what is depicted by the picture.



Solids have particles that are very close together.

Figure 2. Picture A and a corresponding text component

Pictures that serve an organisational function on the other hand organise the text content into a useful, structural framework, making it more coherent to the readers. Such pictures are commonly used alongside procedural texts or those that list down a property's distinct features. Pictures E, F and G in the material are all organisational pictures. Each organises the features of the state of matter it represents into a more coherent structure. Picture G (Figure 3) will be used to illustrate this.



1. In gases, the particles are far apart and can move anywhere by themselves. That's why gases are spread all over the place and fill any container they are in.
2. Particles in a gas are very far apart and move very quickly in all directions. The particles are able to move all over the place because there are no bonds between them.
3. When we squash a gas, the particles are moved closer together. The more the gas is squashed, the closer together the particles will get.

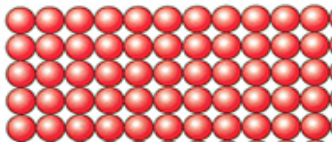
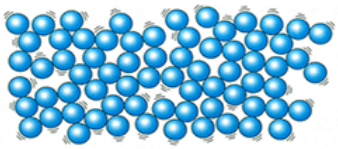
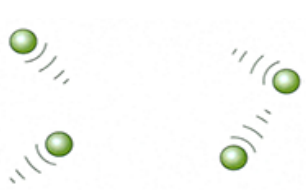
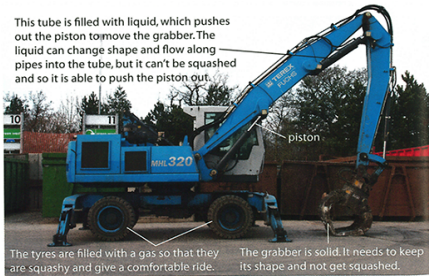
Figure 3. Picture G and three corresponding text components

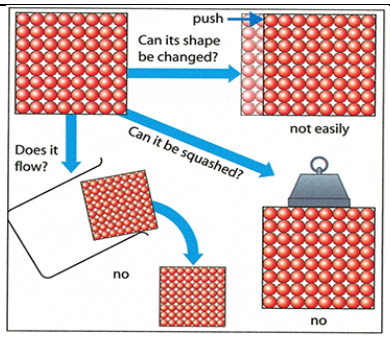
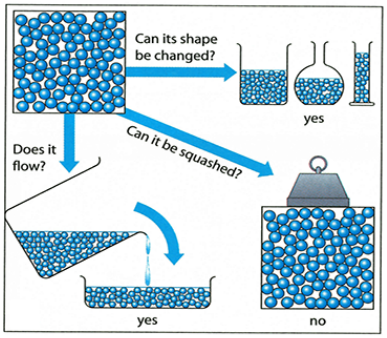
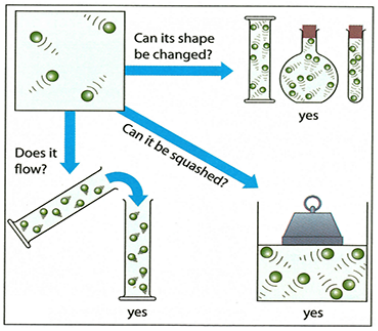
The green circles in Picture G represent the gas particles. There are multiple text components that correspond to this particular picture and all are listed below the picture. All of these components describe the distinct features or properties of gas

particles. Text component 1 describes the distance between particles, their mobility and their ability to fill containers. Text component 2 adds two more details; the speed of the movement and the absence of bonds to hold the particles together. Text component 3 adds yet another property, i.e. that gas can be squashed. Picture G brings together these properties which are actually located at different places on the two-page spread and organises them into just one but more coherent and succinct structure.

Table 2 summarises the classification of the graphic representations based on the two taxonomies. Overall, their variety not only reflects the richness of the material per se, but also the richness of current textbooks in general as the material includes many of the features one commonly encounters in textbooks nowadays.

Table 2. The classification of the graphic representations based on Lohse et al. (1994) and Levin et al.'s (1987) taxonomies

Label	Picture	Structure	Function	Example of corresponding text
A		SD	Representational	Solids have particles that are very close together.
B		SD	Representational	The particles in a liquid are close to one another but can move past each other.
C		SD	Representational	In gases, the particles are far apart and can move anywhere by themselves.
D		Photo	Representational	A recycling centre rubbish mover. This machine is used to move and squash rubbish at a recycling centre.

Label	Picture	Structure	Function	Example of corresponding text
E		PD	Organisational	1. Solids have particles that are very close together. The particles can vibrate… 2. Things can only be squashed if the particles get closer...
F		PD	Organisational	1. The particles in a liquid are close to one another but can move past each other. That's why liquids flow. 2. Liquids cannot be squashed because the particles in a liquid are already close…
G		PD	Organisational	1. Particles in a gas are very far apart and move very quickly in all directions. The particles are able to move all over… 2. When we squash a gas, the particles are moved closer together...

3.5.1.4. Readers' perspectives of the representations

This section presents the readers' perceptions of the various representations derived from verbal reports elicited via a cued retrospective protocol and semi-structured interviews. The reports were analysed using Ainsworth's functional taxonomy of multiple external representations (MERs) (1999, 2006, 2014) to explore the range of purposes the representations served in the eyes of the readers.

DeFT: The functions of multiple representations

This taxonomy was derived from findings of previous research that summarised how MERs are beneficial for learning and problem-solving according to three main functions; by supporting different information and processes, by constraining interpretations and by promoting a deeper understanding. These three ideas developed into the three functions that MERs serve in facilitating learning; to complement, constrain and construct. It is argued that representations can complement each other when each supports a different process or when each contains different information. Complementary representations that support different processes convey the same information but deliver it in different ways, for instance via text and a diagram. Text and diagrams each have unique computational properties that make them computationally non-equivalent. Sometimes, rather than having complex information contained in a single representation, it can be distributed over representations. These are representations that support complementary information. DeFT suggests that having such functional representations can therefore facilitate learning when learners benefit from each individual representation. The section main category in DeFT is when a representation can constrain the interpretation of another in two ways. First, a more familiar representation can constrain the interpretation of a less familiar one. Second, the inherent properties of a specific representation, such as its specificity, can constrain the interpretation of one that is more ambiguous. In both instances, the more familiar or the more specific representation provides a support for readers to understand and not to misinterpret the other that they know less about. The final constructing function is realised when learners integrate information from the different representations to construct deeper understanding of the subject matter. Here, multiple representations are necessary as such achievement would have proven difficult with only one representation. Ainsworth (e.g. 2006, 2014) also notes that a single multiple representational system can serve more than one of these functions simultaneously and she argues that developers' intentions

need to be considered alongside learners' actual response. Thus, this taxonomy was chosen as a method for analysing the verbal reports as it is comprehensive in the sense that in addition to considering the structural form of representations, more importantly it takes into account their specific educational purposes and pedagogical functions as perceived by developers and learners. Moreover, the taxonomy is more readily available to infer from their reports.

The verbal reports

To elicit readers' accounts of the various representations, five participants were recruited. Each received an inconvenience allowance for their participation. Two were native English speakers or L1 readers (1 female) and three were non-native English speakers or L2 readers (2 females). Of the three L2 readers, two reported Indonesian as their L1 and the remaining L1 was Malay. The five participants' average age was 35 years old. All the L2 readers were, at most, good speakers of English and were not visually impaired nor had any reading difficulties. In addition, the two L1s, i.e. Indonesian and Malay, both read from left-to-right, and top-to-bottom matching that of English. Similarly, the L1 readers were non-science students, and neither visually impaired nor suffer any reading difficulties. These sampling criteria were the same ones used to sample participants for the main studies (i.e. Chapters 4 and 0). The two L1 readers henceforth are coded L1-01 and L1-02 whereas the three L2 readers; L2-01, L2-02, and L2-03.

Data elicitation sessions took place on an individual basis. The participants were told that they would be asked to read a science passage and that their eye movements would be recorded. They were also told that after reading is completed, they would be asked to report their thought processes during reading and to answer some questions about the passage. After the participants agreed and gave their consent, they read the passage silently. Afterwards, they verbalised what they were thinking of when they were reading the passage. To help them recall their thought processes while reading, the recording of their eye movements was shown or 'cued'. The participants verbalised their thoughts while watching this recording. This utilises a version of cued retrospective protocol (CRP) or stimulated recall techniques in which some degree of support is provided to facilitate recall (Gass & Mackey, 2000). Each session was video recorded. After the CRPs, a semi-structured interview took place to elicit more information about the participants' perceptions of the material. The verbal reports were then transcribed and then coded using the functional taxonomy from DeFT (see Table 3 for the coding scheme).

Table 3. The coding scheme to explore participants' perceptions of the functions of representations based on Ainsworth's functional taxonomy

Function	Subcategory	Example statement
Complementary	Two or more representations, each contains unique information	- here (*pointing to the structure diagrams (SDs)*), it's quite simple and you're talking about the arrangements... And here (*the process diagrams (PDs)*) these are talking more, I think, this demonstrates more movements rather than just the arrangement…
		- you can easily (*pointing at the first PD*) see squash it, does it flow, it wouldn't really flow. Whereas here *(pointing at the second PD*), you can see the flow and this *(pointing at the third PD*) demonstrates it.
	Two or more representations, each with unique computational properties	- I realised that some of these were sort of reflective questions used obviously for the students to 'explain why the grabber of the rubbish mover is made of a solid.' And by looking at the imagery, it can more or less help the learner understand in a kind of multi-sensory fashion, I suppose because you got an image and you got text.
		- I think the diagrams are quite helpful and they reinforce what I've read… it seems as though I'm using both: the text and the diagrams, to aid comprehension there.
		- And I also paid attention to the pictures (*the SDs*) that could help me to understand the principles of the particles, when it was implemented in the pictures (*pointing to the SDs*).

Function	Subcategory	Example statement
Constraining	Two or more representations, one supporting the interpretation of the other through familiarity or ease of interpretation	- But then I jumped to the picture (*the Photo*) but that picture doesn't seem to be very helpful, so I've gone back to the diagrams (*the first and second PDs*)… I kind of fixated back on the diagram. So this is really the application of the theory. - I came across the word 'squash' (*Content 6*), mmm I'm trying to get what squash means because I thought I might understand it differently… So that's why I kept looking to this picture (*the Photo*) and referring to the text beside it (*Content 6*). - I think to use the illustration of this machine is a great idea, isn't it? Because you got the fact that the tyre is filled up with gas, solid. Illustrates in a real-world context what the theoretical stuff is.
Constructing	Two or more representations, upon integration a deeper understanding of a domain is achieved	Not mentioned

Before commencing with the results, the researcher's own account of the functional relationships between representations is presented first. This will then be contrasted and compared with the participants' perspectives. It must be noted here that it was not possible to know exactly the intention of the textbook authors for having used representations in the first place. To start off, the lesson introduces the particle theory and the concept of particles which make up the three states of matter; solids, liquids and gases. Basically, the lesson revolves around three main ideas: the arrangement of particles, their properties and the real-life application of these concepts. As such, it will be easier to explain the functions of the representations by dividing the material into three sections; Sections 1, 2 and 3. Readers are referred to Table *2* for the graphic representations and Figure 8 for numbering of text chunks.

Section 1 introduces the concept of particle; that solids, liquids and gases are made of differently arranged particles which give them each unique properties. Each matter, its particles, their arrangements and basic properties are described one after another. Following the textual description of each matter, a picture (i.e. a structure diagram or

SD) conveying that information ensues (see Figure 4 for an example). In such instance, each picture functions to complement its corresponding textual representation because both representations convey the same information, i.e. the arrangement of the matter's particles. They however convey this information in different ways. Perceptual cues in the pictures are readily available for readers to utilise, for instance to gauge distance between the particles.

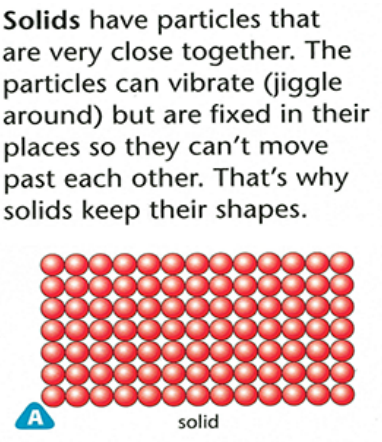


Figure 4. Text describing the arrangement of solid particles and the corresponding Picture A that visualises the arrangement

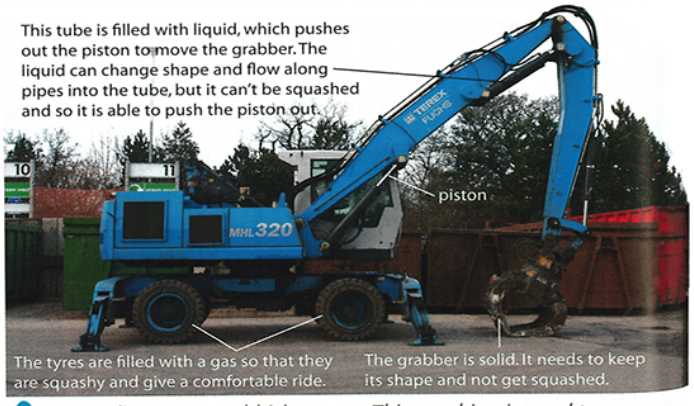
The participants also shared the same view, referencing the SDs for their role in complementing the textual representations in Section 1,

> And I also paid attention to the pictures (Pictures A, B and C) that could help me to understand the principles of the particles, when it was implemented in the pictures (pointing to Pictures A, B and C). [L2-03]

The lesson continues with Section 2 which describes how the particle theory and its concepts are applied in real life. In particular, this section makes reference to a rubbish mover machine and includes Picture D, the photo-realistic picture to visualise this reference (see Figure 5). As such, the text and Picture D both complement each other. L1-02 also agrees with this view:

> So I started to do quite a careful reading of this and then I was distracted by the machine on the right hand-side (Picture D) because it was referred to in the text…

This machine is used to move rubbish around at the recycling centre.



D *A recycling centre rubbish mover. This machine is used to move and squash rubbish at a recycling centre.*

Figure 5. Text referring to the rubbish mover machine, a photo-realistic picture of the machine, along with text captions over and beneath Picture D

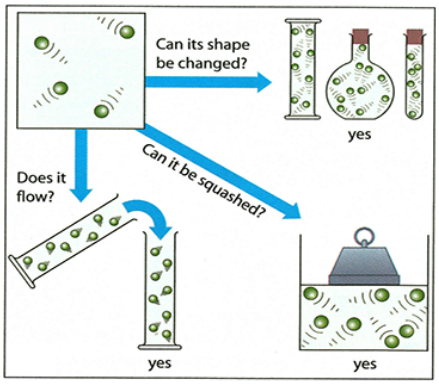
Over and below Picture D, there are text captions referring to several components of the machine. These components are indicated in the picture by lines. The text captions and corresponding Picture D components are therefore complementary as well.

Picture D may also serve a constraining function in relation to its text counterparts (i.e. both the text on the left of Picture D and the text captions). The rubbish mover machine and some of the components referred to in the text captions such as piston and tube may be unfamiliar to some readers. Picture D provides a concrete image for the readers to associate with and therefore might function to constrain the interpretation of their less specific textual correspondences. This was noted by L2-01 who had a little difficulty with the term used:

> This time, when I came across the word 'squash' (Content 6), mmm I'm trying to get what squash means because I thought I might understand it differently compared to what it's actually about so that's why when I read across this one ('squash' in one of Picture D captions), I came to this article (Content 6) again… So that's why I kept looking to this picture (Picture D) and referring to the text beside it (Content 6).

Section 3 further details the unique properties of each matter; the presence/ absence and strength of bonds, whether the matter can flow, be shaped, or squashed. Again,

this is done sequentially, one matter after another. Next to the textual description of each matter, a process diagram (PD) conveying that information is included. The two types of representations are therefore complementary, conveying similar information but in different forms (see Figure 6 for an example).



Particles in a gas are very far apart and move very quickly in all directions. The particles are able to move all over the place because there are no bonds between them.

When we squash a gas, the particles are moved closer together. The more the gas is squashed, the closer together the particles will get.

Figure 6. Picture G that visualises the text describing the properties of gas particles on its right

This relationship can be inferred from L1-02 and L2-03's statements:

> I started to closely consider the text (Content 7) … and I started to look at the image (Picture E) but then I moved quickly back to look at (read) liquids and I was making comparisons between the first part of the text on solids and the second part on liquids… We moved on to look at (read) the gas… And then that, (referring to fixations on Picture F then Picture E) yeah, I started to do a bit of close scanning of those images because I thought well the images and the diagrams are going to help me reinforce what I've learned. [L1-02]

> So I tried to read this one (pointing to Content 7) first about the liquid err, about the solid and then returned back to, OK, I saw the picture first (Picture E). And then after that I read this one (Content 8) and then I tried to relate with the picture (Picture F) also. So basically, I tried to understand this one (text chunks next to Pictures E, F and G) and then this one (Pictures E, F and G) I spent the last time I think to understand this part. [L2-03]

We can see the benefits of having these diagrams in that they can summarise a long list of characteristics into just one concise representation, taking advantage of

relevant perceptual cues like arrows and lines for instance to indicate processes and movements.

L1-02 noted this advantage as well:

> I think these images here (Pictures E, F and G) are very helpful… they kind of succinctly summarise the core concepts. You kind of almost get this flow chart thing, “Can it change its shape?”, “Not easy”, “Yes”, “No”. So I think these kinds of diagram are very helpful for learners… you got a lot of conceptual stuff going here. I was just impressed with the way… it kind of illuminates very complex concepts yet in a very simple clear way. Pictures sometimes speak more than a thousand words, as they say.

So far, the relationship between text and pictures have been identified. However, the various pictures are also related. Pictures A, B and C for instance, are sources of complementary information (refer to Figure 7). Although they all convey information about particle arrangement, each corresponds to a different matter, with its own uniquely arranged particles. Having three distinct representations to convey this information is therefore very useful.

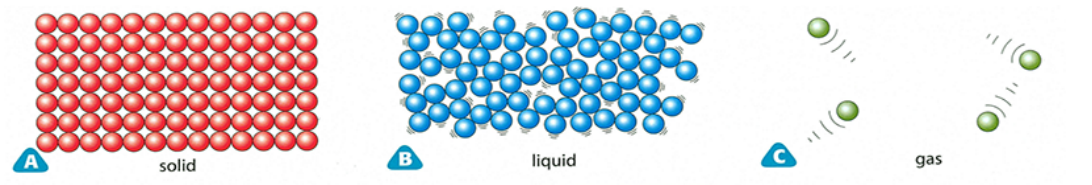


Figure 7. Pictures A, B and C, each representing a different state of matter

L2-03 noted the same relationship as well,

> so the picture (Pictures A, B and C) gave … for example, the solid particles are in this part and the liquids are in this one and the gases are in this part, so something like that.

Pictures E, F and G also serve to complement each other in the same way that the Pictures A, B and C do, each conveying information about the properties of the matter it represents. This is similarly noted by L1-01:

> And here (Pictures E, F and G) these are talking more, I think, this demonstrates more movements rather than just the arrangement. So you can easily (pointing at Picture E) see squash it, does it flow, it wouldn't really flow. Whereas here (pointing at Picture F), you can see the flow and this (pointing at Picture G) demonstrates it.

One of the participants also noted positively the use of visual cues, i.e. colours, to distinguish the three different matters:

> I think it's very good that they put one colour for solid, another for liquid and another for gas, so you can notice it. [L2-02]

Additionally, within each PD (i.e. Pictures E, F and G), four distinct SDs can be seen. These SDs complement each other as they convey different information about the property of the matter represented. Since each matter has multiple properties, it is better to distribute this information over representations rather than having them all combined into one very complex representation. L2-01 also noted this complementary relationship:

> And as you can see (in Picture E), because I didn't focus more on the arrangement of the particles (i.e. the first SD in Picture E) because (so) I just keep this one and also this one (referring to the remaining SDs the picture) because I already know the particles.

Sections 1, 2 and 3 are not isolated sections. The concepts introduced in Section 1 serve as a basis for understanding the next sections and all sections combined contribute to an understanding of the domain. Arguably, this is also applicable to the pictures belonging to each section. For instance, we know that Section 1 talks about the arrangement of particles and this information is depicted in its Pictures A, B and C. We also know that this arrangement determines the unique properties of each matter. Hence understanding the arrangement of particles is the foundation of understanding the properties. These properties are introduced briefly in Section 1 and elaborated in Section 3. This elaboration is further depicted in Pictures E, F and G in Section 3. Arguably, the simpler Pictures A, B and C (or the SDs) are intended to prime the understanding of the more complex Pictures E, F and G (or the PDs) in the same way. As such, Pictures A, B and C constrain the interpretation of Pictures E, F and G.

However, L1-01 commented that the SDs and PDs are complementary as they each contain different information about the particles.

> … here (SDs), it's quite simple and you're talking about the arrangements... And here (PDs) these are talking more, I think, this demonstrates more movements rather than just the arrangement. So you can easily (pointing at the first PD) see squash it, does it flow, it wouldn't really flow… So this (pointing at the PDs) is very important but this (pointing at the SDs) is good because it's simple. Well for me anyway, the concept that I needed from that (the SDs), those pictures, was the fact that the arrangements are close or farther apart. And here (pointing at the PDs), this is about the movement. No, I think they complement each other.

Despite having understood the concepts of particles, their arrangements and unique properties, they are not exactly concrete and easily relatable. Pictures A, B and C in Section 1 and Pictures E, F and G in Section 3 depict these concepts merely as entities composing different coloured circles with different arrangements and properties. Section 2 helps to make concrete these abstract concepts by utilising a real-life object more relatable to the everyday life, i.e. the rubbish mover machine. Solids, the arrangement of their particles and unique properties are depicted by the grabber/ shovel of the machine, liquids and their particles by that which is contained in the machine's tube and gases and their particles by that contained in the tyres. These real-life objects are more familiar to the readers and support their understanding of the abstract particle theory and therefore serve a constraining function to the SDs and PDs.

Such constraining relationship between the representations can be inferred from L1-02 and L2-03:

> I think to use the illustration of this machine is a great idea, isn't it? Because you got the fact that the tyre is filled up with gas, solid. Illustrates in a real-world context what the theoretical stuff is. [L1-02]

> And then this one (pointing to Picture D), this picture helped me a lot, I think. Like to find which one is the solid one, to find which one is the gases and then the liquid ones also. [L2-03]

Another participant, L1-01, however felt differently. She felt that the application of the theory difficult and Picture D not helpful for her. She therefore consulted Pictures E, F and G that she already looked at previously to understand this concept better. Hence, Pictures E, F and G were the representations doing the constraining.

> This (Content 5) was a bit, OK, this I think, is new information, (about) the machine. So it's more, it's applying the theory now, so I think I had to like, "Hang on". But then I jumped to the picture (Picture D) but that picture doesn't seem to be very helpful, so I've gone back to the diagrams (Pictures E and F)… I kind of fixated back on the diagram. So this is really the application of the theory ... Obviously, I find this maybe slightly more difficult than the basic theory.

Ainsworth (2018) also notes the importance of considering representations based on its suitability for a specific task. The material as highlighted in a previous section contains questions spread throughout the two pages that participants can attempt (mentally) as they proceed through the lesson. All answers to these questions are explicitly stated in the material, either within just one sentence or requiring readers to integrate information across multiple sentences or multiple representations. As such, it is interesting to note how participants coordinated information between representations to attempt the questions. They generally suggested that they used the representations in a complementary manner:

> I realised that some of these were sort of reflective questions used obviously for the students to "Explain why the grabber of the rubbish mover is made of a solid." And by looking at the imagery, it can more or less help the learner understand in a kind of multi-sensory fashion, I suppose because you got an image and you got text. [L1-02]

> Note there (fixations on Question 6 then Picture D, then back to Question 6) I'm kind of fixating on the idea of the gas not being able to be used in the tube so I went back to the diagram (Picture G then Picture F). I come back up scanning the text (Content 7 then Content 8) looking for that relevant information again. This issue between why gas can't be used in the tube itself. So I think I was doing a bit of reiterating, checking myself, I got the concepts clear. [L1-02]

> In the questions (Question 5), I tried to figure out the answer by looking at the text (pointing to Content 8) again and also the diagrams (pointing to Picture F). [L2-01]

It can be concluded from both the CRPs and interviews that the participants reported that the multiple representations used in the material generally supported their learning by serving the complementary and constraining functions. Since the topic is not particularly demanding, it appears that there is no need to combine and integrate the representations towards the construction of a deeper understanding of the domain (i.e. the constructing function). However, this absence does not negate the fact that the multiple representations, based on the researcher and the participants' own accounts, are indeed functional to the learning of the material.

3.5.2. Segmentation of Areas of Interest

This section explains how the multi-representational stimulus was segmented in preparation for data analysis. Due to its multi-representational nature, the stimulus was segmented into multiple areas of interest (AOIs). An AOI is an area of the visual stimulus for which eye movement measures are determined and eye movement data are gathered (Holmqvist et al., 2011; Scheiter & Eitel, 2017). In the present research, the AOIs were created using Tobii Studio's Areas of Interest tool (Tobii Studio version 3.4.8). In total, 28 AOIs were created; 21 text and 7 picture AOIs. The classification of text and pictures has been explained earlier (see Section 2.2). The remaining parts of the stimulus that were not covered by AOIs will be referred to as whitespace. The AOI positions were drawn before data collection commenced but some adjustments were necessary because there were small spatial distances between semantic units in certain parts of the stimulus that made it difficult to segment AOIs accurately. Moreover, additional margins had to be included in the original AOIs to account for slight eye tracking inaccuracies. For the record, certain parts of the original scan of the material had been stretched to allow for more spacing between AOIs. This was necessary to ensure that the eye tracker can distinguish between the different AOIs and hence provide accurate analyses of each AOI. Care was taken to ensure that the final stimulus did not differ substantially from the original scan. Figure 8 visualises the reading material with its various AOIs.

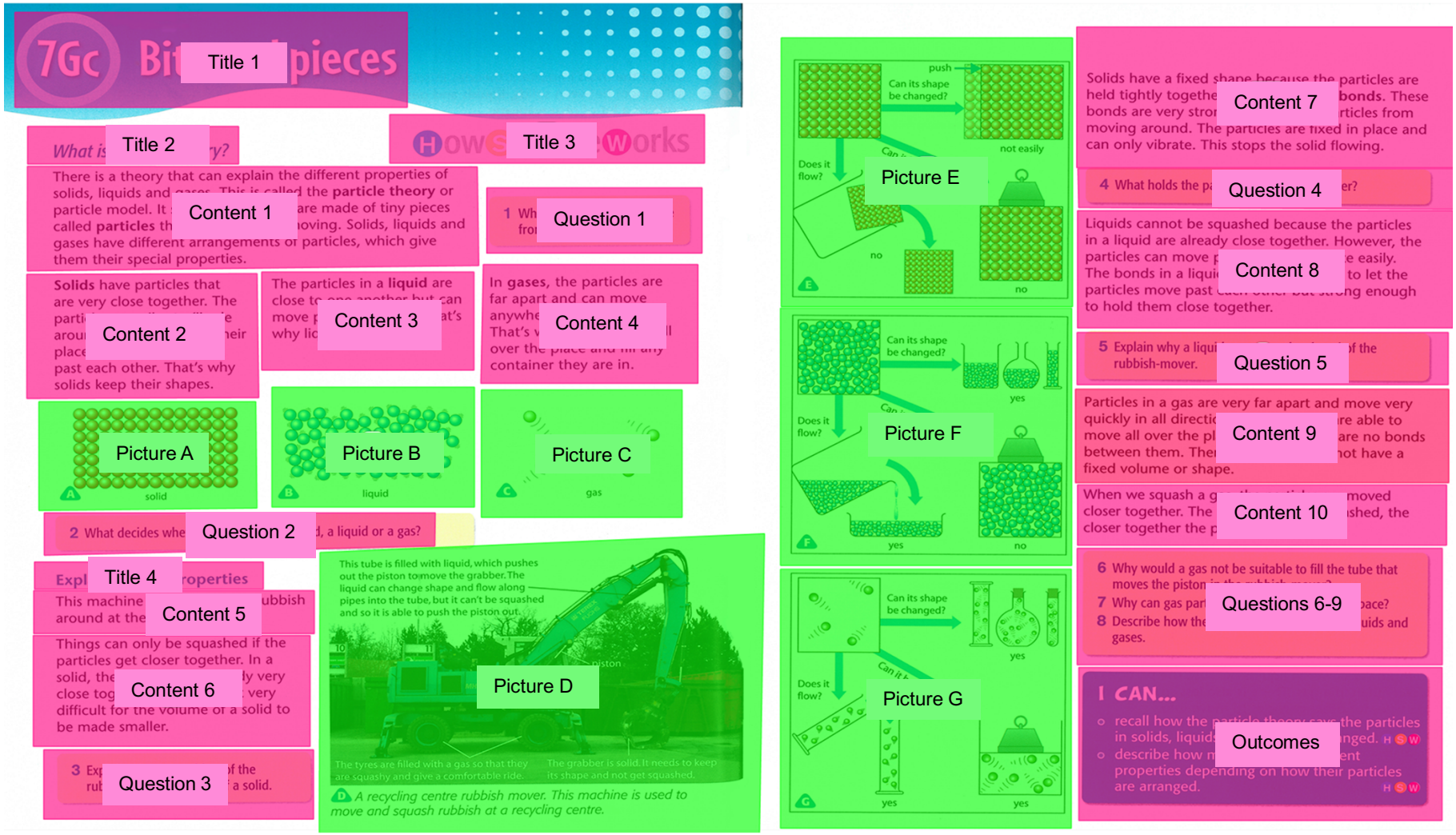


Figure 8. Segmentation of the material into AOIs. Pink AOIs are text AOIs whereas green AOIs are picture AOIs

3.6. METHODS OF DATA COLLECTION

3.6.1. Eye Tracking

This section introduces the eye tracking method, the primary data collection tool used in the research. It also highlights the unique features of the method and justifies how it is relevant and valuable in the context of the research in contrast to other data collection tools. This section also discusses the limitations of eye tracking and describes in detail the eye movement measures used in the research.

3.6.1.1. Introduction

Eye tracking refers to the recording of eye gaze and movements during an examination of a visual stimulus through the use of an eye tracker (Collewijn, 1999). An eye tracker generally consists of a light source (e.g. an infrared light) which is positioned in front of a viewer (remote eye tracking) and directed towards the eye, and a camera that tracks the reflection of the light source (EyeTracking, 2011). It records the position of the viewer's eyes as they move across a visual stimulus, thus allowing detailed examination of visual behaviour.

A basic record of eye movements is made up of a sequence of fixations and saccades (Scheiter et al., 2009). A fixation is the period of time in which the eyes remain still as they focus on and encode new information from a certain area of a visual stimulus (Rayner, 2009; Scheiter & Eitel, 2017). A saccade on the other hand refers to the rapid movement of the eyes between two successive fixations (Lai et al., 2013). Even though the eyes do not encode new information during saccades, they are fundamental due to the rapid decline in visual acuity or resolution away from the foveal region of the eye (Liversedge & Findlay, 2000; Rayner, 2009; Scheiter & Eitel, 2017). Vision in the fovea is best and saccades allow the positioning of the fovea on other parts of visual stimulus (Irwin, 2004; Rayner, 2009).

Eye tracking is frequently used in cognitive psychology because eye movement data provide moment-to-moment indication of cognitive processes underlying complex cognitive tasks such as reading and visual search (Rayner, 1998, 2009). This link between eye movements and cognitive processes is illustrated in Just and Carpenter's (1980) immediacy and eye-mind assumptions. These assumptions,

derived from a study on reading comprehension, propose that during reading readers attempt to interpret content words as they are encountered and this interpretation takes place immediately, without delay. The eyes continue to fixate on a particular word for as long as it is being processed or "until they have processed it as far as they can," (p. 350). Fixations on a visual input hence are an indicator of what the mind is currently processing. The length of these fixations is an indicator of how long the cognitive processing takes place. Eye tracking measures or indices therefore provide spatial and temporal resolution of visual attention which are both quantitative and objective (Scheiter et al., 2009).

The above-mentioned features make eye tracking very valuable in the field of multi-representations where researchers and learners themselves deal with complex materials, comprising more than one external representation. In such cases, examining how learners distribute their visual attention among the many representations during learning may be of interest. Since eye tracking measures provide spatial and temporal resolution of visual attention, the method can therefore supply the relevant information with regard to location and duration of visual attention and questions such as which "representations are visually attended to, in what order, and for how long" (van Gog & Scheiter, 2010, p. 95) can be addressed.

In addition to its use in examining distribution of visual attention across multi-representational materials, eye tracking has also been used extensively to provide insights into the cognitive processing during the learning process (Hyönä, 2010; Scheiter & van Gog, 2009; van Gog & Scheiter, 2010). In such research, the eye-mind and immediacy assumptions are particularly relevant as the spatial and temporal information about fixations and saccades, serves to inform researchers about the cognitive processes that occur during learning, rather than just as indicators of visual attention. At times, these measures are treated "as evidence for when, how long, and in which order externally presented information is processed at the cognitive level" (Scheiter & Eitel, 2017). In these process-oriented studies, eye tracking is commonly applied to investigate the three cognitive processes assumed to take place during learning with multiple representations; the selection, organisation and integration of representations (Mason, Tornatora, et al., 2013; Scheiter & Eitel, 2017).

3.6.1.2. Relevance and limitations of eye tracking

This section will justify the advantages of using the eye tracking methodology in this research as opposed to other techniques. The shortcomings of the methodology are also addressed. Earlier research that examined cognitive processing during reading has mostly utilised the verbal protocol methodology which at the most basic level asks participants to verbalise or externalise their thoughts while reading as speech (Holmqvist et al., 2011; Mason, Tornatora, et al., 2013; Stieff et al., 2011). There are various methods for collecting verbal data but this section will only address the think-aloud approaches which have two forms; concurrent and retrospective. Concurrent and retrospective think-alouds are different in terms of when the verbalisation takes place. The data from concurrent think-aloud approach are elicited during reading whilst those from retrospective approach are elicited after reading. Even though both approaches produce rich amounts of data, the validity and reliability of think-aloud methodology has been challenged. Concurrent think-aloud for instance is seen as intrusive, capable of affecting and slowing down the primary task (Holmqvist et al., 2011; Mason, Tornatora, et al., 2013). Retrospective think-aloud on the other hand is susceptible to participants' inability to recall their thoughts and fabrication of details. It may not slow down the learning process but verbalising after reading may take twice as long as verbalising during reading (Holmqvist et al., 2011). Eye tracking overcomes these limitations well. In the context of the present research, eye tracking measures have already recorded relevant information on participants' perceptual processing, hence their recollections of the reading process were not necessary. The participants were therefore relieved of the secondary think aloud task and could concentrate instead on the primary task; reading. The reading task was left undisturbed and reading behaviour for the most part was authentic and natural. Moreover, since the participants' eyes have done 'all the talking', and data elicited were objective, they were not liable to memory decay or fabrication. These advantages also mean that the reading process and procedural time were not prolonged.

In terms of the nature of the verbal data, Stieff et al. (2011) highlight several weaknesses. They pertain to its incomplete and limited nature when verbalising thought processes for tasks that involve multiple representations. Firstly, the data can be incomplete as participants can finish the task without referring to any particular representation during verbalisation. Additionally, even though participants do refer to a particular representation when verbalising, this does not necessarily indicate that it

is useful for the primary task. The verbal data can also be limited as they are only reflective of representations that participants are conscious of using when they may also attend to those that they do not refer to in their verbalisation. Eye tracking overcomes the first and third limitations as the data provide information regarding where participants are attending to in the stimulus, whether or not they are conscious of doing so. However, the second limitation is also shared by the eye tracking data. The limitations of eye tracking are discussed next.

Despite its relevance and value to the research, eye tracking like every other research tool, also has its shortcomings. These are discussed in the context of the present research. Perhaps one of its major shortcomings is that its data are not self-explanatory and therefore information about participants' cognitive processes employ during learning are incomplete (Scheiter & van Gog, 2009; Stieff et al., 2011). The data can only offer information about which representations participants attend to but not necessarily why they are attended to or how they are used (Stieff et al., 2011). Hence, researchers may find themselves involved in substantial amount of inferences and interpretations to disambiguate the underlying cognitive processes behind certain eye tracking measures (Scheiter & van Gog, 2009; Schwonke et al., 2009). For instance, if a participant attended to a particular representation for a long time, she may have done so for a variety of different reasons: she might have found it confusing, interesting, or particularly useful for the task at hand. This limitation did affect the present research to a certain extent. Since the L2 readers' experience with multi-representational materials is underexplored, insights into why and how they use the representations could provide a very valuable and comprehensive account of their reading behaviour. This information could further enrich the present findings and conclusions regarding differences and similarities of these two groups of readers. Earlier in the thesis, assumptions about how reading of a multi-representational material could be similar or different in L1 and L2 were outlined. In both cases, the possibility that the L2 readers could be experiencing L2 reading difficulties was put forward as a potential factor. The eye tracking data however, can only inform whether L2 reading is similar or different to L1 reading in that it can show that L2 reading for instance is similarly text-dominant like L1 reading. The data however cannot verify that the reason behind this similarity is indeed due to language difficulties that the L2 readers might have experienced during learning.

To alleviate ambiguous interpretations, Scheiter and Eitel (2017) suggest interpreting eye tracking data simply as indicators for participants' visual attention, rather than for

cognitive processing. In other words, the data should be strictly used for what they offer and inform, i.e. the objective measurements of the locations, the durations and the order of visual attention, and not for inferring the rationale behind visual attention. Should researchers proceed with the latter, eye tracking data should be complemented with performance or process measures to arrive at more accurate interpretations of underlying cognitive processing (Hyönä, 2010; Scheiter & Eitel, 2017). This is indeed the standard practice in the literature as researchers typically supplement the online eye tracking data with offline performance measures such as learning outcomes. Such complementary approach not only addresses this limitation but also the second limitation; that eye tracking data cannot inform researchers the end product of visual attention (Hyönä, 2010). To elaborate, even though the eye tracking data can inform researchers about how learners attended to relevant information during learning, it does not enable researchers to conclude that the learners have understood and learned that information well. Simply attending to task-relevant information does not mean successful comprehension. Complementing the data with performance measures therefore helps researchers to make informed decisions about the various eye tracking measures and how they implicate or relate to learning. In line with this suggestion, the present research complemented the data with an offline performance measure; the learning outcome test and related the two together. Acknowledging that other factors such as individual characteristics may moderate this relationship, offline measures of relevant individual characteristics (e.g. English proficiency and prior knowledge) were also supplemented (Scheiter & Eitel, 2017).

However, complementing eye tracking data in such a way still does not provide justifications for why the present research participants behaved in the way they did. If the L2 readers made frequent integrative transitions between text and picture areas of the multi-representational material and they performed well in the learning outcome test, this only means that frequent text-picture transitions are associated with successful learning. It does not automatically mean that the participants did indeed integrate the textual and pictorial information. In this case, a more holistic approach is to supplement the eye tracking data with verbal data. In such form of methodological triangulation, verbal data commonly serves as an information source on learners' cognitive processing, whereas eye tracking data as an information source on their perceptual processing (Holmqvist et al., 2011). Verbal data can disambiguate the interpretations of eye tracking data as participants' own accounts of experienced cognitive processes while performing a task are obtained (Holmqvist et al., 2011;

Schwonke et al., 2009). Unfortunately, this was not possible for the present research, given that the participants would have to verbalise in English, their L2. Requiring them to provide protocols in their L2 may produce incomplete or invalid data as they could potentially withhold from expressing their thoughts rendering their data incomplete. Alternatively, they could confabulate information to compensate for their inability to express their thoughts in the right vocabulary. Both possibilities could jeopardise the validity of the data and render the combination of eye tracking and verbal data counterproductive and eliminate its strengths. Indeed, Gass and Mackey (2000) highlight that one of the problems that L2 researchers continuously face when conducting CRPs in participants' L2 is that the participants have a limited ability to express themselves in the language. This may consequently impact the content of their verbal reports as rather than verbalising a full account of their thought processes, they may end up verbalising only what they can.

Thirdly, the eye-mind assumption fundamental to the cognitive interpretation of eye movement data may not always hold true as what the eyes attend to may not always reflect cognitive processing. Irwin (2004) for instance challenges that locus of cognitive processing is not restricted to fixation locations and can include a larger area. He also argues that fixation location does not always correspond to the locus of cognitive processing because when the eyes move, the locus of cognitive processing shifts first before the eyes do. Third, the position of the eyes is not always under cognitive control as at times they move involuntarily and along with it, attention shifts and cognitive processing takes place. Fourth, cognitive processing may also occur during saccades, not just during fixations. However, despite criticisms against the eye-mind assumptions, most researchers agree that eye movements can indeed be a good indicator of current attentional processes when the available visual environment is relevant and pertinent to the task (Hyönä, 2010; Scheiter & Eitel, 2017) and when participants have the freedom of looking anywhere on the stimulus (Stieff et al., 2011). This suggests that before research participants learn and attend to a visual stimulus, they must be explicitly informed about why they need to do it in the first place and what kind of task follows.

The technique is also liable to produce low quality data due to various reasons such as calibration errors, low precision, participant-specific tracking difficulties or offsets between the true and recorded gaze positions (Holmqvist et al., 2011; Stieff et al., 2011). It is therefore not unusual for researchers to need to exclude a significant amount of data when their quality undermines its suitability for analyses. Holmqvist et

al. (2011) note that they usually had to exclude 2 to 5% of their data. Data loss remains the most challenging issue encountered in the present research with up to 30% data loss occurring. The decision to use a remote eye tracker without a head stabilisation system might have contributed about this issue but the nature of the research did not require such restriction, which is more commonly associated with text reading research. In addition, using authentic material rather than single isolated simpler stimuli provided a greater opportunity for such loss to occur.

3.6.1.3. Eye tracking measures

A number of eye tracking measures were computed to analyse how the L1 and L2 readers read the multi-representational text (refer to

Table *4* for the measures used in this thesis). This combination of measures, with the exception of the look-from picture to picture fixation time (or LF-Pict Pict) was modelled from a very similar study by Mason, Tornatora, et al. (2013). In that study, the measures were used to examine the reading behaviour of young native Italian readers when learning from a simpler illustrated science material on the topic of air. The study has successfully found a significant link between these measures and learning from the material and in a more recent study (Mason, Tornatora, et al., 2015), these findings were replicated. Considering these successful applications, the present research therefore used the study and its measures as a model to analyse the reading behaviour of different populations of readers (i.e. readers of English as an L1 and an L2) when learning from a different material, i.e. an authentic material with multiple graphic representations. Since this material has more than one graphic representation, unlike Mason, Tornatora, et al.'s (2013), the LF-Pict Pict was computed as well.

The two basic measures are fixation count and fixation time. Fixation count refers to the number of fixations on an AOI, whereas fixation time refers to the duration of these fixations. These two form the basis of the remaining measures. Among them are those concerning first-pass processing which provide an insight into readers' behaviour when they first encounter the material. These measures are first-pass fixation time on pictures (or FP-Pict) and first-pass fixation time on text (or FP-Text). They refer to the summed duration of all fixations on picture or text AOIs during initial inspection or reading, before exiting the AOIs (Mason, Tornatora, et al., 2013). These measures,

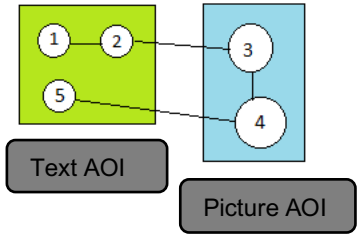
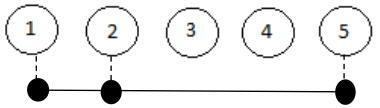
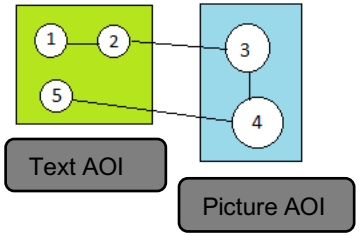
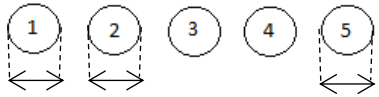
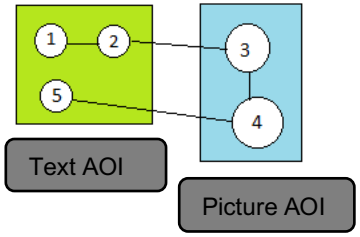
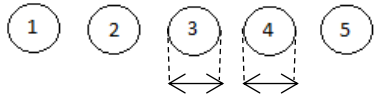
FP-Text in particular, are frequently linked with basic reading processes and comprehension difficulties (Mason, Tornatora, et al., 2015; Scheiter & Eitel, 2017).

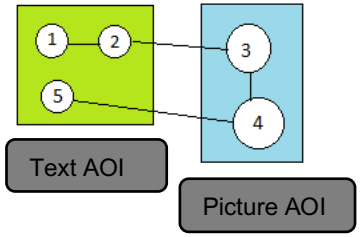
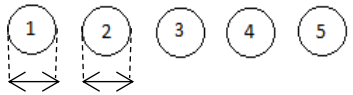
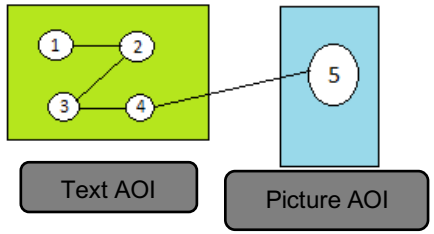
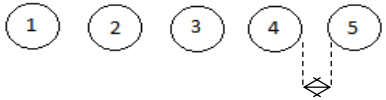
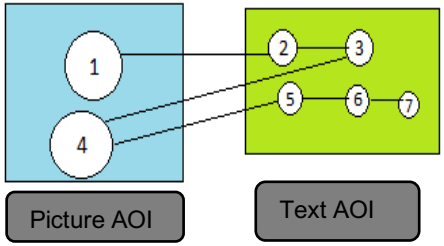
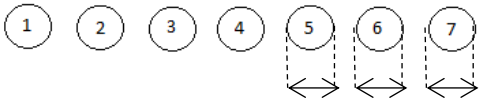
It is essential in this research to understand how participants integrate verbal and pictorial elements in the text. Hence, several text-picture integrative measures or integrations across representation types were collected. They were computed based on the saccades or gaze shifts made between text and picture AOIs (Scheiter & Eitel, 2017). Integrative transitions (or transitions) refers to the saccade from a text AOI to a picture AOI or vice versa. It is a frequency measure indicating the number of times the readers attempt to integrate text and pictures (Johnson & Mayer, 2012; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013; Scheiter & Eitel, 2017). Two temporal measures were also collected to examine the total time spent on text-picture integration. Look-from picture to text fixation time (or LF-Pict Text) refers to the summed duration of all refixations on a text AOIs during re-inspection of a picture AOI, whereas look-from text to picture fixation time (or LF-Text Pict) refers to the summed duration of all refixations on a picture AOI during re-reading of a text AOI (Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013). Both LF-Pict Text and LF-Text Pict are used to indicate "the extent to which a text segment is used as an anchor point for processing the picture segments or a picture segment is used as an anchor point for processing the text segments" (Mason, Tornatora, et al., 2015, p. 859). For LF-Pict Text, it is assumed that a reader is trying to integrate the verbal information she is currently re-reading with the pictorial information she has just re-inspected. The opposite is true for LF-Text Pict.

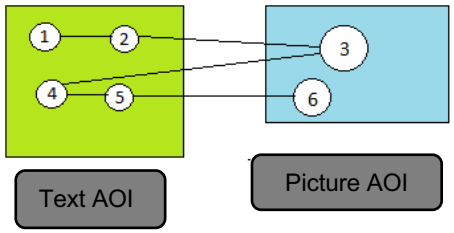
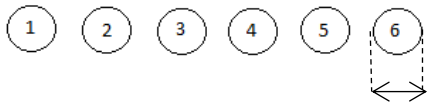
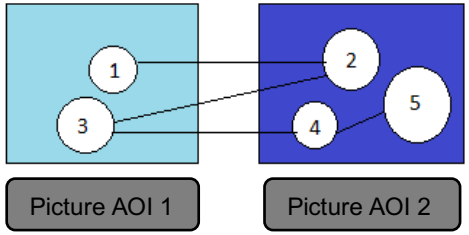
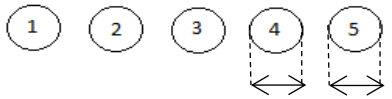
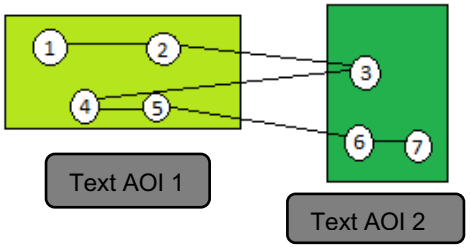
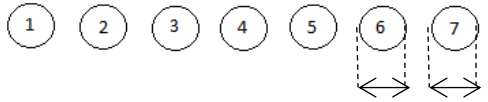
In addition to examining text-picture integrations, integrations within the same representation types were also examined. This refers to integrations made across multiple text AOIs or across multiple picture AOIs. For this type of integration, two temporal measures were computed; look-from picture to picture fixation time (or LF-Pict Pict) and look-from text to text fixation time (or LF-Text Text). LF-Pict Pict refers to the summed duration of all refixations on a picture AOI during re-inspection of another picture AOI whereas LF-Text Text refers to the summed duration of all refixations on a text AOI during re-reading of another text AOI. All look-from fixation time measures (i.e. LF-Pict Text, LF-Text Pict, LF-Pict Pict and LF-Text Text) are second-pass processing measures. They are in fact fine-grained second-pass measures, given that readers need to make at least two visits in both target areas included in each measure. Second-pass processing is argued to reflect a more intentional and purposeful processing of materials as opposed to the initial and more

automatic first-pass processing (Mason, Tornatora, et al., 2015; Scheiter & Eitel, 2017). Second pass processing is therefore associated with more intentional reading behaviour such as attempting to integrate information across AOIs and to resolve comprehension difficulties. Please note that not all second-pass temporal measures were examined in this thesis. Only the finer-grained second pass measures, i.e. the look-from fixation time measures mentioned above were examined as they have been shown in Mason, Tornatora, et al.'s (2015; 2013) studies to influence cluster memberships and learning outcomes. Figure 9 provides an example of how each of these measures could occur during the viewing of the stimulus.

Table 4. Eye tracking measures and their definition largely based on Mason, Tornatora, et al.'s (2013) study

	Eye tracking measure	Definition	Illustration	Specification
1	Fixation count	Total number of fixations on an AOI		 Total fixation count on Text AOI = 3 (fixations 1 + 2 + 5)
2	Fixation time	Total duration of fixations on an AOI		 Total fixation time on Text AOI = duration fixations 1 + 2 + 5
3	First-pass fixation time on picture (FP-Pict)	Total duration of all fixations on picture AOIs during initial inspection		 FP-Pict = duration fixations 3 + 4

4	First-pass fixation time on text (FP-Text)	Total duration of all fixations on text AOIs during initial reading		 FP-Text = duration fixations 1 + 2
5	Integrative transitions (transitions)	The number of gaze shifts or saccades from a text AOI to a picture AOI and vice versa		 Transition
6	Look-from picture to text fixation time (LF-Pict Text)	The duration of all regressive fixations on a text AOI during re-inspection of a picture AOI		 LF-Pict Text = duration fixations 5 + 6 + 7

7	Look-from text to picture fixation time (LF-Text Pict)	The duration of all regressive fixations on a picture AOI during re-reading of a text AOI		 LF-Text Pict = duration fixation 6
8	Look-from picture to picture fixation time (LF-Pict Pict)	The duration of all regressive fixations on a picture AOI during re-inspection of another picture AOI		 LF-Pict Pict = duration fixations 4 + 5
9	Look-from text to text fixation time (LF-Text Text)	The duration of all regressive fixations on a text AOI during re-reading of another text AOI		 LF-Text Text = duration fixations 6 + 7

Note:

1. All temporal measures (i.e. all first-pass and look-from fixation time measures) do not add up to 100% in terms of proportions because fixations on whitespace and other irrelevant measures are not considered
2. All look-from fixation time measures are second-pass processing measures requiring a minimum of two visits in each AOI

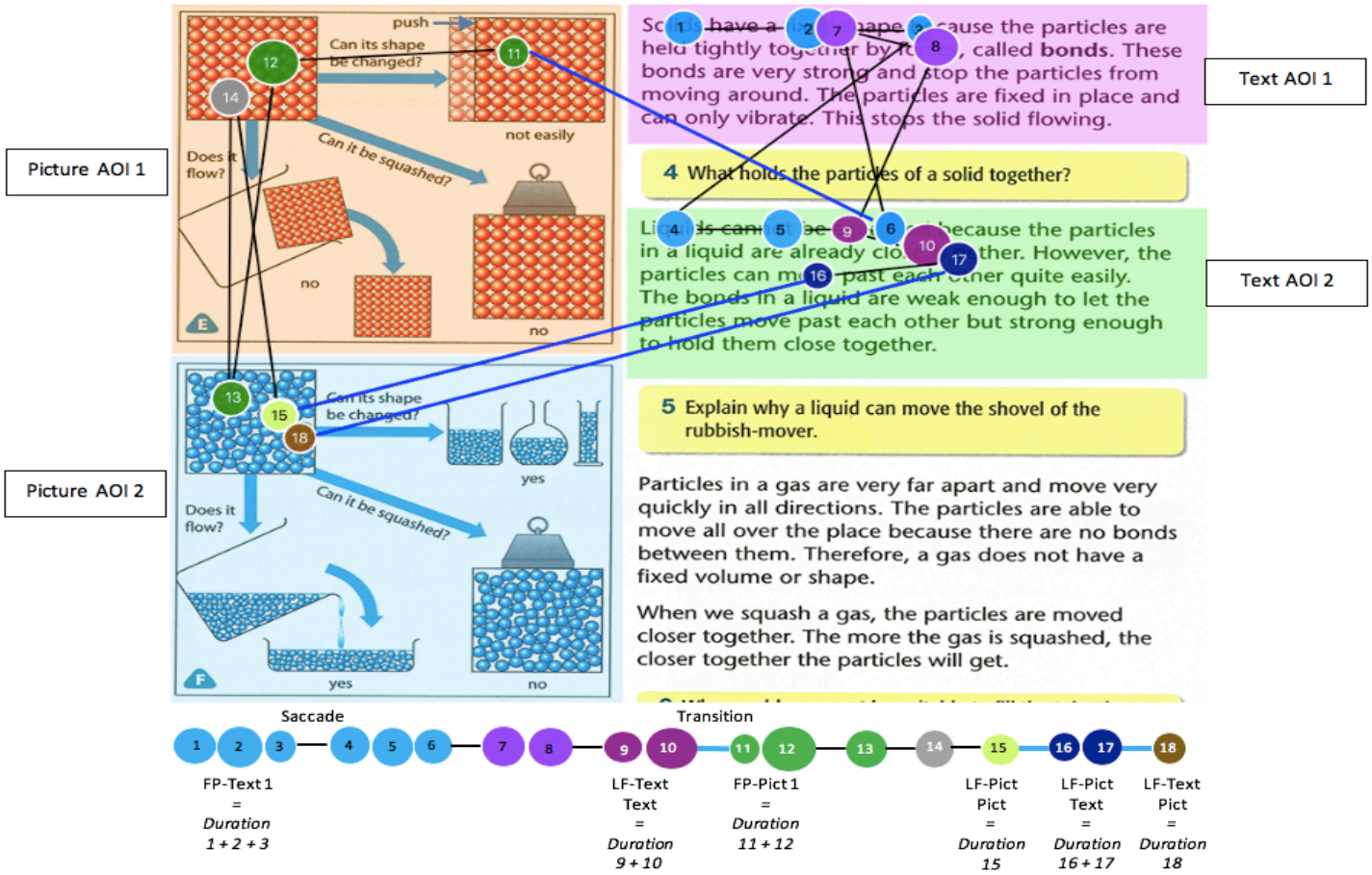


Figure 9. An example of the various eye tracking measures on the stimulus and labelling of the eye tracking measures

3.6.1.4. Apparatus

Eye tracking data were collected using a Tobii Pro TX300 eye tracker (Tobii Technology, Stockholm, Sweden) with a 300 Hz sampling rate. This remote eye tracker is not intrusive and does not require a head stabilisation system. The participants were therefore allowed to make minimal head movements. The stimulus was presented on a 23-inch TFT monitor with a resolution of 1920 x 1080 pixels. Data were recorded and analysed with Tobii Studio software version 3.4.8.

3.6.2. Tests

In addition to eye movements, measures of individual characteristics and learning outcome were also collected. These necessitated the use of tests, both existing ones and adapted, to assess these variables. This section introduces and describes the tests used in this research.

3.6.2.1. English vocabulary knowledge

The Lexical Test for Advanced Learners of English (LexTALE) was used to measure the L2 readers' English vocabulary knowledge (Lemhöfer & Broersma, 2012). LexTALE assesses the vocabulary knowledge of medium to highly proficient English as an L2 speakers in an experimental setting. The test involves an un-speeded yes/no visual lexical decision task in which each participant was presented with a random list of 40 words and 20 nonwords and had to decide whether or not each letter string was an existing English word. The words are low frequency, whereas the nonwords are orthographically legal in English and pronounceable. The way that LexTALE is scored is such as that test-takers are penalised against yes and no-biases. The maximum score is 100. LexTALE was chosen because it is valid, quick, and easy to administer. It was tested on native Korean and native Dutch speakers who speak English as an L2 in a large-scale study (Lemhöfer & Broersma, 2012). In this study, LexTALE was shown to be a good predictor of the English vocabulary knowledge because it correlated highly with another measure of English proficiency (i.e. the Quick Placement Test) and it was superior to a self-ratings measure of English proficiency. Since LexTALE has worked well with the Korean and Dutch speakers with very distant L1s, the findings are assumed to generalise to English L2 users with different L1s. This is particularly relevant for the present research as the L2 readers recruited had

diverse L1 backgrounds. Additionally, LexTALE takes users about 3.5 minutes to complete. It is also easy to administer as it can be administered online.

3.6.2.2. English reading comprehension ability

The L1 and L2 readers' English reading comprehension ability was assessed with the Gray Silent Reading Tests (GSRT) which was normed on 1400 individuals in the USA between 1996 and 1998 (Wiederholt & Blalock, 2000a). The GSRT has two parallel forms (Forms A and B). Form B was used in both studies. It contains 13 developmentally sequenced stories. Each story is a paragraph long and has five multiple-choice comprehension questions. Each question has four options with only one correct answer. The test was administered and scored using the basal and ceiling formats outlined in the GSRT Examiner's Manual (Wiederholt & Blalock, 2000b). A basal was established when all five questions for a story were correctly answered, whereas a ceiling was reached when only two questions were correctly answered. The maximum raw score is 65. The reported Cronbach's reliability coefficient for Form B is .94.

3.6.2.3. Prior knowledge

The L1 and L2 readers' prior scientific knowledge of the properties of solids, liquids and gases was assessed with a paper-and-pencil test comprising 10 multiple-choice items. Each has four options with only one correct answer. The maximum score is 10. All items were adopted from a list of assessment items available on the American Association for the Advancement of Science (AAAS) Project 2061 Science Assessment website under the topic of Atoms, Molecules and States of Matter ("American association for the advancement of science," 2016). All assessment items on the website have undergone a rigorous developmental process, piloted and field-tested. To select the most suitable items for the prior knowledge test, the researcher checked whether they aligned to the content and learning outcomes of the reading material. This resulted in the decision to replace the word 'molecule' in the AAAS items to 'particle' as this is the term used throughout the reading material, and to avoid misleading the participants. This decision was checked with a chemistry teacher and considered appropriate as it is unlikely that false outcomes will materialise from the interpretation of the word 'particle' in the items. An example of the items and its answer options can be found below. The correct answer is italicised. The complete test is provided in Appendix 4: Prior Knowledge Task.

Question 1: In which state of matter are the particles spaced farthest apart?

A. A gas

B. A liquid

C. A solid

D. All are equal.

3.6.2.4. Learning outcome and coding

The measure of learning outcome was a paper-and-pencil task that assessed the participants' recall and comprehension of the reading material. It comprised eight open-ended items which were reproduced from the question section in the reading material. The last item however was slightly different from the original. In addition to describing "the movements of the particles in solids, liquids and gases" in just words, participants were also asked to describe them through drawing(s). This additional element assesses the participants' recall and comprehension of the pictorial information, in particular. All answers to the open-ended items are explicitly stated in the material, either within just one sentence or requiring readers to integrate information across larger distances, such as across multiple sentences or multiple representations (i.e. text and picture).

Since the test was open-ended, a scoring rubric was produced to score the answers. Answers to all eight items were scored according to their correctness and completeness to provide a maximum score of 30. The total score of the first seven items was 14. The items differed in complexity and number of main points. Thus, each item was answerable for different maximum scores. The total score of the final question was 16. This part of the test asked participants to describe the movements of solid, liquid and gas particles and to draw these particles (i.e. how they are arranged). Description and drawing of solid and liquid particles were scored between 0 to 5 each, whereas the description and drawing of gas particles were scored between 0 to 6. (See Appendix 5: Learning Outcome Task for the test items, and Appendix 6: Scoring Rubrics for Learning Outcome Task). The author scored all the participants' learning outcome tests and two independent raters (one rater for each study) scored a subset of seven randomly selected tests for each of the two studies. The intra-class correlation coefficients (a mixed model ICC, absolute type) for Study 1 was .98, and for Study 2, .83, suggesting a high interrater reliability.

The decision to use the existing items to assess the participants' learning from the text was made as it was felt that the textbook authors had designed these items in order to match the content of the reading material most appropriately. Hence, developing a new and different set of questions could have missed the original assessment goal intended for this material. An example of the item and its scoring can be found below.

Question 3: Explain why the grabber of the rubbish-mover is made of a solid.

1 point was awarded for each of the following main points:
- a solid can keep its shape;
- a solid cannot get squashed

3.6.3. Questionnaire

A questionnaire was administered to the L1 and L2 readers to collect their background information and to ensure that they met the inclusion criteria for each study. The questionnaire administered to the L2 readers consists of two sections: language profile and science education profile. For language profiling, information on the L2 readers' language repertoire, English acquisition, education and ability, language of instruction in school, and frequency of English use in everyday activities was collected. The science education profile gathered data about the type of science education (i.e. what science subjects) the participants had undergone previously, the language of instruction used to teach the subjects, when they last studied the subjects and their highest qualification for the subjects. Additionally, information on the participants' clarity of vision and reading disabilities were also collected. The background questionnaire administered to the L1 readers was much shorter as the full language profile section was removed as they only needed to state what their additional language(s) were, if any. Sections on science education profile, clarity of vision and reading disabilities were retained.

3.7. METHODS OF DATA ANALYSIS

3.7.1. Preparation for Data Analysis

After the data collection stage, all data (i.e. eye movements, individual characteristics, learning outcome) were first explored to prepare them for analysis. This step allows identification of potential errors and issues in the subsequent analysis stage and formulation of decisions to overcome them so that they can be later justified (Holmqvist et al., 2011). For eye tracking research, data quality is of utmost importance. Data were first inspected using scanpath plots. Offsets and drifts were identified, and participants were excluded when this value exceeded five percent. Data quality notes taken during recordings were also useful to support these decisions. During eye tracking, researchers normally have a real-time view of participants' eye movements and therefore can note, albeit not very accurately, the quality of the data. When participants passed this initial stage, their data were first trimmed before submitted to statistical analyses. This step was necessary to remove any incorrectly identified information. Firstly, all fixations that lasted shorter than 100 ms were discarded from analyses given the argument that it is unlikely to encode adequate information from a visual display within such a short time frame (Hegarty & Just, 1993). In practice, short fixations are usually discarded given that the brain requires time to encode visual information. Secondly, each participant's first fixation on the stimulus page was also excluded since its presentation was preceded by a fixation cross.

During data exploration, the distribution of data and presence of outliers were also inspected. Many eye tracking measures are not normally distributed and tend to be positively skewed (Holmqvist et al., 2011). This is also the case for most of the eye tracking measures in the present research. Such issue has direct implications on analysis as the methods of analysis normally used for eye tracking data are statistical and many statistical tests require that data be normally distributed. How these issues were addressed is discussed in the following section.

3.7.2. Statistical Considerations

Parametric tests were used throughout the thesis. This set of tests is generally used as it has a greater statistical power than its non-parametric counterparts in detecting that a genuine effect exists in a given data set (Field, 2013). In this thesis, the significance criterion is set at .05. The use of parametric tests necessitates meeting various assumptions so that their results can be accepted and interpreted appropriately. Violations of these assumptions on the other hand can bias the conclusions. As indicated in the previous section, the distributions of many of the eye

tracking data used in this thesis have violated one of the assumptions of parametric tests; i.e. the normality of distribution. Hence, questions about the necessity of using parametric tests may arise at this point. Why insist on parametric tests when upon exploration of data, a potential bias has been identified? Why were the non-parametric equivalents not opted for instead when they make fewer assumptions? It is necessary to point out here that there are means to reduce these biases and correct for problems in data. Furthermore, some parametric tests are robust to violations of assumptions (Glass, Peckham, & Sanders, 1972; Rasch & Guiard, 2004) and this is certainly the case for the tests used to analyse the present data. Finally, although using the non-parametric tests was indeed an option, they are not necessarily the ideal solution. Using them entails ranking the data and hence converting them from interval/ratio to ordinal/ nominal. This consequently leads to loss of information as the magnitude of differences between numbers becomes obscured (Field, 2013; Holmqvist et al., 2011).

3.7.2.1. Handling violations of assumptions

As mentioned above, there are ways to reduce biases and one of them is through data transformation. Ideally, by transforming data, problems with distributions, outliers, linearity and equality of variances can be resolved (Field, 2013). It is not uncommon to transform data in eye tracking studies because eye tracking data tend to be positively skewed. The most commonly used transformation method is log transformation (Holmqvist et al., 2011). In addition to the log method, Field (2013) suggests two other methods that can correct for positively skewed distributions; square root and reciprocal. Out of the three, the square root transformation ($\sqrt{X_i}$) worked best with the eye tracking measures in the present research, as in most cases it corrected for distributional problems. This transformation reduces high values along the right tail of the distribution hence reducing skewness (Field, 2013; Holmqvist et al., 2011).This transformation was therefore used throughout the thesis whenever eye tracking measures did not meet the assumption of normality. Note that transformation was only applied to eye tracking data (that violated the normality assumption) and not for other data, i.e. individual characteristics, learning outcome since transforming eye tracking data is the standard practice used in the studies that this thesis closely modelled. Additionally, by applying this transformation, another potential bias in the eye tracking data, i.e. presence of outliers was reduced. Extreme outliers disappeared when this transformation was applied. Note that in this thesis, outliers were not removed as they were neither extreme nor illegitimate. Transformations undoubtedly

are not ideal for any studies because they may reduce its statistical power and have implications for the interpretation of test results (Field, 2013; Holmqvist et al., 2011).

Additionally, there are parametric tests that are robust to violations of assumptions. T-tests, and ANOVAs are some of the examples and they were used in this thesis to compare two or more means. These tests are robust against non-normal distributions. Rasch and Guiard (2004) for instance noted that "the t-test is so robust against non-normality that there is nearly no need" (p. 75) to opt for its non-parametric counterpart. With regard to ANOVA, Glass et al.'s (1972) review of literature on the robustness of ANOVA concluded that "skewed populations have very little effect on either the level of significance or the power of the fixed-effects model *F*-test" (p. 273). This statement is also applicable when the distributions are both skewed and sample sizes are unequal. Using this argument, t-test and ANOVA were the preferred methods used to compare means. In cases when transformation did not fully correct for the skewed eye tracking data, t-tests and ANOVAs were still carried out in line with this argument.

T-test was the parametric test used when comparing two means on one dependent variable. When more than two means were compared on one dependent variable, a one-way ANOVA was used. With multiple comparisons, a post hoc analysis is required to identify where significant differences between groups are and by how much. In this analysis called pairwise comparisons, all different combinations of groups are compared. To avoid inflating Type I error rate caused by multiple tests being carried out, these comparisons correct for the significance level for each test so that the overall Type I error rate across all comparisons remains at the nominal level (Field, 2013). When choosing a post hoc test to follow up on significant multiple comparisons for this research, several factors were considered. Since the cluster sizes derived from the cluster analyses were unequal (e.g. Study 1's *n*'s = 21: 11, Study 2's *n*'s = 7: 17: 9), the post hoc test of choice therefore needs to perform well under this condition. Hence, when effect of clusters on dependent variable(s) was being scrutinised, the Hochberg's GT2 post hoc test was chosen. Additionally, there were instances when the assumption of homogeneity of variance was violated. Since Hochberg's GT2 is not robust to this violation of assumption, the Games-Howell was selected as it performs well not only when variances were heterogeneous but also when the group sizes are unequal. These criteria of selection are based on Field's (2013) guidelines.

When means were compared on more than one dependent variable, MANOVA was the statistical test used. Among the four test statistics available, Pillai-Bartlett trace or Pillai's trace was selected as it is accurate with small or unequal sample sizes and controls the Type I error rate best (Field, 2013; Park, Cho, & Ki, 2009). Additionally, it is also useful when the homogeneity of covariance assumption is violated (Park et al., 2009). In the main study chapters, effect sizes in addition to *p*-values are reported to indicate the magnitude of effects. Cohen's *d* is reported for t-tests whereas partial eta squared (η_p^2) is reported for both ANOVAs and MANOVAs.

3.8. ETHICAL PROCESSES FOLLOWED

The research complied with the University of Nottingham's Code of Research Conduct and Research Ethics informed by the British Educational Research Association's Revised Ethical Guidelines for Educational Research (BERA, 2011). Ethical approval was attained from the Ethics Committee of the School of Education prior to data collection. As a laboratory research that involved the participation of volunteering University Students, the research imposed very minimal risk on the part of the participants as well as the researcher. Additionally, many of the issues encountered in educational research in more naturalistic settings are not relevant. These include gaining access to the research site and participants from gatekeepers, respecting the setting of the research site by imposing minimal intrusion as well as political influences or pressures on researchers (Creswell, 2014a; Robson & McCartan, 2016).

Upon establishing contact with potential participants, they were emailed an information sheet that briefly describes the study, its purpose and procedures. Upon their arrival in the laboratory, they were again briefed about the study purpose and procedures, this time in greater detail. They were given the opportunity to ask questions and were informed that they were free to withdraw from the research at any time. Once they agreed to participate, an informed consent form was given to be read and signed. In this form, the researcher's and her supervisors' contact details were listed along with the research purpose, procedures and tasks that the participants were asked to perform. Additionally, all participants were informed that their participation was voluntary, and they had the right to withdraw their data from the research at any time and remain anonymous. Their data will be kept confidential and anonymised, and only the researcher and her supervisors will have access to them.

Finally, as the stimulus was taken from a textbook, issue of copyright might arise. The research complies with the United Kingdom's guidelines of using copyright materials. The guidelines state an exemption from seeking the permission of the copyright owners in instances where the materials are used for non-commercial research and private studies. Details can be found on https://www.gov.uk/guidance/exceptions-to-copyright#non-commercial-research-and-private-study.

4. STUDY 1: THE L2 STUDY

4.1. INTRODUCTION

This chapter describes the first experimental study and concerns second language (L2) readers of English. The first two chapters have highlighted the importance of studying this group of readers. To summarise, the large body of literature on learning multi-representational texts has primarily focused on people reading multi-representational texts in their native or first language (L1). The literature has not sufficiently examined those reading in their L2. This is rather unfortunate given that more and more students are being educated in their non-native language. Moreover, studies comparing L1 and L2 reading have found that reading in both languages differs. L2 reading patterns for instance are usually characterised by longer sentence reading times, more fixations and shorter saccades (Cop et al., 2015). L2 readers' experience with multi-representational texts may therefore be different from their native-reader counterparts'. For example, their difficulty in comprehending unfamiliar words in their L2 could guide them to use the pictures as a source, allowing them to gain better understanding of the text (Lee & Mayer, 2015).

An exploratory study was deemed appropriate to address the above-mentioned gap. Study 1 was therefore designed to explore the ways in which learners read and learn a multi-representational text in their L2. The study utilised the eye-tracking methodology to examine the L2 reading patterns. In particular, it examines whether reading of a multi-representational text in L2 was text or picture-dominant. Additionally, it investigated whether the L2 readers exhibit distinct eye movement patterns and whether these are related to prior knowledge and aspects of L2 proficiency, i.e. reading comprehension ability and vocabulary knowledge. The relationship between eye movements and learning outcomes was also examined.

4.1.1. The Nature of Reading Multi-Representational Texts

It has been observed in Chapter 2 that reading multi-representational texts in the L1 is generally text-dominant as readers predominantly attend to text areas and sometimes fail to devote sufficient amount of attention to inspect picture areas. Hannus and Hyönä's (1999) study, for instance, found that young Finnish learners generally spent only six percent (around 16 to 20 seconds) of total study time on the

pictures of authentic Biology textbook passages. Native German speakers in Schmidt-Weigand et al.'s (2010) study also demonstrated similar tendency when learning an animation accompanied with written text spent as they spent more time attending to text than on the dynamic visualisations across two experiments. Similar findings have been found in many other studies (Cromley et al., 2010; Ozcelik et al., 2010; Rayner et al., 2001; Schwonke et al., 2009).

Unfortunately, multimedia researchers argue that pictures should not be neglected as successful comprehension and learning of multi-representational texts necessitate the readers to integrate information from both representations. Text-picture integration is fundamental as it allows readers to construct a coherent mental representation of the learning content (Ainsworth & Lowe, 2012; Hegarty & Just, 1993; Mayer, 2005b; Renkl & Scheiter, 2015; Seufert, 2003). Much of the available literature suggests that the process of integrating text and picture is usually guided by the text. Readers often read or even reread the verbal information first before attending to the pictorial information and process the text extensively to form a global understanding of the material content (Hegarty & Just, 1993; Hochpöchler et al., 2013). Such approach is described in detail by Hegarty and Just (1993) who presented readers with a textual description of a pulley system along with its corresponding diagram. They found that readers processed the textual description in increments, by first integrating information at a text-based level, and then building a spatial mental model with the diagram as an aid. The process of switching between the textual description and the diagram was highly interleaved. The readers started with reading or rereading one or more information about the components before inspecting the corresponding part in the diagram. They then continued reading about the next components.

Although the task of integrating representations may seem uncomplicated, it is neither easy nor automatic and is arguably the most challenging task in learning with multi-representational texts (Ainsworth, 2014; Bartholomé & Bromme, 2009; Schnotz et al., 2017; Won, Yoon, & Treagust, 2014). Accordingly, the literature has highlighted that not all learners attempt to integrate representations when reading. In some studies, readers are grouped according to how much they integrated (or did not integrate) representations (Mason, Tornatora, et al., 2013) or how strongly they integrated the representations (Mason, Tornatora, et al., 2015). They can also be grouped according to their referencing strategies (Jian & Wu, 2015).

Even though the literature has generally concluded that the native readers generally display a text-dominant and text-guided approach in reading multi-representational texts in their native tongue, the extent to which they integrate text and pictures differs. When it comes to reading multi-representational texts in an L2, it could be possible that readers will also exhibit similar reading patterns. However, the reasons behind these patterns may be different. One may argue that the difficulty in comprehending certain L2 vocabulary leads the L2 readers to predominatly attend to the text, consequently increasing the time reading and re-reading the verbal information in an effort to establish understanding. Indeed, when compared with more skilled readers, those who have poorer reading skills have been shown to spend longer amount of time reading text upon first encounter with a passage in L1 (Mason, Tornatora, et al., 2013).

Alternatively, comprehension difficulties could lead L2 readers to switch their attention to the picture in an attempt to resolve their vocabulary problem. In other words, the L2 readers could use the picture as an information source to help them gain better understanding of the text (Ainsworth's constraining function, 2006). Lee and Mayer (2015) propose that pictorial information that is redundant yet representational in nature can be beneficial when learning takes place in one's L2, especially when learners struggle with word meaning. They believe that adding this redundant information in pictorial format can help learners determine word meaning quickly by building referential connections between the verbal and pictorial representations of the words in question. This will consequently help them with the comprehension of the overall lesson. In their study, adding a concurrent video to an audio lesson in Korean learners' L2 (i.e. English) led to better comprehension performance than merely learning from audio.

Several hypotheses can therefore be formulated about the L2 reading patterns, hence the exploratory nature of Study 1. Nevertheless, it is expected that reading of multi-representational texts will also be largely text-dominant like in L1. Fixation frequencies and duration will largely be spent on the text portions of the stimulus so that the L2 readers can successfully comprehend the overall content. In contrast, the picture areas are expected to be processed less extensively. Study 1 also predicts that L2 readers will not completely ignore the pictures and will integrate verbal and pictorial information. The extent to which they integrate these two information however will vary, consequently leading to distinct patterns of eye movements among them. L2 proficiency factors (i.e. reading comprehension and vocabulary knowledge in English)

and prior knowledge are expected to correlate with relevant eye tracking measures. The role of these three individual characteristics are discussed in a subsequent section.

4.1.2. Text-picture Integration and Learning

Study 1 also examines the relationship between reading behaviour and learning from multi-representational texts in L2. In terms of text-picture integration, the literature has recorded systematic correlations between reading behaviour and various outcomes of learning. In accordance with the view that integrating verbal and pictorial information is crucial for learning with multi-representational texts, research has found that greater integration of text and pictures is associated with better learning outcomes. In eye-tracking research, a commonly used measure that examines this integrative processing is the integrative transitions. This refers to the number of saccades or gaze shifts made between text and picture. Mason and colleagues in a series of studies with Italian school children for instance have consistently found positive correlations between this measure and learning outcomes. In general, they found that the more frequent transitions were made between text and pictures when reading illustrated texts in the children's native tongue, the better they performed in the learning outcome tasks (Mason, Pluchino, Tornatora, & Ariasi, 2013; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013). Similar findings were also found with regard to another learning environment, i.e. simulation that involves integration of various graphic representations (O'Keefe et al., 2014).

There are however studies that did not find such associations. In Scheiter and Eitel (2015) for instance, learners who performed better in a text-diagram integration task did not make more frequent transitions between text and diagrams. Arndt et al. (2015) also found similar results. It has been argued that making frequent transitions between different representations could also suggest confusion (Rau, Michaelis, & Fay, 2015) or difficulties in integrating them (Holsanova, Holmberg, & Holmqvist, 2009). Integrative transitions may therefore indicate only superficial processing and may not be fine-grained enough to guarantee successful text-picture integration (Rau et al., 2015). Finer-grained integrative measures that examine more strategic integrative processing such as integrative transitions between corresponding or semantically-related text and picture elements as well as look-from fixation times either from text to picture or from picture to text were then proposed.

In one of the studies that utilised these finer-grained measures, Italian children who read an illustrated text in their native language were found to display three distinct eye movement patterns and were labelled high, intermediate or low integrators (Mason, Tornatora, et al., 2013). Those who displayed the highest level of text-picture integration performed better than intermediate and/ or low integrators in several learning tasks. Unlike the other two types of readers, the high integrators' gaze behaviour was characterised by greater integrative measures, i.e. more frequent integrative transitions, longer look-from picture to text, text to picture as well as text to text fixation times. The low integrators' gaze behaviour on the other hand, was associated with the poorest performance. More importantly, Mason, Tornatora, et al. (2015) found that the degree of text-picture integration was predictive of learning outcomes. The more frequent and the longer readers integrated text and pictures during second-pass reading, the better they performed in learning tasks.

Alternatively, and especially for simple multi-representational materials, learners may construct a coherent mental model of text and pictures from memory so that they do not have to make frequent transitions between text and pictures (Scheiter & Eitel, 2015; Schüler, 2017). While fixating on a text section, they may retrieve pictorial information from memory without having to switch attention to the corresponding picture. Supporting this proposition, Schüler (2017) found longer reading times on text and more revisits to text and pictures in her study on the processing of inconsistent verbal and pictorial information. With materials that contain lengthier texts like the one used in this research, it seems more plausible to assume that such integration may take place when learners re-read and integrate the lengthy text segments. As evident in Mason, Tornatora, et al. (2013), look-from text to text fixation time along with text-picture integrative measures are correlated positively with learning outcome.

Using Mason, Tornatora, et al. (2013)'s study as a model, several eye tracking measures were used to examine the process of text-picture integration and to relate them to learning outcomes. These are the frequency of transitions and look-from text to picture, and picture to text fixation times. Since learners may also attempt integration during re-reading of texts and previous research has found its crucial role in learning outcome, look-from text to text fixation time is also obtained. Finally, participants' look-from picture to picture fixation time is analysed to measure the integration of pictorial information. This is part of an effort to examine the distribution of attention on authentic multi-representational materials that are commonly used in the classroom setting. It is predicted that a positive relationship will emerge between

gaze behaviour and learning of the multi-representational text in L2. L2 readers who integrated text and pictures as well as text sections to a greater extent will have better learning outcomes than those who integrated them to a lesser extent.

As pointed out by Hannus and Hyönä (1999), the materials used in majority of the relevant studies have used rather simple materials consisting of text portions accompanied by only one picture. Authentic textbook materials however are much more complex and it is common to come across more than one visual representation per science textbook page, for instance (Cromley, Bergey, et al., 2013). Visual representations that appear in current textbooks have also increased in variety (Slough et al., 2010). Learners therefore encounter complex cognitive tasks when presented with such materials (Ainsworth, 2006; Hannus & Hyönä, 1999). Having participants read and learn from simplified materials may not therefore reflect the complexity of the learning tasks and environment that students immediately face in the classroom. Furthermore, learning performance examined from interaction with these materials may also not be applicable to the real-life learning environment. Study 1 uses an authentic passage of a science textbook to better reflect the complexity of processing and learning from such texts.

4.1.3. Reading of Multi-representational Texts and Individual Characteristics

Chapter 2 identified three individual characteristics that would be further examined in this thesis. English reading comprehension ability and prior knowledge as seen in Section 2.3 are linked to the ways readers engage and learn with multi-representational materials. They have been linked to readers' comprehension of text, graphics and also to their gaze behaviour when reading and can work either independently or collaboratively. The role of English vocabulary knowledge is also examined. As shown in Section 2.4.3, this variable is a significant predictor of L2 reading and research has shown that L1 and L2 learners differ in terms of their vocabulary knowledge in the target language. L2 learners have smaller L2 vocabularies which are usually less extensive compared to L1 learners' (Droop & Verhoeven, 2003). Their vocabulary knowledge may also differ qualitatively as they have weaker semantic networks due to the lack of associative links between words. These limitations may therefore impede their L2 reading and influence their gaze behaviour when reading in L2. English vocabulary knowledge together with English reading comprehension ability will be used to measure the L2 readers' English proficiency.

Based on the literature, Study 1 predicts several relationships between the L2 readers' individual characteristics and eye tracking measures. Specifically, a positive relationship is expected between prior knowledge and integrative transitions. Conversely, it is hypothesised that both English reading comprehension ability and vocabulary knowledge would be negatively associated with first-pass reading on the text areas of the reading material.

4.2. HYPOTHESES

To summarise, Study 1 attempts to explore the reading and learning of an authentic multi-representational text in L2 through the examination of L2 readers' eye movement patterns. Several hypotheses are derived based on the literature:
H_1: Reading of multi-representational texts in L2 will be text-dominant: Fixation frequencies and duration will be largely spent on the text portions rather than the pictures.

H_2: There will be distinct patterns of eye movements when reading a multi-representational text in an L2, varying in terms of the extent to which text and pictures are integrated.

H_3: There will be a relationship between online processing of the multi-representational material in L2 and individual characteristics of interest: Both reading comprehension ability and vocabulary knowledge in English will correlate negatively with first-pass fixation time on text. Prior knowledge will correlate positively with text-picture integrative measures.

H_4: There will be a relationship between online processing of the multi-representational material in L2 and learning: L2 readers who integrate text and pictures to a greater extent will be those who learn better from the material.

4.3. METHOD

4.3.1. Participants

Fifty university students who reported English as their second language took part in this study and received an inconvenience allowance for their participation. Due to calibration errors and drifts in gaze data when reading a long authentic text for an average of 5.93 minutes, only the data from 32 participants could be analysed. Their age ranged from 20 to 40 (M = 22.7, SD = 4.34). 20 participants reported Chinese as their L1, 8 reported Malay, 2 reported Spanish and 1 each for French and Korean. In total, there were 5 males and 27 females. (For sampling criteria, see Section 3.4).

4.3.2. Materials

The materials used in this study were introduced and detailed in Sections 3.5.1, 3.6.2 and 3.6.3. For the learning outcome test, the researcher scored all the participants' scripts and another rater independently scored a subset of seven randomly selected scripts to eliminate a measurement error. The intra-class correlation coefficient (a mixed model ICC, absolute type) between the two raters was .93, suggesting a very high interrater reliability.

4.3.3. Eye Tracking Measures

A number of eye tracking measures were computed to analyse how the L2 readers read the authentic multi-representational text. To determine how the L2 readers distribute their visual attention across the text and picture areas, the number and duration of fixations on each AOI group (text AOI group and picture AOI group) were computed. Proportion values were used to correct for the differences in total surface area occupied by the AOI groups. This was necessary as the AOI group with larger surface area may be attended to more extensively. The total area occupied by the text AOIs was larger; 962,325-pixel square (46.4%) in comparison with the picture AOIs' 732,436 pixel-square (35.3%). Fixation frequencies on either AOI group were therefore expressed as proportions of the total fixation count on the overall stimulus, whereas fixation durations as proportions of the total fixation duration on the overall stimulus.

The remaining seven eye tracking measures were adopted from Mason, Tornatora, et al. (2013)'s study and were defined and explained in Section 3.6.1.3.

4.3.4. Apparatus

The apparatus used in this study was described in Section 3.6.1.4.

4.3.5. Procedure

All participants were tested individually. The study comprised three stages; pre-reading, reading and post-reading. All were carried out in a single session. After obtaining consent from each participant, the pre-reading stage commenced with the administration of the English reading comprehension test (GSRT) followed by the prior knowledge test. Since all participants were above 17 years old, they started the GSRT with Story 8 and proceeded to the next and more difficult stories until they attained a ceiling. Once the ceiling was attained, they were told to stop. If a basal had not been established during this process, the participants were asked to read and attempt Story 7 (and subsequent less difficult stories) until the basal was established.

In the reading stage, the participants sat in front of a monitor with the built-in eye-tracking system and were asked to read the science passage silently and carefully as they would have to answer questions afterwards. There was no time limit and all participants read at their own pace. Before the reading experiment commenced, the system was calibrated on a nine-point grid. Following successful calibration, the participants were shown five successive pages; a blank screen, an instruction page, a fixation cross, another blank screen and the science passage. Once they have finished reading, they clicked on their mouse to end the experiment. Despite not imposing a time limit, all participants finished reading the stimulus under 10 minutes.

The post-reading stage entailed the administration of the learning outcome test, the English vocabulary task (LexTALE) and the background questionnaire. All participants completed the three stages under 1 hour and 45 minutes.

4.4. DATA ANALYSIS

Since the stimulus page was presented after a fixation cross, the first fixation of each participant was excluded from analyses. Also excluded were fixations that were shorter than 100ms. The eye tracking measures were transformed using the square root transformation due to violations of assumptions.

4.5. RESULTS

4.5.1. Distribution of Attention across text and Picture Areas

To address H_1 that reading of multi-representational texts in L2 will be text-dominant, proportion values for the number and duration of fixations on each AOI type (text AOI, picture AOI and whitespace AOI) were computed. Table 5 reports the means and standard deviations of these proportional values. In general, the text areas received more and longer fixations than the picture areas. Paired sample t-tests conducted on the square root transformed values indicated that these differences were significant. For proportion of fixation frequencies the difference was .35, $t(31) = 17.02$, $p < .001$, $d = 5.53$. For proportion of fixation duration the difference was .34, $t(31) = 15.23$, $p < .001$, $d = 5.38$.

Table 5. Mean proportion of fixation measures by AOI type

AOI type	Proportion of fixation frequency		Proportion of fixation duration	
	M	*SD*	*M*	*SD*
Text	.73	.07	.72	.08
Picture	.25	.07	.27	.07
Whitespace	.02	.01	.02	.01

4.5.2. Patterns of Eye Movements

The next set of analyses addresses H_2 which predicts that distinct patterns of eye movements would be exhibited when reading in L2. A cluster analysis was first carried out to examine whether there are distinct groups of readers in the sample. This analysis looks for a natural structure among observations and group them on the basis of their similarity (Hair, Black, Babin, & Anderson, 2010). As a result, observations in the same cluster are more similar and homogeneous than they are to objects in different clusters. The resulting clusters therefore have high within-group homogeneity and high between-group heterogeneity.

Seven eye tracking measures were selected as the cluster variables, as indicated by a previous study (Mason, Tornatora, et al., 2013). These are first-pass fixation time on picture (FP-Pict), first-pass fixation time on text (FP-Text), integrative transitions (transitions), look-from picture to picture fixation time (LF-Pict Pict), look-from picture

to text fixation time (LF-Pict Text), look-from text to picture fixation time (LF-Text Pict) and look-from text to text fixation time (LF-Text Text). The cluster analysis would therefore group the participants by measuring their similarity across all seven measures. For Study 1, the interobject similarity across the variables is calculated using the squared or absolute Euclidean distance.

An agglomerative hierarchical procedure was used to cluster the L2 readers. In an agglomerative hierarchical clustering, each observation begins as its own cluster and is joined successively until all observations end up in one single cluster. In each stage, the two most similar clusters are joined one at a time. Hierarchical clustering procedures merge clusters based on a clustering algorithm that defines how similarity is measured between clusters of multiple members. The clustering algorithm selected is the Ward's method which measures similarity based on the minimum variance criterion.

To illustrate, in Ward's hierarchical clustering procedure (just like in other hierarchical procedures), each of the 32 L2 readers begins at the initial stage of the analysis as one cluster. In the second (and subsequent) stages, the method looks for a pair of clusters that produces the least increase in within-cluster variance so that the variability within the cluster would be minimal after merging. The aim of Ward's method is therefore to produce a cluster solution with minimal within-cluster variance.

The relevant eye tracking measures that have been transformed using the square root transformation were first standardised for the clustering procedure. To determine the final cluster solution, three aspects were looked at; the proportion of change in heterogeneity, visual inspection of the dendrogram and clarity of interpretation. Based on these aspects, a two-pattern cluster solution was chosen. Figure 10 displays the dendrogram whereas Table 6 reports means and standard errors of the eye tracking measures for both clusters. Standard errors are used as the sample size for each cluster differs. Figure 11 illustrates how the two clusters' means were noticeably different from each other. The means of each variable crosses the x-axis, producing a striking visual pattern. Note that values of eye tracking measures in tables presented henceforth in this chapter are transformed values.

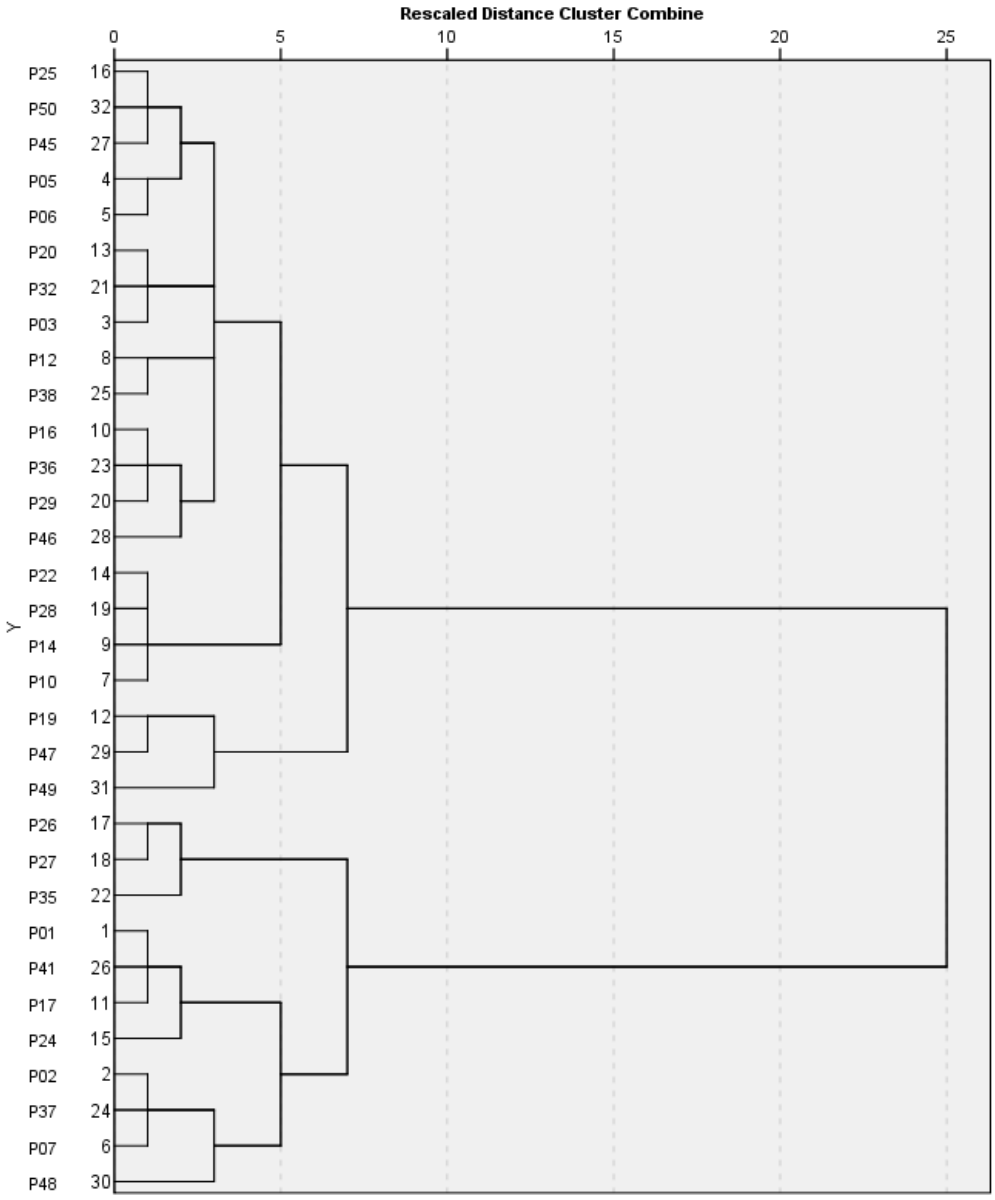


Figure 10. Dendrogram using Ward's linkage for Study 1

Table 6. Descriptive statistics of eye tracking measures by cluster

Measure	Cluster 1 (*n* = 21)		Cluster 2 (*n* = 11)	
	M	*SE*	*M*	*SE*
FP-Pict	-0.30	0.13	0.58	0.41
FP-Text	-0.23	0.21	0.43	0.30
Transitions	0.50	0.17	-0.95	0.18
LF-Pict Pict	0.11	0.20	-0.22	0.36
LF-Pict Text	0.59	0.13	-1.12	0.17
LF-Text Pict	0.54	0.12	-1.03	0.26
LF-Text Text	0.40	0.20	-0.76	0.19

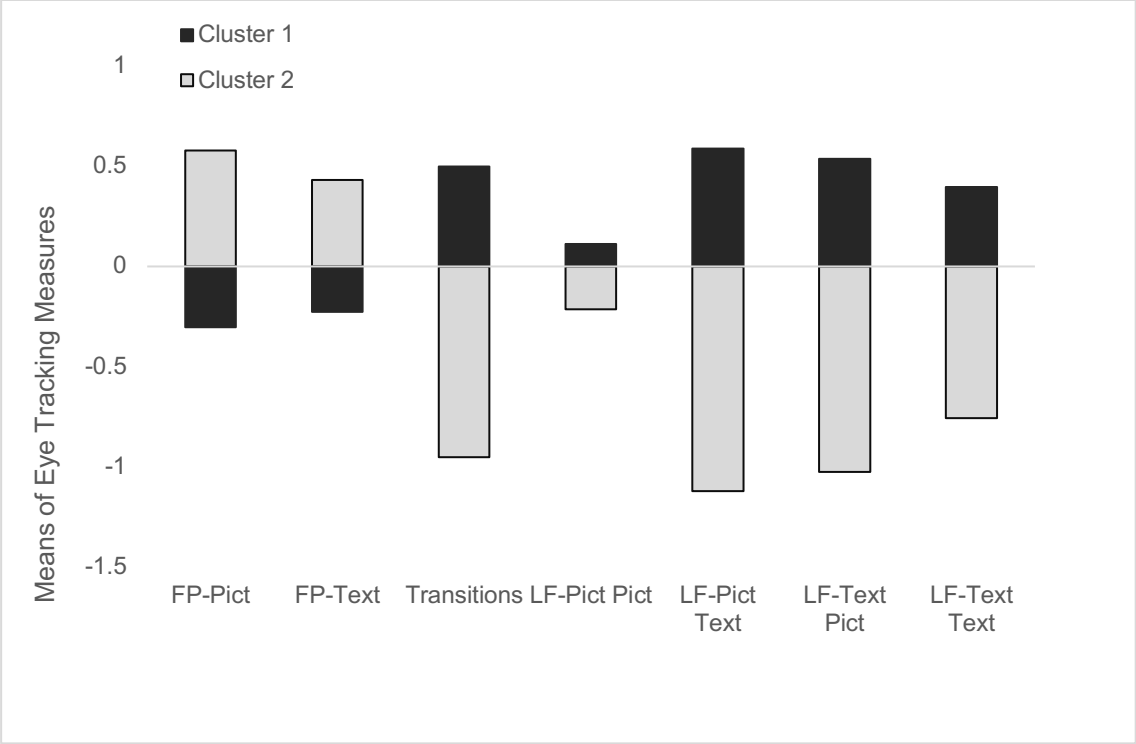


Figure 11. Means of eye tracking measures by cluster

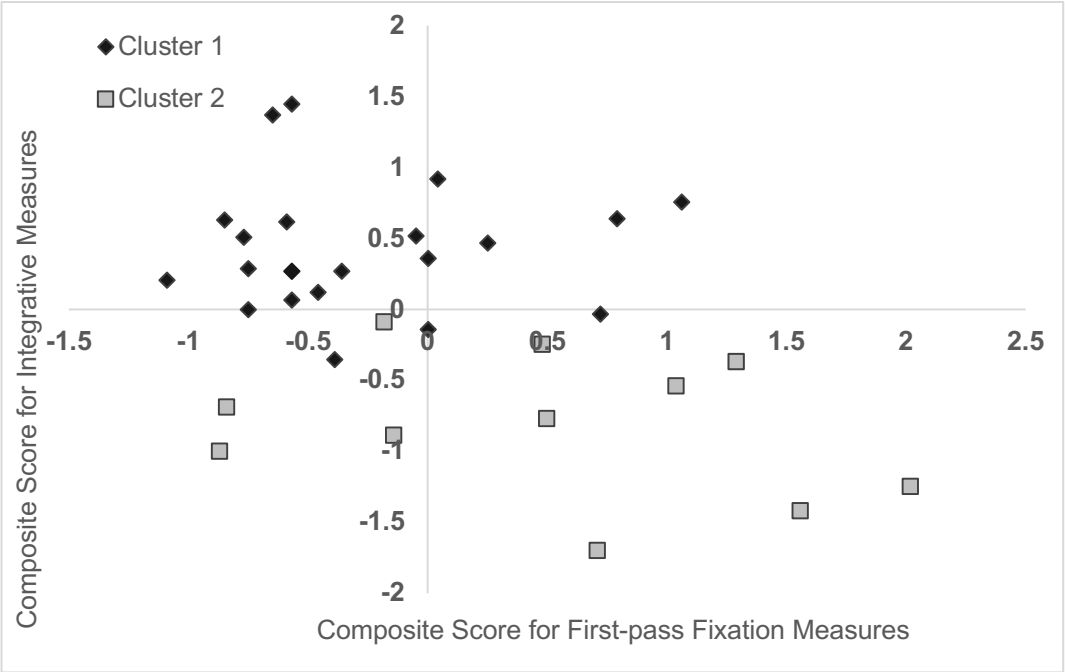


Figure 12. Scatterplot of composite scores for first-pass fixation measures and integrative measures

Figure 12 presents a scatterplot that visualises how dissimilar the participants in each cluster behaved. In this scatterplot, the composite score for first-pass fixation measures (i.e. means of FP-Pict and FP-Text) on the x-axis is plotted against the composite score for integrative measures (i.e. means of transitions, LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text) on the y-axis. The rationale is that a contrasting pattern appears to emerge between the two types of measures (refer to their means in Table 6). To illustrate, Cluster 1 readers spent less time on text and picture areas upon initial reading/ inspection but integrated across and within the representations more frequently and for a longer time. In contrast, Cluster 2 readers spent more time on text and picture areas upon initial reading/ inspection yet integrated across and within representations less frequently and for a shorter time. Due to this seemingly contrasting pattern, it was decided to plot them against one another on a single scatterplot so that the dissimilarity between the two clusters can be visualised more effectively.

A MANOVA test was performed to evaluate whether the two clusters differed for all seven eye movement measures. It revealed a large main effect of cluster, Pillai's Trace = .81, F (7, 24) = 14.19, p < .001; η_p^2 = .81. Univariate tests showed significant

differences only for five measures; FP-Pict, transitions, LF-Pict Text, LF-Text Pict, and LF-Text Text. These results are presented in Table 7.

Table 7. Analysis of variance for individual eye tracking measures

Measure	*df*	*F*	η_p^2	*p*
FP-Pict	30	6.69	.18	.02
FP-Text	30	3.14	.10	.08
Transitions	30	28.93	.49	.000
LF-Pict Pict	30	.77	.03	.39
LF-Pict Text	30	64.55	.68	.000
LF-Text Pict	30	39.84	.57	.000
LF-Text Text	30	13.51	.31	.001

Overall, L2 readers in Cluster 1 made more gaze shifts or transitions between text and picture areas compared to those in Cluster 2. Furthermore, these readers spent more time re-reading text areas while re-inspecting picture areas and vice versa. Another significant difference is that L2 readers in Cluster 1 also spent more time re-reading a text area while re-reading another. In contrast, Cluster 1 readers spent significantly less time than Cluster 2 readers inspecting the pictures upon first contact. They also spent less time reading the text areas upon first contact but this difference in first-pass fixation time on text was only marginally significant.

Based on the analyses, it could be concluded that there are two distinct groups of L2 readers in the sample that differ significantly in terms of integrative processing. This processing is indicated by the frequency of attempts at integrating text and pictures and the time spent integrating these representations during the second-pass processing. They also differ significantly in terms of the time spent integrating text areas during the second-pass processing. Another significant factor is the time spent inspecting picture areas during the first-pass processing. Since the L2 readers in Cluster 1 have higher mean scores with respect to integrative processing, they are thus labelled as ‘high’ integrators, whereas Cluster 2 readers are labelled ‘low’ integrators due to their lower mean scores on these measures. This difference is illustrated in Figure 13 whereby a typical example of scan path for each cluster is presented. In the top figure, a high integrator’s greater integrative processing can be seen, for instance, from the large number of transitions (as shown in by the saccades or lines) between the text and picture segments. In contrast, a low integrator’s low

integrative processing can be seen from the few transitions between the text and picture segments, as seen in the bottom figure.

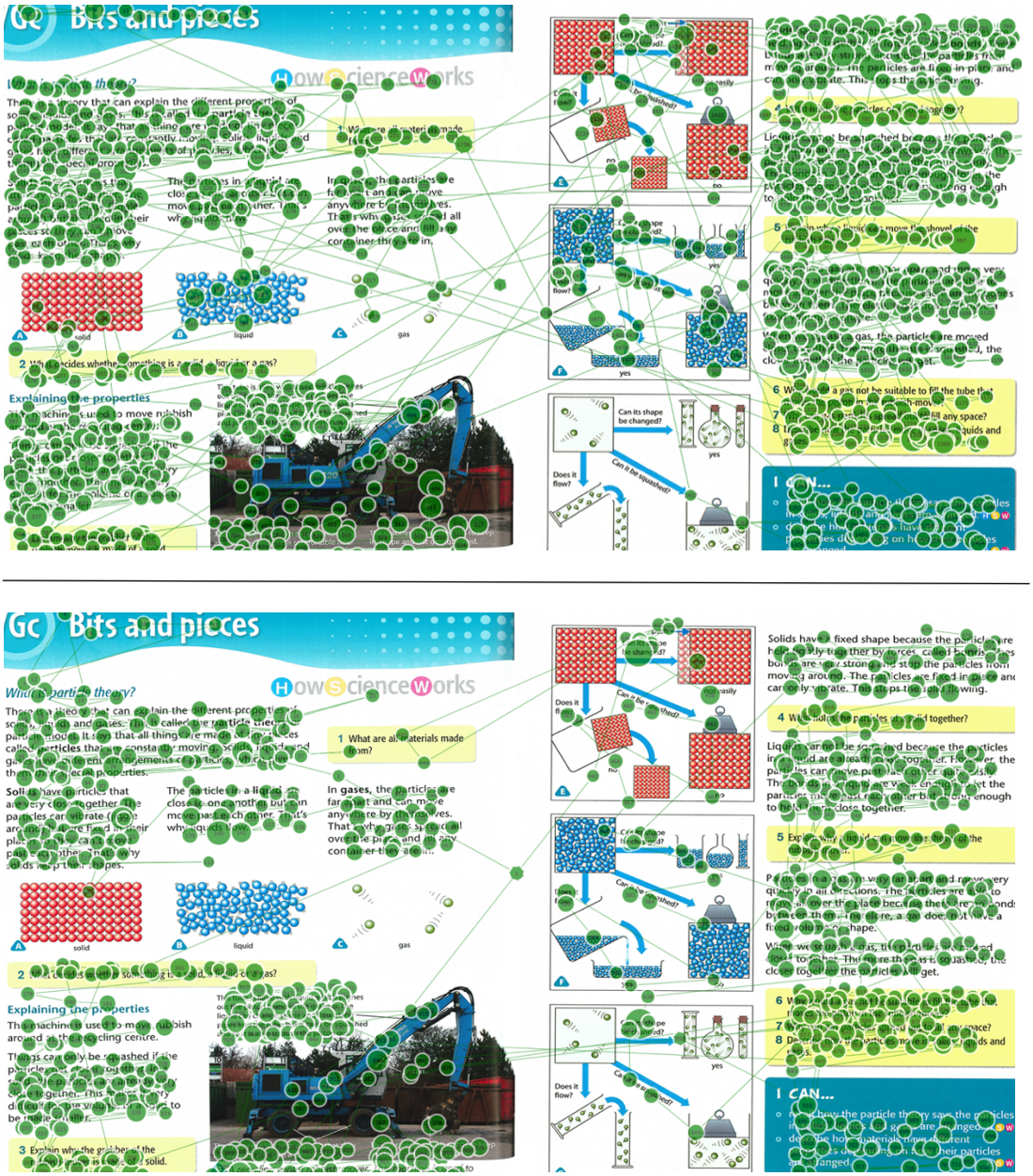


Figure 13. Typical examples of a high integrator's scan path (top) and a low integrator's scan path (bottom)

4.5.3. Eye Movements and Individual Characteristics

H_3 predicts that there would be a relationship between online processing of the multi-representational material in L2 and individual characteristics of interest. To address this, the high integrators' means for each individual characteristic test were compared with the lower integrators'. Table 8 displays the descriptive statistics of both individual characteristic tests for each cluster.

Table 8. Descriptive statistics of individual characteristic tests by cluster

Individual characteristic	Cluster 1 *n* = 21		Cluster 2 *n* = 11	
	M	*SE*	*M*	*SE*
Reading comprehension ability (scored out of 65)	50.86	1.09	49.18	1.66
Vocabulary knowledge (scored out of 100)	74.82	3.48	70.57	4.17
Prior knowledge (scored out of 10)	7.67	0.52	7.09	0.60

A MANOVA was carried out to evaluate whether or not the means of the two clusters significantly differed. No statistical difference emerged between the two clusters and the combined dependent variables, Pillai's Trace = .04, F (3, 28) = .43, p = .73; η_p^2 = .04. Table 9 reports the univariate tests.

Table 9. Analysis of variance for individual characteristic test

Individual characteristic	*df*	*F*	η_p^2	*p*
Reading comprehension ability in English	30	.76	.03	.39
Vocabulary knowledge in English	30	.56	.02	.46
Prior knowledge	30	.47	.02	.50

Given these results, a finer-grained analysis was performed to examine the relationship between each of the seven eye tracking measures and each individual characteristic. Again, no significant associations emerged, although the negative relationship between vocabulary knowledge in English and FP-Text was marginally significant (r = -.32, p = .08), see Table 10.

Table 10. Correlations between eye tracking and individual characteristic measures

Measure	Reading comprehension ability		Vocabulary knowledge		Prior knowledge	
	r	*p*	*r*	*p*	*r*	*p*
FP-Pict	.13	.50	.16	.38	- .04	.84
FP-Text	- .19	.29	- .32	.08	- .30	.10
Transitions	.18	.34	.22	.22	.26	.15
LF-Pict Pict	.15	.41	.06	.75	- .02	.91
LF-Pict Text	.22	.24	- .10	.60	.21	.26
LF-Text Pict	- .02	.93	- .01	.95	.22	.22
LF-Text Text	- .00	.99	.13	.49	.12	.53

Although the language group construct is not specifically examined in this study as sampling was by opportunity and not by specific L1, it is nevertheless useful to examine whether any language group is particularly dominant in either of the two clusters. Table 11 displays the breakdown of Clusters 1 and 2 participants' first languages. The high integrators in Cluster 1 were mostly Chinese readers (67%), followed by the Malay readers (24%). The low integrators were, again, mostly Chinese readers (55%), followed yet again by the Malay readers (27%). However, this pattern of distribution of different L1 speakers reflects the proportions of each L1 in the sample with the Chinese readers making up 63% of the sample, followed by the Malay readers (25%).

Table 11. Breakdown of cluster by language group

L1	Cluster 1 *n* = 21	Cluster 2 *n* = 11	Total
Spanish	1	1	2
Chinese	14	6	20
Malay	5	3	8
French	1	0	1
Korean	0	1	1

4.5.4. Eye Movements and Learning Outcome

H_4 predicts a relationship between online processing of the multi-representational material in L2 and learning of the material. Specifically, L2 readers who integrated

text and pictures to a greater extent were predicted to be those who learned better from the material.

Before comparing the performance of each cluster in the learning outcome test, the relationship between each individual characteristic and learning outcome was first examined. Table 12 details the correlations. No significant relationship was found between L2 readers' individual characteristics and learning outcome, although its relationship with prior knowledge of the properties of materials was marginally significant (r = .32, p = .07). Since none of the individual characteristics correlated significantly with learning outcome, they were not considered further in the study.

Table 12. Correlations between individual characteristic and learning outcome test

Reading comprehension ability		Vocabulary knowledge		Prior knowledge	
r	*p*	*r*	*p*	*r*	*p*
- .15	.42	.12	.50	.32	.07

An independent *t*-test was performed to determine whether the high and low integrators differ significantly with regard to the learning outcome. On average, the high integrators performed better (M = 18.95, SE = 0.58) than the low integrators (M = 16.05, SE = 1.26). This difference was significant $t(30) = 2.40$, p = .02, d = .83.

4.6. DISCUSSION

Study 1 aims to explore the reading and learning of an authentic multi-representational text in an L2. To accomplish this both online (eye tracking) and offline (individual characteristics and learning outcome) data were collected and analysed.

H_1: Reading of multi-representational texts in L2 will be text-dominant: Fixation frequencies and duration will be largely spent on the text portions rather than the pictures.

The first objective sought to determine whether similar to reading and processing of multi-representational texts in L1, reading and processing in L2 would also be heavily text-based. After controlling for differences in total surface area occupied by the text chunks and pictures, fixation data indicate that just like what has been observed in

the L1 literature (Hannus & Hyönä, 1999; Rayner et al., 2001; Schmidt-Weigand et al., 2010), L2 reading of the multi-representational text also proceeded in a text-dominant fashion. In general, the text areas were attended to almost three times more than the picture areas. Specifically, the L2 readers spent an average of 72% of their learning time on the text areas and only 27% on the pictures. Similarly, 73% of their total fixations were on the text areas, whereas only 25% were on the pictures. Effect sizes were large, d = 5.99 for proportion of fixation duration and 6.85 for proportion of fixation frequency. These findings suggest that irrespective of whether the material is in one's L1 or L2, multi-representational reading in general, is processed in a text-dominant fashion. They also extend the text-dominant tendency observed among language learners presented with multi-representational materials (Bisson et al., 2014b) and among those who are learning with illustrated narratives in their L2 (Pellicer-Sanchez et al., 2018; Yusof et al., 2019) to learning with more expository materials.

H_2: There will be distinct patterns of eye movements when reading a multi-representational text in an L2, varying in terms of the extent to which text and pictures are integrated.

The next objective was to investigate whether the L2 readers displayed distinct patterns of eye movements. Despite predicting that reading of the material would be text-dominant, the L2 readers were expected to pay (some amount) of attention to the pictures, perhaps as an information source to help them resolve comprehension issues when reading in an L2 (Lee & Mayer, 2015). The extent to which such attempts at text-picture integration occurred, however were expected to differ, similar to what Mason, Tornatora, et al. (2013; 2015) observed. The cluster analysis modelled from Mason, Tornatora, et al.'s (2013) study, confirmed this prediction. Two groups of readers emerged and were labelled 'high' or 'low' integrators according to the extent they integrated verbal and pictorial information. Both groups also differed in terms of the time they spent integrating the verbal information during re-reading (LF-Text Text). Another significant difference was in the time they inspected the pictures during initial reading (FP-Pict). To be more specific, the high integrators' pattern of eye movements was characterised by more attempts to integrate text and picture areas (transitions) and longer time integrating text and picture areas during re-reading of text (LF-Text Pict) or re-inspection of pictures (LF-Pict Text). In contrast, the low integrators' eye movement patterns were characterised by fewer attempts to integrate text and picture

areas and shorter time integrating text and picture areas during re-reading of text or re-inspection of pictures.

Mason, Tornatora, et al. (2013) also found that their three groups of Italian readers significantly differed in terms of these five eye tracking measures. However, one difference is with regard to FP-Pict. In their study, the high integrators spent significantly more time inspecting the one picture in the illustrated text during initial reading. Study 1 found the opposite as the high integrators spent less time on the pictures. This could be due to the inclusion of text in picture AOIs in the present study. The two clusters might have actually differed in terms of the time they spent reading the text in the picture AOIs with the low integrators displaying a more laborious processing than the high integrators. Alternatively, the different pattern could be due to the different ways in which second-pass measures were computed in the two studies computed their second-pass measures. The present study defines a second-pass fixation as any refixation on an AOI after it has been exited once, and re-entered. In Mason, Tornatora, et al.'s (2013) study, a second-pass fixation on an AOI would only be counted as such if all other AOIs have been attended to and exited. This would appear appropriate with such simpler-designed material, however, given the complexity of the material in this research and its number of AOIs, the former computation was deemed more appropriate in reflecting the participants' reading behaviour. Having said that, the remaining measures, all of which are integrative, show identical patterns. Hence, despite the differential patterns in FP-Pict, it would be important to note that Study 1 essentially replicated what has been found by Mason, Tornatora, et al. (2013).

H_3: There will be a relationship between online processing of the multi-representational material in L2 and individual characteristics of interest: Both reading comprehension ability and vocabulary knowledge in English will correlate negatively with first-pass fixation time on text. Prior knowledge will correlate positively with text-picture integrative measures.

The relationship between these distinct eye movement patterns and the three individual characteristics were scrutinised. One of the predictions was that the L2 readers' reading comprehension ability and vocabulary knowledge in English would each correlate negatively with the time they spent reading the text areas during the first-pass. One of the reasons behind this prediction is that the participants were reading in a language that is less familiar to them. As this was the case, their first-

pass fixation time on text areas were expected to be longer, which normally reflects difficulties in understanding the verbal information (Mason, Tornatora, et al., 2013; Rayner et al., 2006). Study 1 also predicted that greater text and picture integration would be associated with higher prior knowledge since it equips readers with a schema which could be used to integrate representations (Scheiter et al., 2014). However, no significant differences emerged for any of the participants' pre-existing individual characteristics. A finer-grained analysis that examined the relationships between the individual characteristics and each eye tracking measure also revealed no significant results. The three individual characteristics examined in Study 1 therefore cannot account for how an authentic multi-representational text was read and processed in L2.

Due to the complexity of the reading material it was expected that L2 proficiency and pre-existing knowledge would have some impact on eye movements. The lack of significant results is therefore rather surprising. In terms of reading comprehension ability and vocabulary knowledge in English, it could be that the passage was not demanding enough in terms of readability for the readers. Hence even though the L2 readers varied considerably on both variables, they could still comprehend the text segments and form a coherent understanding of the passage successfully regardless of their level of proficiency. As understanding of verbal information can guide readers' attention towards relevant pictures, readers can then attempt to comprehend the pictures and integrate verbal and pictorial information to form a mental model of the multi-representational text. Note that there is a significant correlation between look-from fixation time on text and second-pass fixation time on picture ($r = .379$, $p = .033$), suggesting those who spent more time re-reading and integrating text segments during re-reading also spent more time inspecting the pictures during re-inspection. Note, however, that there was a trend towards a significant negative relationship between first-pass fixation on text and vocabulary knowledge in English ($p = .077$). This suggests that there is a tendency for readers with less English vocabulary knowledge to read the text segments for a longer time during the initial reading compared to those with more vocabulary.

With regard to prior knowledge, Schnotz et al. (2017) argue that text-picture integration competence is a domain-general construct and the mapping procedures between verbal and pictorial information do not require domain-specific prior knowledge. Similar to Schnotz et al.'s (2017) study, it was not necessary for the

participants in the present study to have prior knowledge about the properties of solids, liquids and gases to integrate representations in the reading material.

Alternatively, lack of variation in the scores could have caused this lack of significant relationships. For prior knowledge, the sample size was rather small to attain enough variation in the test scores. Similarly, with regard to English reading comprehension ability and vocabulary knowledge, given that the L2 readers were UK university students with a good overall command of English, there may not have been enough variation in the test scores to arrive at significant relationships.

The issue of the L2 readers' first languages is also worth considering. It is plausible that a reader's L2 reading behaviour is related to their L1 and perhaps existing cultural practices. The sample consisted of students who spoke a variety of L1s. This was not the basis for sampling and so consequently the number of speakers of any particular L1 varies considerably. Hence, it is not possible to analyse these data inferentially. However, a descriptive analysis was conducted to explore how each L1 was distributed in the two clusters. As seen in the results section, although the Chinese readers were the most frequent L1 in the high integrator cluster, they were also the most frequent L1 in the sample as a whole. In fact, the breakdown of each of the L1s into clusters was remarkably similar to their proportions in the overall sample as a whole. This suggests that no one language group was particularly dominant in any cluster. Differences in the high and low integrators' reading behaviour therefore were unlikely to be solely due to existing language or cultural practices relevant to the L2 readers' first language.

H_4: There will be a relationship between online processing of the multi-representational material in L2 and learning: L2 readers who integrate text and pictures to a greater extent will be those who learn better from the material.

Finally, the study predicted that there would be a significant association between eye movement patterns and learning from the multi-representational text in L2 just like what has been observed for multi-representational learning in L1. In particular, the pattern of eye movements comprising more frequent and longer integrative processing would be associated with better learning performance, whilst the opposite pattern would be associated with poorer learning performance. As expected, the two groups of readers performed significantly different in the learning outcome test and the high integrator group (M = 18.95 or 63.17%) who integrated text and picture areas

to a greater extent performed better compared to the low integrator group (*M* = 16.05 or 53.5%). These findings suggest that the eye movement pattern of the high integrators characterised by more frequent and longer attempts to integrate text and picture areas might be more useful and strategic to learn from multiple representations. This is in line with prior L1 studies that found similar relationships (Hannus & Hyönä, 1999; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013; O'Keefe et al., 2014). Additionally, since the material also contained substantial amount of text, it might also be crucial to re-read and integrate the text segments to understand the content successfully. From a more technical perspective, these findings demonstrate that less and more fine-grained measures of integration (i.e. transitions and look-from fixation times) are both useful to measure integrative processing. Overall, the findings are in line with the view that integrating text and pictures is fundamental to learn successfully from multi-representational texts. Integration of external representations therefore might also be useful for learning in the L2 as it is for learning in the L1.

4.7. CONCLUSION

Study 1 aims at extending current research on multi-representational learning by exploring readers' engagement with an authentic multi-representational text in their L2. An examination of the L2 readers' eye movements revealed that they displayed similar patterns of processing as those reading in their L1. Firstly, the L2 readers' processing of the text was heavily text-dominant and they attended the text portions almost three times longer than they attended the pictures. More importantly, similar to L1 readers, the L2 readers also demonstrated distinct patterns of eye movements that differed, essentially, in terms of the extent verbal and pictorial information were integrated. Due to the complexity of the reading material it was expected that L2 proficiency and pre-existing knowledge would have some impact on eye movements. However, none of these participant-specific factors correlated significantly with the eye movements. The two groups of L2 readers also performed differently in the learning outcome test with the high integrators performing better, complementing the view that integrating text and pictures is essential to learning from multiple representations in authentic science textbooks. It thus appears that the integration process when learning in one's L2 might be as equally important as it is in one's L1. In conclusion, Study 1 revealed striking similarities in the way that L1 and L2 readers read, process and learn from multi-representational texts. However further evidence

is needed to corroborate this observation. The next chapter presents a subsequent study that examines the processing and learning of the same multi-representational text in the L1.

5. STUDY 2: THE L1 STUDY

5.1. INTRODUCTION

Study 1 found that when L2 readers read a multi-representational text, they appear to display similar visual behaviour, characteristic of their L1 counterparts. To summarise, Study 1 found that L2 readers were more text rather than picture-dominant, in that they fixated on the text portions of the passage more and longer than they did on the pictures. The findings also support the importance of integrating verbal and visual information to learn successfully from multi-representational texts. In particular, the L2 readers who benefitted more from the passage integrated text and pictures to a greater extent as shown by the high integrator cluster. They also used text to guide their text-picture integration. Moreover, those who learned better spent more time integrating different text portions. These findings are consistent with what has been found with regard to reading multi-representational texts in one's L1. However, it would be misleading to claim that the L1 and the L2 readers display similar visual behaviour when reading multi-representational texts unless a direct contrast is undertaken. Study 2 is an initial step towards this comparative study, examining the ways in which the L1 English readers read and learn the same multi-representational material.

5.2. HYPOTHESES

Study 2 explores the reading and learning of an authentic multi-representational text in L1 through the examination of L1 readers' eye movement patterns. Several hypotheses are derived based on the literature:

H_1: Reading of multi-representational texts in L1 will be text-dominant.
Fixation frequencies and duration will be largely spent on the text portions rather than the pictures.

H_2: There will be distinct patterns of eye movements when reading a multi-representational text in an L1, varying in terms of the extent to which text and pictures are integrated.

H_3: There will be a relationship between online processing of the multi-representational material in L1 and individual characteristics of interest: Reading comprehension ability in English will correlate negatively with first-pass fixation time on text. Prior knowledge will correlate positively with text-picture integrative measures.

H_4: There will be a relationship between online processing of the multi-representational material in L1 and learning: L1 readers who integrate text and pictures to a greater extent will be those who learn better from the material.

5.3. METHOD

5.3.1. Participants

42 native English university students took part in this study and received an inconvenience allowance for their participation. Due to calibration errors and drifts in gaze data when reading the long authentic text for an average of 3.61 minutes, only the data from 33 participants could be analysed. Their age ranged from 18 to 41 (*M* = 22.12, *SD* = 4.99). Some of the criteria that were set for recruiting L2 readers also applied when recruiting the L1 readers to ensure that the two samples (i.e. L1 and L2 readers) were as homogeneous as possible. Specifically, the L1 readers were also those whose science education is below the standards of A-level or its equivalence and they were neither visually impaired nor had any reading difficulties.

5.3.2. Materials

Study 2 used the same materials as Study 1 with two exceptions. The LexTALE was excluded from the study, as it is meant for English as a second language speakers. Additionally, the background questionnaire administered was slightly different as the language profile was removed. For language profiling, the L1 readers only needed to state what their additional language(s) were, if any. Sections on their clarity of vision, reading disabilities and science education profile were retained.

5.3.3. Eye Tracking Measures and Apparatus

Study 2 used the same eye tracking measures and apparatus as Study 1.

5.3.4. Procedure

The procedure was the same as in Study 1, except that the participants did not have to undertake the LexTALE task. Hence, for the post-reading stage, only the learning outcome test and the background questionnaire were administered. All participants completed the three stages of pre-reading, reading and post-reading under 1 hour and 15 minutes.

5.4. RESULTS

5.4.1. Distribution of Attention across Text and Picture Areas

To address H_1 that reading of multi-representational texts in L1 will be text-dominant, proportion values for the number and duration of fixations on each AOI group (text AOI, picture AOI and whitespace AOI) were computed. Table 13 reports the means and standard deviations of these proportion values. In general, the text areas received more and longer fixations than the picture areas. Paired sample t-tests conducted on the square root transformed values indicated that these differences were significant. For proportion of fixation frequency, the difference was 0.34, $t(32) = 17.28$, $p < .001$, $d = 5.96$. For proportion of fixation duration, the difference was 0.32, $t(32) = 15.08$, $p < .001$, $d = 5.22$.

Table 13. Mean proportion of fixation measures by AOI type

AOI type	Proportion of fixation frequency		Proportion of fixation duration	
	M	*SD*	*M*	*SD*
Text	.72	.07	.71	.08
Picture	.26	.07	.27	.07
Whitespace	.02	.01	.02	.01

5.4.2. Patterns of Eye Movements

The next set of analyses addresses H_2 that predicts that distinct patterns of eye movements would be exhibited when reading in L1. A cluster analysis was first carried out to examine whether there are distinct groups of readers in the sample. The eye

tracking measures used in Study 1 as the cluster variables were retained. They were first transformed using the square root transformation, then standardised for the clustering procedure. To determine the final cluster solution, three aspects were looked at; the proportion of change in heterogeneity, visual inspection of the dendrogram (Figure 14) and clarity of interpretation. Based on these aspects, a three-pattern cluster solution was chosen.

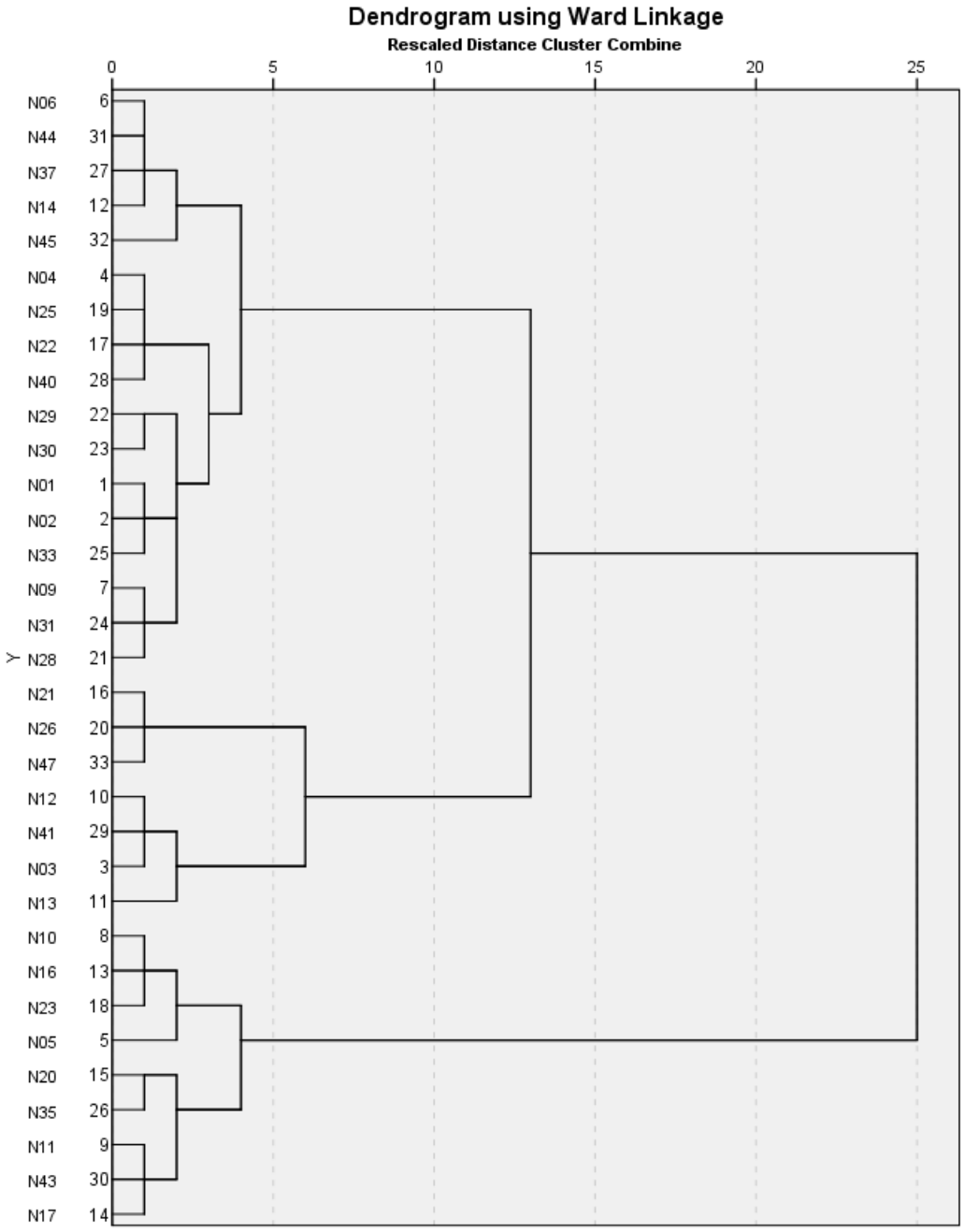


Figure 14. Dendrogram using Ward's linkage for Study 2

Table 14 reports means and standard errors of the eye tracking measures for each cluster. Note that values of eye tracking measures in tables presented henceforth in this chapter are transformed values.

Cluster 1 groups readers who displayed the most text-picture integrative behaviour. Specifically, they shifted between text and picture areas most frequently and spent the most time re-reading text while re-inspecting pictures and vice versa. They also spent the most time integrating within the same type of representations. In contrast, they spent less time on text and picture areas upon initial reading/ inspection. They are thus the 'high' integrators. Cluster 2, on the other hand, groups the 'intermediate' integrators whose integration of text and pictures is at the intermediate level, between Clusters 1 and 3. Compared to the 'high' integrators, the intermediate integrators shifted between text and picture areas less frequently and spent less time re-reading text while re-inspecting pictures and vice versa. They also spent less time integrating within the same type of representations (i.e. from picture to picture or from text to text). With regard to first-pass fixation measures, these readers spent less time on text and picture areas upon initial reading/ inspection. Those in Cluster 3 are the 'low' integrators. They integrated across and within the representations least frequently and spent the least time doing so. They however spent the most time on text and picture areas upon initial reading/ inspection.

Table 14. Descriptive statistics of eye tracking measures by cluster

Measure	High integrators (*n* = 7)		Intermediate integrators (*n* = 17)		Low integrators (*n* = 9)	
	M	*SE*	*M*	*SE*	*M*	*SE*
FP-Pict	-0.56	0.22	-0.31	0.15	1.03	0.38
FP-Text	0.13	0.53	-0.28	0.22	0.42	0.25
Transitions	1.19	0.43	-0.01	0.14	-0.92	0.15
LF-Pict Pict	0.95	0.26	0.13	0.18	-0.98	0.27
LF-Pict Text	1.41	0.21	-0.11	0.15	-0.89	0.22
LF-Text Pict	1.16	0.19	0.20	0.13	-1.28	0.13
LF-Text Text	1.32	0.18	-0.11	0.18	-0.81	0.22

Figure 15 illustrates how the three clusters' means differ from each other.

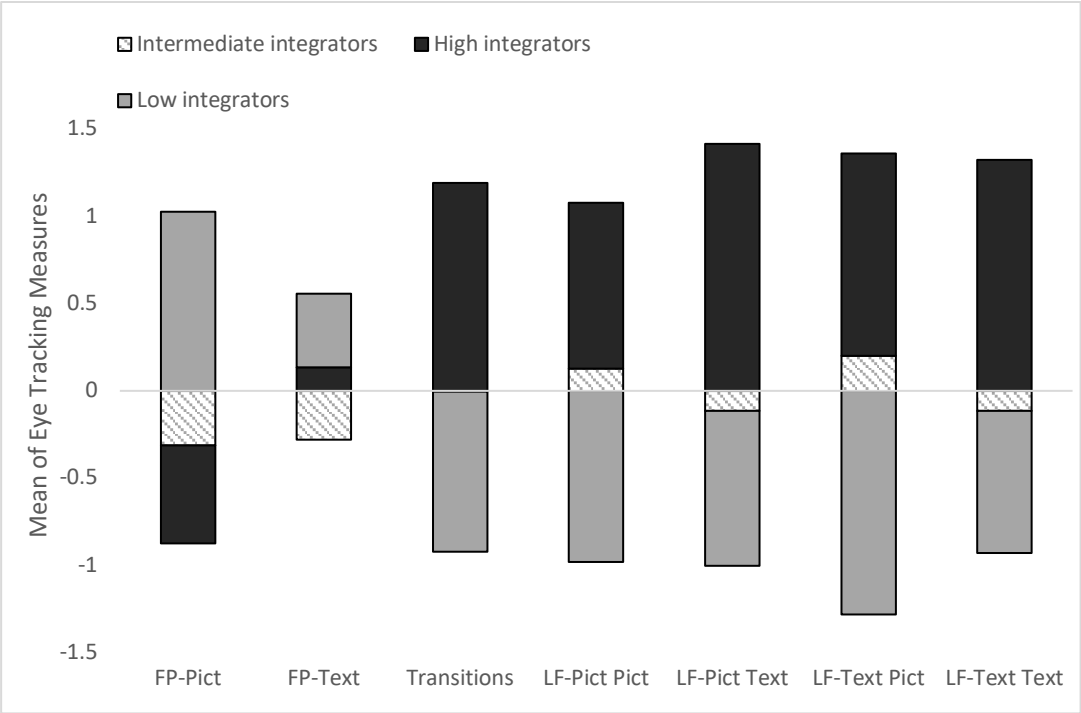


Figure 15. Means of eye tracking measures by cluster

A scatterplot (Figure 16) is also provided to visualise how dissimilar the participants in each cluster behaved. In this scatterplot, the composite score for first-pass fixation measures (i.e. means of FP-Pict and FP-Text) on the x-axis is plotted against the composite score for integrative measures (i.e. means of transitions, LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text) on the y-axis. The rationale is that a contrasting pattern appears to emerge between both types of measures. This is especially true for the high and low integrators (refer to their means in Table 14). To illustrate, the high integrators integrated across and within the representations most frequently and for the longest time but most of them spent less time on text and picture areas upon initial reading/ inspection. In contrast, the low integrators integrated across and within representations least frequently and for the shortest time, yet many spent more time on text and picture areas upon initial reading/ inspection. Due to this contrasting pattern, both types of measures were plotted against one another on a single scatterplot so that the dissimilarity between the clusters can be visualised more effectively.

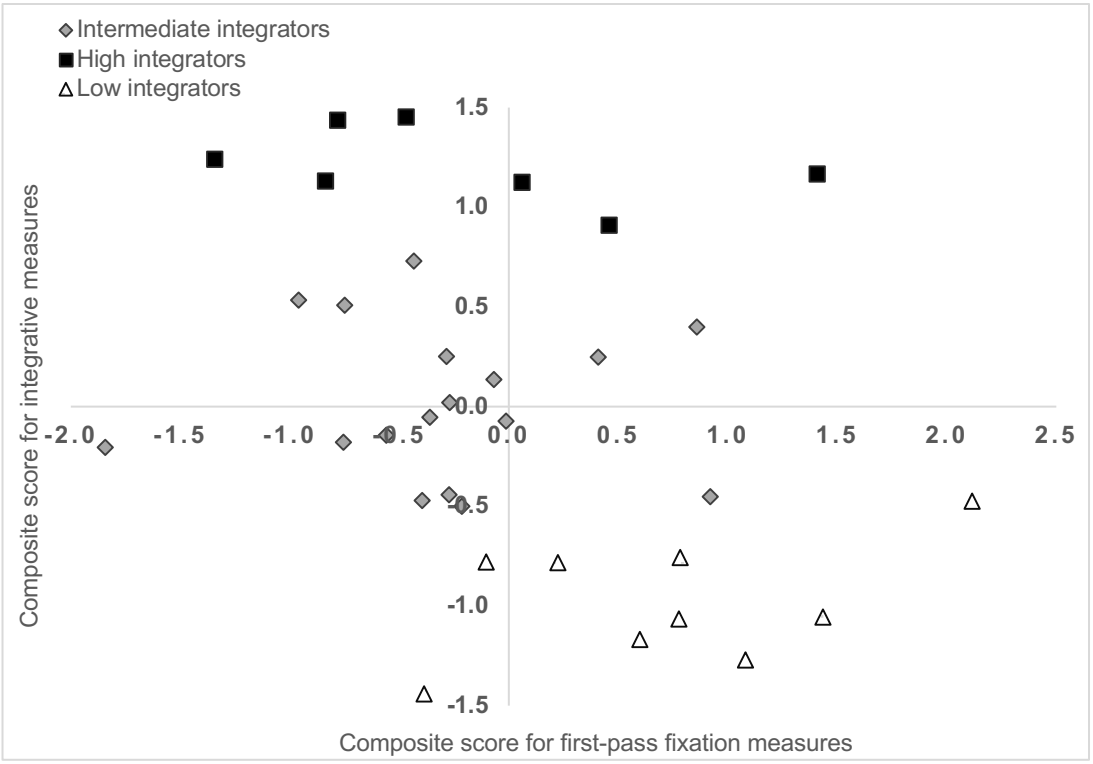


Figure 16. Scatterplot of composite scores for first-pass fixation measures and integrative measures

A MANOVA test was performed to evaluate whether the differences in eye tracking measures across the three clusters were significant. It revealed a large main effect of cluster, Pillai's Trace = 1.25, F (14, 50) = 5.92, p < .0001; η_p^2 = .62. Separate univariate ANOVAs showed significant differences for six out of seven measures. The clusters did not differ significantly only for the FP-Text. The results are presented in Table 15.

Table 15. Analysis of variance for eye tracking measures

Measure	*df*	*F*	η_p^2	*p*
FP-Pict	30	10.69	.42	.000
FP-Text	30	1.58	.10	.224
Transitions	30	18.06	.55	.000
LF-Pict Pict	30	13.63	.48	.000
LF-Pict Text	30	29.90	.67	.000
LF-Text Pict	30	52.15	.78	.000
LF-Text Text	30	20.33	.58	.000

Hochberg post-hoc test was reported due to the very different sample size across the clusters (Field, 2013). When the assumption of equality of variances was violated, Games-Howell post-hoc test was reported instead.

1. FP-Pict

Games-Howell post-hoc tests showed that the low integrators made significantly longer FP-Pict than both the high and intermediate integrators (*ps* = .009 and .020, respectively). The difference between the intermediate and the high integrators was not significant (*p* = .632).

2. Transitions

Games-Howell post-hoc tests showed that the low integrators made significantly less transitions than both the high and intermediate integrators (*ps* = .005 and .001, respectively). Again, there was no significant difference between the intermediate and the high integrators (*p* = .071).

3. LF-Pict Pict

Hochberg post-hoc tests showed that again the low integrators spent significantly shorter LF-Pict Pict than both the high and the intermediate integrators (*ps* < .001 and .003, respectively). The difference between the intermediate and the high integrators was marginal (*p* = .060) with the intermediate integrators spending shorter LF-Pict Pict.

4. LF-Pict Text

Hochberg post-hoc tests showed that the low integrators spent significantly shorter LF-Pict Text than both the high and the intermediate integrators (*ps* < .001 and .011). The intermediate integrators also spent significantly shorter LF-Pict Text than the high integrators (*p* <.001).

5. LF-Text Pict

Hochberg post-hoc tests revealed the same trend; the low integrators spent significantly shorter LF-Text Pict than the high and the intermediate integrators (both *ps* < .001). The intermediate integrators also spent significantly shorter LF-Text Pict than the high integrators (*p* <.001).

6. LF-Text Text

Hochberg post-hoc tests revealed that the low integrators spent significantly shorter LF-Text Text than the high and the intermediate integrators (p < .001 and = .05, respectively). The intermediate integrators also spent significantly shorter LF-Text Pict than the high integrators (p <.001).

These findings suggest that the three clusters of L1 readers mainly differ in terms of integrative processing. It appears that the high integrators displayed the most strategic reading behaviour associated with learning from multi-representational texts compared to the other two clusters. More specifically, they made the most transitions across representations and spent the longest time integrating across and within representations. On the other hand, the reading behaviour displayed by the low integrators suggests that it is the least strategic for learning as they integrated across and within representations the least and for the shortest amount of time. The intermediate integrators' reading behaviour appeared to be somewhere in the middle between the other two clusters. These differences can be seen in Figure 17, whereby a typical example of scan path for each of the clusters is presented. In the top figure, a high integrator's extensive integrative processing can be seen, for instance, from the large number of transitions (as shown by the saccades or lines) between text and picture segments. A low integrator whose integrative processing was the least extensive can be seen from the very few transitions made between text and picture segments, as seen in the bottom figure. As stated above, there seems to be a contrast between integrative processing and first-pass processing measures, which is most evident in the low and high integrators. The low integrators made the least integrative processing but paid most attention to text and pictures during the first-pass. Contrastively, the high integrators made the most integrative processing but paid less attention on the representations during the first-pass. How this seemingly contradictory behaviour was related to learning is examined in H_4.

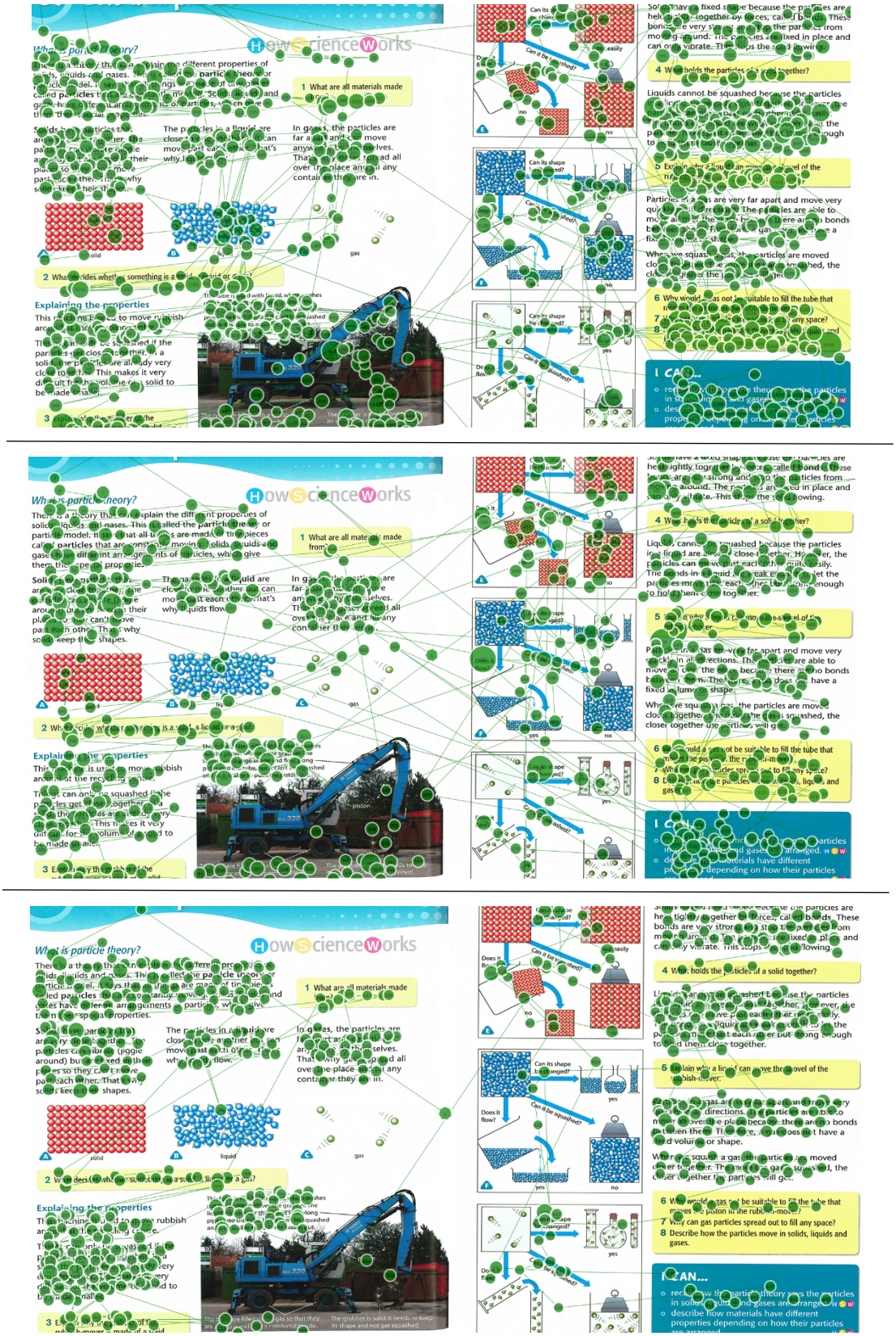


Figure 17. Typical examples of a high integrator's scan path (top), an intermediate integrator's scan path (middle) and a low integrator's scan path (bottom)

5.4.3. Eye Movements and Individual Characteristics

H_3 predicts that there would be a relationship between online processing of the multi-representational material in L1 and individual characteristics of interest; reading comprehension ability in English and prior knowledge. Table 16 displays the descriptive statistics of both individual characteristic tests for each cluster.

Table 16. Descriptive statistics of individual characteristic tests by cluster

Individual characteristic	High integrators *n* = 7		Intermediate integrators *n* = 17		Low integrators *n* = 9	
	M	*SE*	*M*	*SE*	*M*	*SE*
Reading comprehension ability	54.71	3.56	56.06	1.29	55.89	2.81
Prior knowledge	6.57	0.53	6.41	0.49	6.22	0.81

A MANOVA was carried out to evaluate whether or not the means of the three clusters significantly differed. No statistical difference emerged between the three clusters and the combined dependent variables, F (4, 60) = .09, p = .986; Pillai's Trace = .01, η_p^2 = .01. Table 17 reports the univariate tests.

Table 17. Analysis of variance for individual characteristic tests

Individual characteristic	*df*	*F*	η_p^2	*p*
Reading comprehension ability	30	.09	.01	.91
Prior knowledge	30	.06	.00	.94

A finer-grained analysis was performed to examine the relationship between each of the seven eye tracking measures and each individual characteristic. Again, no significant associations emerged (Table 18).

Table 18. Correlations between eye tracking measures and individual characteristic

Measure	Reading comprehension ability		Prior knowledge	
	r	*p*	*r*	*p*
FP-Pict	-.24	.18	-.15	.41
FP-Text	-.17	.35	.04	.82
Transitions	-.09	.62	-.02	.93
LF-Pict Pict	-.11	.53	.22	.21
LF-Pict Text	-.15	.40	.00	.98
LF-Text Pict	-.10	.58	.09	.63
LF-Text Text	.08	.65	.18	.32

* eye tracking measures were transformed

5.4.4. Eye Movements and Learning Outcome

H_4 predicts a relationship between online processing of the multi-representational material in L1 and learning of the material. Specifically, L1 readers who integrated text and pictures to a greater extent would be those who learned better from the material. Table 19 displays the descriptive statistics of the learning outcome test for each cluster.

Table 19. Descriptive statistics of learning outcome test by clusters

High integrators *n* = 7		Intermediate integrators *n* = 17		Low integrators *n* = 9	
M	*SE*	*M*	*SE*	*M*	*SE*
19.07	1.38	16.53	0.86	14.5	0.95

Before comparing the performance of each cluster in the learning outcome test, the relationship between individual characteristics and learning outcome was first examined. *Table 20* details the correlations. Learning outcome did not significantly correlate with prior knowledge but did so with reading comprehension ability. Better readers were therefore those who learned better from the multi-representational text. Due to this significant correlation, reading comprehension ability was considered as a covariate in the next analysis.

Table 20. Correlations between individual characteristic and learning outcome test

Reading comprehension ability		Prior knowledge	
r	*p*	*r*	*p*
.39	.02*	.16	.39

The covariate was significantly related to learning outcome scores, $F(1, 29) = 8.15$, $p = .008$, $\eta_p^2 = .22$. After controlling for this ability, a significant difference between the clusters was found, $F(2, 29) = 5.15$, $p = .012$, $\eta_p^2 = .26$. Pairwise comparisons revealed that the high integrators ($M = 19.07$, $SE = 1.38$) performed significantly better in the test compared to the low integrators ($M = 14.61$, $SE = 0.89$). The mean difference is 4.71, $p = .01$. However, the high integrators did not perform significantly better than the intermediate integrators ($M = 16.65$, $SE = .81$), nor did the latter significantly outperformed the low integrators. The mean differences are 2.71, $p = .143$ and 2.00, $p = .317$, respectively.

5.5. DISCUSSION

Study 2 aimed to explore the reading and learning of an authentic multi-representational text in an L1 as an initial step towards comparing multi-representational learning of the same material in the L1 and the L2. To accomplish this, the L1 readers read the same material as the L2 readers in the previous study and with the exception of a couple of tasks (only relevant for L2 samples), went through the same procedures.

H_1: Reading of multi-representational texts in L1 will be text-dominant: Fixation frequencies and duration will be largely spent on the text portions rather than the pictures.

The large body of research on multi-representational learning in the L1 has demonstrated that multi-representational materials is read in a text-dominant fashion (Hannus & Hyönä, 1999; Johnson & Mayer, 2012; Rayner et al., 2001; Schmidt-Weigand et al., 2010). The L1 readers in this study were no exception. After controlling for differences in total surface area occupied by the text chunks and pictures, fixation data indicate that the L1 readers spent an average of 71% of their learning time on

the text areas and only 27% on the pictures. Additionally, 72% of their total fixations were on the text areas, whereas only 26% were on the pictures. These figures are very similar to what has been observed among the L2 readers in the previous study (refer to Section 4.5.1). These results provide an initial indication that L1 and L2 reading of the same authentic multi-representational material was similarly heavily text-based with the text areas attended to almost three times more than the picture areas.

H_2: There will be distinct patterns of eye movements when reading a multi-representational text in an L1, varying in terms of the extent to which text and pictures are integrated.

The modelled study by Mason, Tornatora, et al. (2013) has found that young Italian children displayed three distinct patterns of eye movements when reading a simple illustrated science text in their L1. Another study by the same authors found two distinct patterns. Regardless of the different numbers, in both studies, the patterns varied with regard to the extent to which text and pictures are integrated. Study 2 corroborated these findings, identifying three distinct patterns of eye movements among the native English speakers, mainly differing in terms of integrative processing. In fact, the main effects of clusters were evident for all five integrative processing measures including LF-Pict Pict. Accordingly, they were labelled high, intermediate or low integrators. The three types of integrators also differed with regard to first-pass processing of pictures. However, just like what was observed in Study 1, the pattern for this variable differed between Study 2 and Mason, Tornatora, et al.'s (2013). In Study 2, the high integrators spent significantly less time inspecting the picture areas whilst in the modelled study, the high integrators did the opposite.

The patterns found in Studies 1 and 2 were essentially similar despite the extra pattern and the main effect of clusters on LF-Pict Pict found in the present study. These results, again, indirectly indicate that the ways in which L1 and L2 readers processed the same authentic multi-representational material were rather similar. In each group, there would be those who extensively integrate text and pictures and those who do this less extensively.

H_3: There will be a relationship between online processing of the multi-representational material in L1 and individual characteristics of interest: Reading comprehension ability in English will correlate negatively with first-pass fixation time

on text. Prior knowledge will correlate positively with text-picture integrative measures.

Given that previous research has found reading comprehension ability and prior knowledge to predict or influence online processing of multi-representational materials (Mason, Tornatora, et al., 2013; 2015), both were expected to predict the cluster membership found for Study 2. This hypothesis however was not supported as the three clusters performed similarly on each individual variable measure. A finer-grained correlational analysis that examined the relationships between the individual characteristics and each eye tracking measure also revealed no significant results. Hence, just like what was found in Study 1, and contradictory to Mason, Tornatora, et al.'s (2013; 2015) studies, the individual characteristics examined could not account for how an authentic multi-representational text was read and processed. This lack of significant findings could be due to the small size of the L1 sample which made it difficult to find enough variation in the scores of each variable.

H_4: There will be a relationship between online processing of the multi-representational material in L1 and learning: L1 readers who integrate text and pictures to a greater extent will be those who learn better from the material.

Finally, it was predicted that a significant association between eye movement patterns and learning from the multi-representational text in L1 would emerge just like what has been observed in the literature (Hannus & Hyönä, 1999; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013; O'Keefe et al., 2014). Hence, the high integrators with the most strategic reading pattern, i.e. more frequent and longer integrative processing, were expected to outperform the intermediate and low integrators. Confirming this hypothesis, the high integrators (*M* = 19.07 or 63.57%) significantly outperformed the low integrators (*M* = 14.5 or 48.33%). They however did not outperform the intermediate integrators (*M* = 16.53 or 55.1%) nor did the latter outperformed the low integrators. Overall, these findings suggest that the eye movement pattern of the high integrators characterised by more frequent and longer attempts to integrate within and across representation types is more useful and strategic to learn from multiple representations.

5.6. CONCLUSION

Study 2 has demonstrated how an authentic multi-representational text is read and learned in the L1. In general, the study found many similarities in the ways L1 and L2 English readers in this study and the previous study read and learned the same material. Both studies found that multi-representational reading was heavily text-dominant with the text areas attended to almost three times more than the pictures were. Additionally, both studies found distinct reading patterns demonstrated by their participants. In general, the patterns varied for integrative processing, or the extent to which different representations, either within or across representation types. Another similarity observed was in the way that the cluster memberships were not influenced by individual characteristic variables; English reading comprehension ability and prior knowledge. Finally, both studies lent further support to the notion that integrating text and pictures is essential to learning from multiple representations. This was reflected in the superior performance of the high integrators compared to the low integrators. Although striking similarities can be observed in the ways that the two groups read, processed and learned from the same material, a direct comparison is needed to corroborate these findings. The following chapter presents the comparison study that essentially identifies the similarities and differences in the way the material is read, processed and learned in L1 and L2.

6. STUDY 3: THE COMPARISON STUDY

6.1. INTRODUCTION

Studies 1 and 2 have explored the reading and learning of an authentic multi-representational text in English by examining the eye movement patterns of two different language groups; the non-native English speakers (or the L2 readers) in Study 1 and the native English speakers (or the L1 readers) in Study 2. Findings from both studies suggest that both groups displayed similar reading behaviour. Reading in both cases was text-dominant and those who benefitted more from the multi-representational text integrated text and pictures to a greater extent. The next step is to directly compare and contrast the two groups in order to examine to what extent reading and learning of the multi-representational text in L1 and L2 is similar. Each dataset from the preceding studies was therefore combined for this purpose and analysed for Study 3. Since the ways in which readers engage and learn with multi-representational materials have been associated with their reading comprehension ability and prior scientific knowledge in the domain (or prior knowledge), Study 3 will first compare the two groups with regard to these two individual characteristics. In Study 1, neither measure significantly influenced the reading behaviour nor the learning outcome performance of the L2 readers. In Study 2, English reading comprehension ability did influence the L1 readers' learning outcome performance but not their reading behaviour. For the current study, the L1 readers are predicted to outperform their L2 counterparts in the English reading comprehension ability test given that it measures reading comprehension ability in the L1 readers' native language and the L2's weaker language. With respect to prior knowledge, since both groups were recruited from the same low-science background population, their level of prior knowledge is not expected to differ.

Unlike the preceding studies, Study 3 will examine the role of reading time on the readers' reading behaviour and learning outcome performance. It is highly plausible that the L1 readers would spend less time reading the material compared to the L2 readers because the former has the advantage of reading in their native language. In Studies 1 and 2, reading time was not accounted for given that the participants in each study were either all reading in their L2 or in their L1, therefore reading time within each group was not expected to vary by much. In Study 3 however, reading time is considered a factor given that the two groups are now combined and the L1

readers are more likely to read the material faster. The difference in reading time would likely lead to difference in reading behaviour since the group with the slower reading time would potentially process the material more than the faster group. If the two groups did significantly differ for average reading time, there would be a need to perform two sets of analysis, one that controls for variation in reading time and one that does not. This would be necessary as the L2 readers whom we might expect to have a longer average reading time could potentially use this extra time integrating within and across representations and end up being the group with the greater integrative processing.

In general, Study 3 predicts the L2 readers to spend a longer time on the material and display a greater extent of integrative processing than the L1 readers. Similarly, they are also predicted to spend longer times on first-pass processing of both text and pictures. Despite the predictions that the L2 readers would spend a longer time on the material and process it more, they however are not expected to outperform the L1 readers. The L2 readers are assumed to spend a longer time on the material and process it more to compensate and make up for having to read in their weaker language. Thus, engaging more in first-pass and integrative processing may compensate for this language disadvantage but not necessarily make them learn more than the L1 readers.

The preceding two studies have examined and found distinct clusters of readers mainly differing in terms of text-picture integrative processing. Furthermore, cluster membership has been found to predict the readers' learning outcome. Given these findings, a cluster analysis will be performed on the now combined L1 and L2 readers to examine whether such clusters are also distinguishable on a similar basis and subsequently predict the readers' learning outcome. In addition, Study 3 will examine whether or not this membership is dependent on the readers' language group. Similar to the preceding studies, the role of individual characteristics; English reading comprehension ability and prior knowledge on cluster membership will also be examined. It is expected that Study 3 clusters like the preceding two studies, will also differ in terms of text-picture integrative processing. However, the cluster membership is not expected to be dependent on language group. In other words, a participant's membership in a cluster does not depend on whether they are an L1 or an L2 reader. Additionally, neither individual characteristic is predicted to be related to cluster membership. Hence, the clusters as mentioned previously are expected to significantly differ strictly for reading behaviour but not for language group nor

individual characteristics of interest. Finally, the influence of cluster membership on the readers' learning outcome performance will be assessed. Since both of the preceding studies have shown the superiority of the high integrator clusters, the same results are also expected for Study 3. Specifically, the high integrator cluster who integrated text and pictures to a greater extent is predicted to outperform the low integrator cluster(s) in the learning outcome test.

6.2. METHOD

Study 3 combines the data from Studies 1 and 2. The participants were the thirty-two L2 readers from Study 1 and thirty-three L1 readers from Study 2, $N = 65$. They were all non-science students with low background knowledge in the science domains. For more details about their background, as well as the materials, procedure and measures involved, see chapters 4 and 5.

6.2.1. RESULTS

6.2.2. Comparing the L1 and L2 Readers

6.2.2.1. Differences in individual characteristics

Separate t-tests were run on the L1 and L2 readers' English reading comprehension ability and prior knowledge scores. Table 21 displays the means and standard errors. As predicted, the L1 readers had better English reading comprehension ability than the L2 readers, $t(63) = 3.57$, $p < .001$, $d = 0.89$. Unexpectedly, the L1 readers' prior knowledge was lower than the L2 readers', $t(63) = 2.06$, $p = .043$, $d = 0.51$.

Table 21. Descriptive statistics for individual characteristic variables by language group

Variable	L1 *n* = 33		L2 *n* = 32	
	M	*SE*	*M*	*SE*
English reading comprehension ability (scored out of 65)	55.73	1.22	50.28	0.91
Prior knowledge (scored out of 10)	6.39	0.34	7.47	.39

6.2.2.2. Differences in eye movements

To compare the L1 and L2 readers' reading behaviour, the same seven eye tracking measures used to compare the clusters in the preceding two studies were again used. Before the comparison was carried out, the possibility that the two groups would spend different amounts of time on the material was considered. It is highly plausible that the L1 readers would spend less time reading the material compared to the L2 readers, given that the former was reading in their native language. A t-test was hence performed and supporting this prediction, the L1 readers (M = 216.86 seconds, SE = 13.45 seconds) indeed spent less time on the material than the L2 readers (M = 358.78 seconds, SE = 18.45 seconds), $t(63) = -6.22$, $p < .001$, $d = 1.55$.

In line with these results, two sets of analyses were carried out. The first set analysed the eye tracking measures in absolute or raw reading time. Thus, the two groups' reading behaviour was compared irrespective of variation in reading time. To control for this variation, the second set of analysis standardised each of these measures and comparison was performed using proportional values instead. For convenience, the first analysis using absolute values will be called 'absolute' analysis, whereas the one using proportional values will be called 'proportional' analysis.

These two analyses address slightly different questions. The absolute analysis compares how long or how frequent each language group spent on the respective measures of reading behaviour. Since the L2 readers' reading time was longer, they could potentially process the material more than the L1 readers and hence display higher values for each of the measures. This comparison is fairly straightforward. If for instance the L2 readers spent an average of 11 000ms on first-pass processing of picture areas (i.e. FP-Pict), while the L1 readers spent an average of 9 000ms, then we can say that the L2 readers had a higher FP-Pict. The proportional analysis, on the other hand, takes into account the variation of reading time between the two groups. Therefore, the actual question addressed is 'how different did the L1 and L2 readers use and allocate their reading time (or saccades in the case of transitions) on each measure of the reading behaviour?'. In this case, the slower group will not necessarily display higher values on the individual measures. This further leads to interesting questions such as whether or not the readers used this reading time making proportionally more strategic reading behaviour and which group made better use of their reading time.

Absolute analysis

A MANOVA that included language group as the between-subject factor and the seven eye tracking measures in their raw values as dependent variables was performed. The dependent variables were first transformed using the square-root transformation due to their non-normal distributions. Table 22 displays the means and standard errors.

Table 22. Descriptive statistics for eye tracking measures by language groups

Measure	L1 *n* = 33		L2 *n* = 32	
	M	*SE*	*M*	*SE*
FP-Pict	88.33	6.77	82.35	7.78
FP-Text	193.10	6.10	209.98	9.09
Transitions	5.41	0.23	6.83	0.23
LF-Pict Pict	81.04	8.93	87.68	10.07
LF-Pict Text	115.64	9.77	165.94	9.15
LF-Text Pict	116.47	10.12	190.20	11.71
LF- Text Text	209.48	12.31	288.37	12.76

A large effect of language group was found, Pillai's trace= .421, F (7, 57) = 5.91, p < .001; η_p^2 = .42. The L1 and the L2 readers exhibited different reading behaviour when reading the multi-representational text. However, separate univariate ANOVAs showed significant language group effects only on four measures; transitions, LF-Pict Text, LF-Text Pict and LF-Text Text (refer to Table 23). The results indicate that overall, the L2 readers demonstrated a greater text-picture integrative processing than the L1 readers by making more frequent transitions between text and picture areas and spending a longer time re-inspecting picture areas during the re-reading of text areas and vice versa. They also spent more time on LF-Text Text, re-reading a text area during the re-reading of another. With regard to another integrative measure, i.e. LF-Pict Pict, no significant difference was found. Similarly, no significant differences were found for the two first-pass fixation measures indicating that the two language groups spent similar amount of time on both text and picture areas upon first encounter with the material.

Table 23. Analysis of variance for individual eye tracking measures

Measure	*df*	*F*	η_p^2	*p*
FP-Pict	63	0.34	.01	.564
FP-Text	63	2.41	.04	.126
Transitions	63	14.29	.19	.001
LF-Pict Pict	63	0.24	.00	.623
LF-Pict Text	63	14.10	.18	.001
LF-Text Pict	63	22.78	.27	.001
LF-Text Text	63	19.83	.24	.001

Proportional analysis

To control for variation in reading time, all seven eye tracking measures were expressed as proportions. All temporal measures (i.e. first-pass and look-from fixation times) are expressed as proportions of the total fixation duration on the material (Mason, Pluchino, et al., 2013) whereas the integrative transitions as proportions of all saccades on the material (Holsanova et al., 2009). After standardising these values, they were all transformed using the square-root transformation due to non-normal distributions.

Table 24 displays the means and standard errors for each language group. Do note however that the combined means of the six temporal measures in Table 24 (i.e. proportions of FP-Pict, FP-Text, LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text) do not add up to 1 for the following reasons. One, they were square-root transformed and two, not all second-pass temporal measures were analysed. In theory, all first-pass and second-pass temporal measures should add up to this value but for this research, only the finer-grained second pass measures were examined. These are the look-from fixation time measures which have been shown in previous studies to influence cluster memberships and learning outcomes.

Table 24. Descriptive statistics for eye tracking measures (proportions) by language groups

Measure	L1 *n* = 33		L2 *n* = 32	
	M	*SE*	*M*	*SE*
Proportion of FP-Pict	.22	.02	.16	.02
Proportion of FP-Text	.48	.02	.40	.02
Proportion of transitions	.20	.01	.20	.02
Proportion of LF-Pict Pict	.19	.02	.16	.02
Proportion of LF-Pict Text	.27	.02	.31	.01
Proportion of LF-Text Pict	.27	.02	.35	.02
Proportion of LF-Text Text	.49	.02	.54	.02

These proportional measures were then submitted to a MANOVA with language group as the between-subject factor. The test revealed a marginally significant effect of language group on the combined dependent variables, Pillai's trace = .20, F (7, 57) = 2.09, p = .06. The effect size however was large, η_p^2 = .204. Separate univariate ANOVAs showed significant effects only on three measures; proportions of FP-Pict, FP-Text and LF-Text Pict. The differences in proportions of LF-Pict Text and LF-Text Text between the two groups were also marginally significant (see Table 25).

Table 25. Analysis of variance for individual eye tracking measures (proportions)

Measure	*df*	*F*	η_p^2	*p*
Proportion of FP-Pict	63	5.49	.08	.022
Proportion of FP-Text	63	7.52	.11	.008
Proportion of transitions	63	0.36	.01	.551
Proportion of LF-Pict Pict	63	0.72	.01	.400
Proportion of LF-Pict Text	63	3.42	.05	.069
Proportion of LF-Text Pict	63	7.77	.11	.007
Proportion of LF-Text Text	63	3.55	.05	.064

These results suggest that when eye movements were compared relative to reading time, differences between the two groups were still evident. Again, the L2 readers as predicted displayed a greater extent of integrative processing than their L1 counterparts. In particular, they spent more of their reading time re-inspecting picture areas during the re-reading of text areas and vice versa. They also spent more of their reading time re-reading a text area during the re-reading of another. However, they

spent equal proportion of reading time as the L1 readers, re-inspecting a picture area while re-inspecting another. The two groups also made an equal proportion of saccades between text and picture areas. When the means of their eye tracking measures were scrutinised, we could see that the L2 readers spent more than half of their reading time (approximately 57%) on integrative processing within and across representation types[4]. Even though the L1 readers also spent a fair amount of their reading time (approximately 47%) on this processing, they did so for a significantly shorter time.

The two groups also displayed different amounts of first-pass processing[5] and contrary to earlier prediction, the L1 readers were the ones displaying greater processing on both text and picture areas. They spent more than a quarter of their reading time, approximately 30% on this processing compared to L2 readers' 21%. To summarise, these findings found that overall, the L2 readers used and allocated significantly more of their reading time than the L1 readers did on the behaviour that has been associated with positive learning outcomes. The next analysis examines whether differences in the two groups' reading behaviour are reflected in their learning of the multi-representational text.

6.2.2.3. Differences in learning outcome

The L1 and L2 readers' scores on the learning outcome test were compared and the t-test results found that the L1 (M = 16.61, SE = 0.61) and the L2 readers (M = 17.95, SE = 0.62) performed similarly on the test, $t(63) = -1.55$, $p = .125$, $d = 0.38$.

Considering that these scores may be influenced by the readers' English reading comprehension ability, prior knowledge and reading time, Pearson's correlations were carried out. Table 26 details the correlation results.

[4] Integrative processing measures here included proportions of LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text which were all expressed as proportions of the total fixation duration. Proportion of integrations was excluded as it was expressed as proportion of total saccades.

[5] First-pass processing measures included proportions of FP-Pict and FP-Text which were expressed as proportions of the total fixation duration.

Table 26. Correlations between learning outcome test and individual characteristic and reading time

Reading comprehension ability		Prior knowledge		Reading time	
r	*p*	*r*	*p*	*r*	*p*
.07	.582	.28	.026	.44	.000

Overall, performance on the learning outcome test was positively related with prior knowledge and reading time. Those who learned more from the material were also those who had more prior knowledge and those who spent more time reading the material. Both variables were therefore entered as covariates in the subsequent ANCOVA that compared the two language groups' performances in the learning outcome test. Overall, prior knowledge did not predict the learning outcome performance, $F(1, 61) = 3.10$, $p = .084$, $\eta_p^2 = .05$ but reading time did, $F(1, 61) = 11.58$, $p = .001$, $\eta_p^2 = .16$. There was no significant effect of language group on learning outcome performance after controlling for the effects of both covariates, $F(1, 61) = 1.24$, $p = .269$, $\eta_p^2 = .02$. These results indicate that the L1 and the L2 readers performed similarly in the learning outcome test.

To summarise, this first section has directly compared the reading of an authentic multi-representational text in both L1 and L2. The L1 and L2 readers were not equal to begin with. The L1 readers obviously were better and more skilful at English reading comprehension, but unexpectedly had a lower level of prior scientific knowledge in the domain. They also read the multi-representational text faster than the L2 readers. Analyses conducted found significant differences in terms of how the two groups read the text. In particular, the L2 readers displayed the more strategic reading behaviour with respect to learning outcomes by making greater integrative processing both within and across representation types than the L1 readers. This was true even when reading time was controlled for, although the main effect of language group did decline for the proportional analysis. Surprisingly, in spite of the advantage of reading in their native tongue, the L1 readers spent more of their shorter reading time on first-pass processing than the L2 readers. Despite overall differences in reading behaviour, the two groups performed similarly in the learning outcome test. This suggests that in order to attain the same learning outcome as the L1 readers, the L2 readers needed to spend more time on the text and use this time not only to re-read

and integrate verbal elements of the text but also to integrate them with pictorial elements to a greater extent than the L1 readers did.

6.2.3. Patterns of Reading Behaviour

The preceding two studies have examined and found distinct clusters of readers mainly differing in terms of text-picture integrative processing. Furthermore, cluster membership has been found to predict the readers' learning outcome. Given these findings, clustering procedures were carried out to examine whether distinct subgroups could be identified on a similar basis, irrespective of whether English is one's L1 or L2. The same seven eye tracking measures used as the cluster variables in the earlier studies were therefore retained to cluster the now combined L1 and L2 readers. Two cluster analyses were carried out. The first one, i.e. the 'absolute' cluster analysis used the raw eye tracking measures as cluster variables, whereas the second analysis, i.e. the 'proportional' cluster analysis used the proportional measures. The readers were therefore clustered slightly differently. In the absolute cluster analysis, clusters were formed based on how long and frequent readers spent on each eye tracking measure. In the proportional cluster analysis however, clusters were formed on the basis of how each reader allocated their unique, individual reading time (or saccades in the case of transition) on each of these measures. The results of each cluster analysis and the effects of clusters on reading behaviour, individual characteristics and learning outcome performance will be presented in the following subsections.

6.2.3.1. Absolute cluster analysis

All seven eye tacking measures were first transformed using the square root transformation, then standardised for the clustering procedure. To determine the final cluster solution, the dendrogram (see Figure 18) and proportions of change in heterogeneity were examined. As suggested in Hair et al. (2010), a possible solution materialises when the percentage of heterogeneity resulted from the continuous combination of clusters, is markedly different from one stage to the next (say from Stage 61 to Stage 62), and researchers are advised to opt for the previous stage (i.e. Stage 61). This is because the stage with the markedly different percentage (i.e. Stage 62) signals that two rather different clusters are being combined. In this absolute cluster analysis, Stage 62 which would cluster the participants into three, has the percentage of heterogeneity at 19.08, a markedly higher percentage than the

previous stage's 11.78%. Following Hair et al.'s suggestion, the four-pattern cluster solution with the 11.78% of heterogeneity was chosen. Table 27 reports each of the four clusters' means and standard errors for the eye tracking measures.

Table 27. Descriptive statistics for eye tracking measures by cluster

Measure	Cluster 1 (*n* = 19)		Cluster 2 (*n* = 14)		Cluster 3 (*n* = 25)		Cluster 4 (*n* = 7)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
FP-Pict	-0.54	0.13	-0.07	0.17	-0.18	0.12	2.26	0.21
FP-Text	-0.72	0.12	1.26	0.16	-0.27	0.16	0.39	0.31
Transitions	1.10	0.18	0.12	0.15	-0.67	0.11	-0.84	0.18
LF-Pict Pict	0.28	0.24	0.56	0.23	-0.24	0.17	-1.03	0.29
LF-Pict Text	0.88	0.11	0.64	0.21	-0.73	0.11	-1.10	0.17
LF-Text Pict	0.73	0.15	0.75	0.16	-0.71	0.14	-0.95	0.25
LF-Text Text	0.65	0.19	0.64	0.21	-0.61	0.14	-0.88	0.32

Similar to the preceding studies, the differences in both integrative and first-pass processing measures among clusters were examined to profile the clusters. However, in terms of labelling the clusters, a slightly different approach was taken. The labelling of Study 3 clusters would include both integrative and first-pass processing, rather than just the former. First-pass processing could be more influential for the present study than the preceding ones since Study 3 compared the L1 and the L2 readers directly. In addition to their integrative processing, first-pass processing was also expected to differ and affect cluster membership due to the L2 readers spending a longer time on the material, and their language disadvantage. In what follows, a description of each cluster is given. Although cluster order is arbitrary, it has been organised on the assumptions set out earlier, i.e. a strategic reading behaviour entails great extent of text-picture integrative processing and short first-pass processing time, whereas a poor reading behaviour entails low extent of text-picture integrative processing and long first-pass processing time. The clusters were therefore ordered from the best to the least strategic.

Cluster 1 groups readers who displayed a relatively high level of integrative processing as illustrated by the high means for transitions, LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text. They however spent the shortest amounts of time on

the first-pass processing of both text and picture areas. Cluster 1 readers are thus the High Integrator/ Low FP readers.

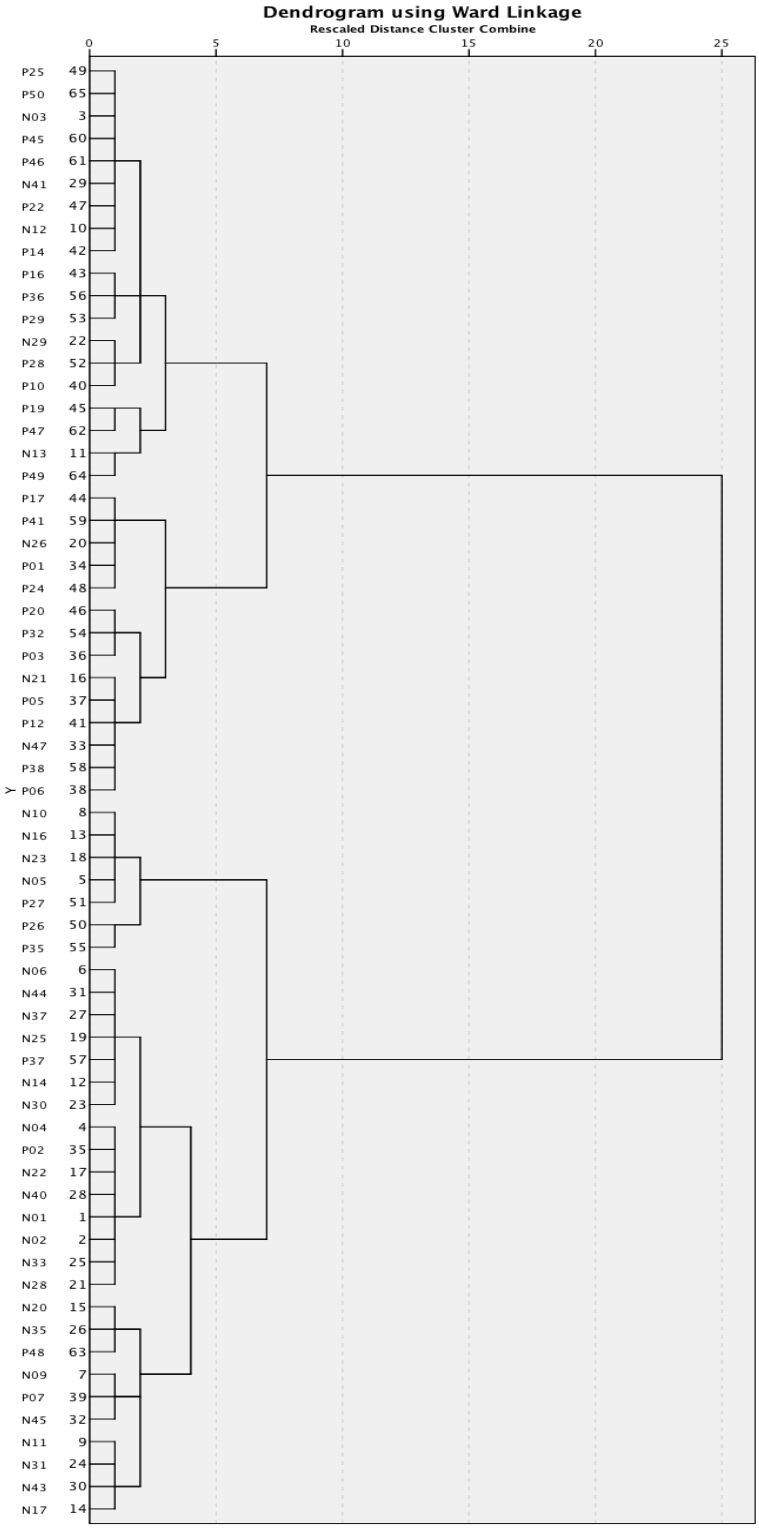


Figure 18. Dendrogram using Ward's linkage for the absolute cluster analysis

Cluster 2 readers, similar to those in Cluster 1 also demonstrated a relatively high level of integrative processing. Unlike those in Cluster 1, Cluster 2 readers also spent relatively long amounts of time on the first-pass processing of both text and picture areas as indicated by FP-Pict and FP-Text. Cluster 2 readers are thus the High Integrator/ High FP readers.

Those in Cluster 3 demonstrated a reading behaviour opposite to that of Cluster 2's. In particular, Cluster 3 readers' level of integrative processing was relatively low. The same goes for their first-pass processing of text and picture areas. Thus, they are the Low Integrator/ Low FP readers.

Cluster 4 readers' reading behaviour on the other hand was the opposite of those in Cluster 1. Their level of integrative processing is at the lowest among the four clusters and they spent relatively long amounts of time on the first-pass processing of both text and picture areas. They are therefore the Low Integrator/ High FP readers.

Subsequently, a MANOVA that included clusters as the between-subject factor, and the seven eye tracking measures as dependent variables was performed to evaluate whether the differences in eye tracking measures across the clusters were significant. It revealed a large main effect of clusters, Pillai's Trace = 1.98, F (21, 171) = 15.71, p < .001; η_p^2 = .66. Separate univariate ANOVAs showed significant effect of clusters on all seven eye tracking measures (refer to Table 28), indicating that the clusters differ not only for text-picture integrative processing but also for within-representation type integrative processing and first-pass processing.

Table 28. Analysis of variance for individual eye tracking measures

Measure	*df*	*F*	η_p^2	*p*
FP-Pict	61	38.92	.66	.000
FP-Text	61	24.63	.55	.000
Transitions	61	32.60	.62	.000
LF-Pict Pict	61	6.14	.23	.001
LF-Pict Text	61	39.44	.66	.000
LF-Text Pict	61	27.64	.58	.000
LF-Text Text	61	16.17	.44	.000

Post-hoc tests were carried out to assess pairwise differences among the clusters. Hochberg post-hoc test results were reported due to the very different sample size across the clusters (Field, 2013).

1. First-pass fixation time on pictures (FP-Pict)

The tests showed that the FP-Pict of the Low Integrator/ High FP readers was significantly higher compared to those from the other clusters, all *p*s < .001. No other significant differences were found, $p = .164$ between the High Integrator/ Low FP and the High Integrator/ High FP readers, and $p = .282$ between the High Integrator/ Low FP and the Low Integrator/ Low FP readers. The difference between the High Integrator/ High FP and the Low Integrator/ Low FP readers was also not significant, $p = .994$.

2. First-pass fixation time on text (FP-Text)

The High Integrator/ High FP readers' FP-Text was significantly higher than those from the other clusters, both *p*s < .001 with the High Integrator/ Low FP and the Low Integrator/ Low FP readers, and $p = .048$ with the Low Integrator/ High FP readers. The Low Integrator/ High FP readers' FP-Text was also significantly higher than the High Integrator/ Low FP readers', $p = .003$, but did not significantly differ with the Low Integrator/ Low FP readers, $p = .162$. Finally, the High Integrator/ Low FP and the Low Integrator/ Low FP readers also did not significantly differ on this measure, $p = .175$.

3. Integrative transitions (transitions)

The High Integrator/ Low FP readers' text-picture transitions was significantly higher than those from the other clusters, all *p*s < .001. The High Integrator/ High FP readers also made significantly more text-picture transitions than the Low Integrator/ Low FP readers ($p = .003$) and the Low Integrator/ High FP readers ($p = .01$). No significant difference in the amount of transitions emerged between the Low Integrator/ Low FP and the Low Integrator/ High FP readers, $p = .987$.

4. Look-from picture to picture fixation time (LF-Pict Pict)

The LF-Pict Pict of the Low Integrator/ High FP readers was significantly lower than the High Integrator/ Low FP ($p = .009$) and the High Integrator/ High FP readers' ($p = .002$), but not significantly lower than the Low Integrator/ Low FP readers', $p = .228$. No other significant differences were found, $p = .941$ between the High Integrator/ Low FP and the High Integrator/ High FP readers, and $p = .303$ between the High Integrator/ Low FP and the Low Integrator/ Low FP readers. The difference between

the High Integrator/ High FP and the Low Integrator/ Low FP readers was marginally significant, $p = .056$.

5. Look-from picture to text fixation time (LF-Pict Text)

The High Integrator/ Low FP readers' LF-Pict Text was significantly higher than the Low Integrator/ Low FP and the Low Integrator/ High FP readers' (both ps < .001), but not significantly higher than the High Integrator/ High FP readers', $p = .827$. The High Integrator/ High FP readers' LF-Pict Text was also significantly higher than the Low Integrator/ Low FP and the Low Integrator/ High FP readers' (both ps < .001). Finally, the difference between the Low Integrator/ Low FP and the Low Integrator/ High FP readers was not significant, $p = .614$.

6. Look-from text to picture fixation time (LF-Text Pict)

The same pattern was also evident for the LF-Text Pict with the High Integrator/ Low FP readers displaying significantly higher LF-Text Pict than the Low Integrator/ Low FP readers and the Low Integrator/ High FP readers (both ps < .001). Their LF-Text Pict however was not significantly higher than the High Integrator/ High FP readers', $p = 1.00$, even though the latter's was higher than the Low Integrator/ Low FP and the Low Integrator/ High FP readers' (both ps < .001). Finally, the difference between the Low Integrator/ Low FP and the Low Integrator/ High FP readers was not significant, $p = .955$.

7. Look-from text to text fixation time (LF-Text Text)

Again, the same pattern was evident in the case of LF-Text Text. The High Integrator/ Low FP readers' was significantly higher than the Low Integrator/ Low FP and the Low Integrator/ High FP readers' (both ps < .001), but not significantly higher than the High Integrator/ High FP readers', $p = 1.00$. The High Integrator/ High FP readers' LF-Text Text was also significantly higher than the Low Integrator/ Low FP and the Low Integrator/ High FP readers' (both ps < .001). Finally, the difference between the Low Integrator/ Low FP and the Low Integrator/ High FP readers was not significant, $p = .955$.

To summarise, a four-cluster solution was selected from the cluster analysis and the MANOVA that was subsequently carried out confirmed the main effect of clusters on the combined eye tracking measures. Unlike the preceding two studies which found clusters distinguished mainly for text-picture integrative processing, the present four clusters were distinctive respective of all integrative and first-pass processing

measures. This was why, unlike Studies 1 and 2), the four clusters were labelled according to their members' levels of integrative and first-pass processing. Out of these four clusters, the post-hoc results indicate that the reading behaviour demonstrated by the Low Integrator/ High FP readers was very different from those demonstrated by the High Integrator/ Low FP and the High Integrator/ High FP readers. The Low Integrator/ High FP readers made a less integrative processing both within and across representation types during their reading of the material but those in the high integrator clusters made significantly more. In contrast, when the Low Integrator/ High FP readers spent a longer time on first-pass processing of text and picture areas, the high integrators spent a significantly shorter time. Such differences however were less evident between the high integrator clusters and another low integrator cluster, i.e. the Low Integrator/ Low FP readers. Similarly, the reading behaviours of readers in the two high integrator clusters and in the two low integrator clusters were less distinguishable from one another.

Overall, the Low Integrator/ High FP readers appeared to display what has been found to be a reading behaviour associated with poor learning outcomes. Studies 1 and 2 have found low integrator readers to perform most poorly in the learning of the multi-representational text. It is therefore predicted that the Low Integrator/ High FP readers would perform more poorly than the High Integrator/ Low FP and the High Integrator/ High FP readers due to their least strategic reading behaviour. The High Integrator/ Low FP readers on the other hand, since they demonstrated a reading behaviour associated with successful learning outcomes were expected to outperform readers from the two low integrator clusters. Before the clusters' performances in the learning outcome test were compared, the influences of language group, and respective individual characteristic variables on the cluster membership were first examined.

The role of language group and individual characteristics

There were in total 33 L1 readers and 32 L2 readers. Their distribution into clusters is reported in Table 29. As can be seen from this table, the first two clusters, i.e. the high integrator clusters, were dominated by the L2 readers. L1 readers on the other hand, dominated the low integrator clusters. To statistically determine whether this cluster membership was associated with the readers' language group, a Fisher's exact test was carried out. This test was used since some of the cells consisted of less than five participants. The test revealed a statistically significant association between the two variables, $p < .001$, suggesting that the distribution of participants across the clusters

was associated with whether they were an L1 or L2 reader. This distribution was therefore not random.

Table 29. Distribution of language groups by clusters

Cluster	L1	L2	Total
High Integrator/ Low FP	5	14	19
High Integrator/ High FP	3	11	14
Low Integrator/ Low FP	21	4	25
Low Integrator/ High FP	4	3	7

To test whether the cluster membership was influenced by individual characteristics, separate ANOVAs which included the clusters as the between-subject factor and English reading comprehension ability and prior knowledge as dependent variables were performed. Table 30 displays the mean scores of both individual characteristic tests for each cluster.

Table 30. Descriptive statistics of individual characteristic tests by clusters

Individual characteristic	High Integrator/ Low FP (*n* = 19)		High Integrator/ High FP (*n* = 14)		Low Integrator/ Low FP (*n* = 25)		Low Integrator/ High FP (*n* = 7)	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Reading comprehension ability	52.84	1.29	49.93	1.87	55.68	1.16	50.43	3.47
Prior knowledge	7.37	0.53	7.0	0.48	6.8	0.40	6.0	1.09

A significant effect of cluster was found on English reading comprehension ability, F (3, 61) = 2.91, p = .042. Hochberg post-hoc test found that the Low Integrator/ Low FP readers (M = 55.68, SD = 5.79) were more skilful in English reading comprehension than the High Integrator/ High FP readers (M = 49.93, SD = 1.87), but this difference was only marginally significant (p = .054). Given that the distribution of the language groups across the clusters was not random, this result was not surprising. With regard to prior knowledge, the effect of clusters was not significant, indicating that the four clusters as predicted have similar levels of prior knowledge, F (3, 61) = 0.72, p = .542.

The four clusters were then compared to see whether they performed differently in the learning outcome test. As detailed above, the Low Integrator/ High FP readers were predicted to perform more poorly than those in the High integrator clusters. In contrast, the High Integrator/ Low FP readers were predicted to outperform those in the low integrator clusters. Each cluster's mean for the learning outcome test is displayed in Table 31.

Table 31. Descriptive statistics of learning outcome test by clusters

High Integrator/ Low FP (*n* = 19)		High Integrator/ High FP (*n* = 14)		Low Integrator/ Low FP (*n* = 25)		Low Integrator/ High FP (*n* = 7)	
M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
19.37	0.72	18.11	0.82	16.56	0.59	12.43	0.65

Given the previously reported effects of English reading comprehension ability and prior knowledge, both variables were considered as covariates in the subsequent analysis. The ANCOVA with clusters as the between-subject factor and scores in the learning outcome test as the dependent variable found that neither covariates were significantly related to learning outcome scores. For English reading comprehension ability, $F(1, 59) = 0.52$, $p = .474$, $\eta_p^2 = .01$ and for prior knowledge, $F(1, 59) = 2.79$, $p = .100$, $\eta_p^2 = .05$. A large effect of clusters was found after controlling for both covariates, $F(3, 59) = 9.18$, $p < .001$, $\eta_p^2 = .32$. Pairwise comparisons revealed that as predicted, the Low Integrator/ High FP readers who displayed the least strategic reading behaviour performed more poorly than the High Integrator/ High FP ($p = .001$) and the High Integrator/ Low FP readers ($p < .001$). They also performed more poorly than the Low Integrator/ Low FP readers ($p = .036$). Also, as predicted, the High Integrator/ Low FP readers with the most strategic reading behaviour outperformed the other low integrator cluster, i.e. the Low Integrator/ Low FP ($p = .021$). They however did not outperform the other high integrator readers, i.e. the High Integrator/ High FP ($p = 1.00$). This section has looked at the clusters of L1 and L2 readers formed based on the raw or absolute eye tracking measures. Four distinct clusters were identified and main effects of clusters were found on each of these measures. Cluster membership however was dependent on language group with the L2 readers

dominating the high integrator clusters, and the L1 readers, the low integrator clusters. In terms of individual characteristics, a main effect of cluster was found on English reading comprehension ability but not for prior knowledge. This effect was also significant on the readers' learning outcome performance with the Low Integrator/ High FP readers performing more poorly than the other clusters and the High Integrator/ Low FP readers outperforming both the low integrator clusters. The next section reports the results of the proportional cluster analysis. Results of subsequent analyses following the clustering procedure were also detailed.

6.2.3.2. Proportional cluster analysis

For this analysis, clusters were formed based on how each reader allocated their unique, individual reading time (or saccades in the case of transition) on the selected measures of reading behaviour. The seven eye tracking measures used in the absolute analysis were therefore computed into proportions and used as the cluster variables. They were first transformed using the square-root transformation, then standardised for the clustering procedure. Again, the dendrogram (see Figure 19) and proportions of change in heterogeneity were examined the proportions of change in heterogeneity were examined to determine the final cluster solution. In this case, a three-pattern cluster solution was selected because a two-pattern cluster solution was associated with a markedly different percentage of heterogeneity than the previous combination of three clusters (i.e. 10.64% at a three-cluster solution versus 18.29% at a two-cluster solution). Table 32 reports each cluster's means and standard errors for the eye tracking measures.

Table 32. Descriptive statistics for proportions of eye tracking measures by clusters

Measure	Cluster 1 (*n* = 36)		Cluster 2 (*n* = 16)		Cluster 3 (*n* = 13)	
	M	*SE*	*M*	*SE*	*M*	*SE*
FP-Pict	-.59	.07	.16	.17	1.43	.27
FP-Text	-.69	.11	.50	.10	1.29	.19
Transitions	.26	.16	.04	.21	-.78	.24
LF-Pict Pict	.10	.14	.54	.21	-.94	.26
LF-Pict Text	.57	.10	-.46	.21	-1.02	.27
LF-Text Pict	.52	.10	.05	.18	-1.50	.18
LF-Text Text	.46	.13	-.35	.23	-.85	.28

Cluster 1 readers demonstrated the highest level of integrative processing compared to those in the other two clusters. This can be seen from the high means for proportions of transitions, LF-Pict Pict, LF-Pict Text, LF-Text Pict and LF-Text Text. In contrast to their integrative processing, they spent the least proportions of reading time on the first-pass processing of text and picture areas. These readers were therefore the High Integrator/ Low FP readers.

Cluster 2 groups readers whose proportions of integrative processing was at the intermediate level, between Clusters 1 and 3. Similar to their integrative processing, these readers' proportions of first-pass processing on both text and picture areas was at the intermediate level. They were thus, the Intermediate Integrator/ Intermediate FP readers.

Cluster 3 readers' reading behaviour was the opposite of those from Cluster 1 as their proportions of integrative processing was at the lowest level. Additionally, they spent the highest proportions of reading time on the first-pass processing of text and picture areas. They were therefore the Low Integrator/ High FP readers.

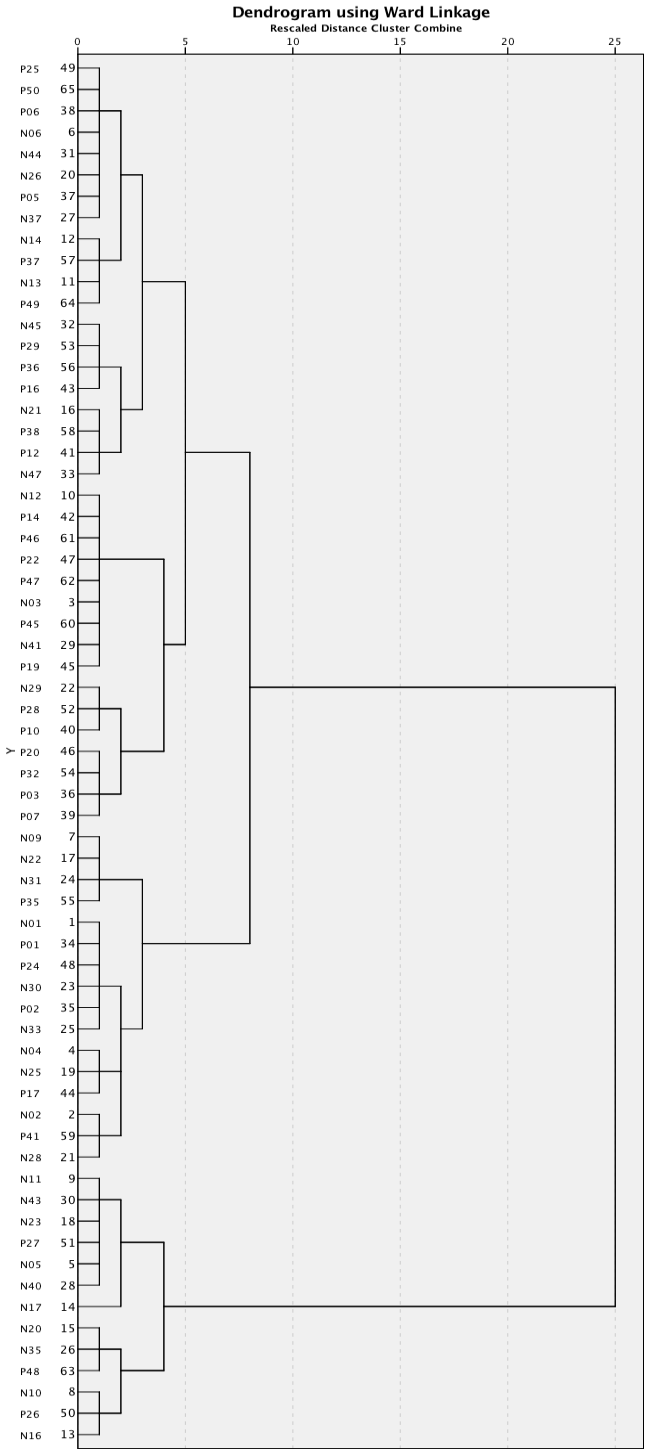


Figure 19. Dendrogram using Ward's linkage for the proportional cluster analysis

Subsequently, a MANOVA that included clusters as between-subject factor, and the seven eye tracking measures as dependent variables was performed to evaluate whether the differences in eye tracking measures across the clusters were significant. It revealed a large main effect of clusters, Pillai's Trace = 1.21, F (14, 114) = 12.51, p < .001; η_p^2 = .61. Separate univariate ANOVAs showed significant effects on all seven eye tracking measures (refer to Table 33).

Table 33. Analysis of variance for individual proportions of eye tracking measures

Measure	*df*	*F*	η_p^2	*p*
FP-Pict	62	50.01	.62	.000
FP-Text	62	61.23	.66	.000
Transitions	62	6.01	.16	.004
LF-Pict Pict	62	10.64	.26	.000
LF-Pict Text	62	25.35	.45	.000
LF-Text Pict	62	48.97	.61	.000
LF-Text Text	62	13.25	.30	.000

Post-hoc tests were carried out to assess pairwise differences among the clusters. Hochberg post-hoc test was reported due to the very different sample size across the clusters (Field 2013). For the proportions of FP-Pict however, Games-Howell post-hoc test was reported instead due to the assumption of equality of variances not being met.

1. Proportions of FP-Pict

The Games-Howell post-hoc test showed that the Low Integrator/ High FP readers spent a significantly more proportion of time on FP-Pict than the Intermediate Integrator/ Intermediate FP (p = .002) and the High Integrator/ Low FP (p < .001) readers. Additionally, the Intermediate Integrator/ Intermediate FP readers spent a significantly greater proportion of time on FP-Pict than the High Integrator/ Low FP readers, p =.002.

2. Proportions of FP-Text

Post-hoc tests showed the same pattern for the proportions of FP-Text with the Low Integrator/ High FP readers spending a significantly greater proportion of time on FP-Text than the Intermediate Integrator/ Intermediate FP (p = .002) and the High Integrator/ Low FP readers (p < .001). Additionally, the Intermediate Integrator/

Intermediate FP readers spent a significantly greater proportion of time on FP-Text than the High Integrator/ Low FP readers, $p < .001$.

3. Proportions of transitions

Post-hoc tests showed that the High Integrator/ Low FP readers made a significantly greater proportion of text-picture transitions than the Low Integrator/ High FP readers, $p = .003$. No other significant differences were found, $p = .798$ between the High Integrator/ Low FP and Intermediate Integrator/ Intermediate FP readers, and $p = .065$ between the Intermediate Integrator/ Intermediate FP and the Low Integrator/ High FP readers.

4. Proportions of LF-Pict Pict

Post-hoc tests showed that the Low Integrator/ High FP readers spent a significantly smaller proportion of time on LF-Pict Pict than the Intermediate Integrator/ Intermediate FP ($p < .001$) and the High Integrator/ Low FP readers ($p = .002$). There was no significant difference between the Intermediate Integrator/ Intermediate FP and the High Integrator/ Low FP readers on this measure, $p = .269$.

5. Proportions of LF-Pict Text

Post-hoc tests showed that the High Integrator/ Low FP readers spent a significantly greater proportion of time on LF-Pict Text than both the Low Integrator/ High FP and the Intermediate Integrator/ Intermediate FP readers (both ps $< .001$). There was no significant difference between the Intermediate Integrator/ Intermediate FP and the Low Integrator/ High FP readers ($p = .138$).

6. Proportions of LF-Text Pict

Post-hoc tests showed that the High Integrator/ Low FP readers spent a significantly greater proportion of time on LF-Text Pict than both the Low Integrator/ High FP ($p < .001$) and the Intermediate Integrator/ Intermediate FP readers ($p = .05$). The Intermediate Integrator/ Intermediate FP readers also spent a significantly more proportion of time on this measure than the Low Integrator/ High FP readers, $p < .001$.

7. Proportions of LF-Text Text

Post-hoc tests showed that the High Integrator/ Low FP readers spent a significantly greater proportion of time on LF-Text Text than the Intermediate Integrator/ Intermediate FP ($p = .007$) and the Low Integrator/ High FP readers ($p < .001$). There

was no significant difference between the Intermediate Integrator/ Intermediate FP and the Low Integrator/ High FP readers, p = .311.

To summarise, a three-cluster solution was selected from the cluster analysis and the MANOVA that was subsequently carried out confirmed the main effect of clusters on the combined eye tracking measures. Separate ANOVAs showed that the three clusters significantly differed for all integrative and first-pass processing measures. Out of these three clusters, the post-hoc results indicate that the reading behaviour of the High Integrator/ Low FP and Low Integrator/ High FP readers were very different from one another. The High Integrator/ Low FP readers made a greater proportion of integrative processing both within and across representation types but the Low Integrator/ High FP readers' was significantly lower. In contrast, when the Low Integrator/ High FP readers spent more of their reading time on first-pass processing of text and picture areas, the High Integrator/ Low FP readers made significantly less. The reading behaviour of the Intermediate Integrator/ Intermediate FP readers on the other hand, was constantly somewhere in between the High Integrator/ Low FP and Low Integrator/ High FP readers', but not always distinguishable from the latter two. Only for first-pass processing did the Intermediate Integrator/ Intermediate FP readers differed significantly from the High Integrator/ Low FP and Low Integrator/ High FP readers as their proportions of first-pass processing was higher than the High Integrator/ Low FP readers' but lower than the Low Integrator/ High FP readers'. For integrative processing, the differences were not as conclusive.

The High Integrator/ Low FP readers overall appeared to display the most strategic reading behaviour than the Intermediate Integrator/ Intermediate FP and Low Integrator/ High FP readers and were therefore expected to perform better in the learning outcome test. Significant difference was only predicted between the High Integrator/ Low FP and Low Integrator/ High FP readers because the Intermediate Integrator/ Intermediate FP readers' reading behaviour as mentioned above was less distinguishable from the rest. Before the clusters' performances in the learning outcome test were compared, the influences of language group, and respective individual characteristic variables on the cluster membership were first examined.

The role of language group and individual characteristics

The distribution of L1 and L2 readers into the proportional clusters is reported in Table 34. As can be seen from this table, majority of the L2 readers were in the High

Integrator/ Low FP cluster. The L1 readers on the other hand, were more equally distributed across the three clusters.

Table 34. Distribution of language group by clusters

Cluster	L1	L2	Total
High Integrator/ Low FP	13	23	36
Intermediate Integrator/ Intermediate FP	10	6	16
Low Integrator/ High FP	10	3	13

To statistically determine whether this cluster membership was associated with the readers' language group, a Chi-square test was carried out. The test revealed a statistically significant association between the two variables, $X^2 = 7.53$, $p = .023$, suggesting that the distribution of participants across the clusters was associated with whether they were an L1 or L2 reader. This distribution was therefore not random.

The three clusters were then compared to see whether they were significantly different with regard to English reading comprehension ability and prior knowledge. Table 35 displays each cluster's mean scores on both individual characteristic tests.

Table 35. Descriptive statistics of individual characteristic tests by clusters

Individual characteristic	High Integrator/ Low FP (n = 36)		Intermediate Integrator/ Intermediate FP (n = 16)		Low Integrator/ High FP (n = 13)	
	M	*SE*	*M*	*SE*	*M*	*SE*
English reading comprehension ability	51.97	1.09	53.94	1.59	54.9	2.05
Prior knowledge	7.31	0.36	6.50	0.46	6.38	0.67

Separate ANOVAs that included the clusters as the between-subject factor and the scores on the English reading comprehension ability and prior knowledge tests as dependent variables were performed. Results revealed that the effects of clusters on both variables were not significant. For English reading comprehension ability, F (2, 62) = 1.12, p = .332 whereas for prior knowledge, F (2, 62) = 1.30, p = .281. The three clusters therefore did not vary for their levels of English reading comprehension ability and prior knowledge.

The three clusters were then compared to see whether they performed differently in the learning outcome test. As mentioned earlier, the High Integrator/ Low FP readers who displayed the most strategic reading behaviour were predicted to outperform the Low Integrator/ High FP readers with the least strategic reading behaviour. Table 36 reports each cluster's mean score and standard errors on the learning outcome test.

Table 36. Descriptive statistics of learning outcome test by clusters

High Integrator/ Low FP (n = 36)		Intermediate Integrator/ Intermediate FP (n = 16)		Low Integrator/ High FP (n = 13)	
M	*SE*	*M*	*SE*	*M*	*SE*
18.57	0.48	17.06	0.97	13.92	0.75

Given that performance on the learning outcome test was previously reported to significantly correlate with prior knowledge (r = .28, p = .026), it was considered as a covariate in the subsequent analysis. The ANCOVA with clusters as the between-subject factor and scores in the learning outcome test as the dependent variable found that the covariate was not significantly related to learning outcome scores, F(1, 61) = 3.23, p = .077, η_p^2 = .05. A large effect of clusters was found after controlling for this covariate, F(2, 61) = 9.53, p < .001, η_p^2 = .24. Pairwise comparisons revealed that as expected, the Low Integrator/ High FP readers performed worse than the High Integrator/ Low FP readers (p < .001). In addition, they also performed worse than the Intermediate Integrator/ Intermediate FP readers did (p = .024). The difference between the Intermediate Integrator/ Intermediate FP and the High Integrator/ Low FP readers was not significant (p = .547).

6.3. DISCUSSION

Study 3 aimed to directly compare and contrast the ways in which the L1 and L2 readers read and learn from an authentic multi-representational text in English. Findings from the preceding Studies 1 and 2 have indirectly suggested that these two groups displayed similar reading behaviour. Results of Study 3 revealed that this was not entirely the case.

To compare the two groups' reading behaviour, difference in average reading time was first taken into account. The difference was significant, hence, reading behaviour was compared from two angles; when reading time varied and eye movements were therefore examined in absolute or raw time (i.e. the absolute analysis), and when reading time was controlled for and eye movements were proportionate to reading time (i.e. the proportional analysis). These two analyses addressed slightly different questions. With the absolute analysis, the question of interest was whether the L1 and L2 readers differed in terms of how long or how frequently they displayed each reading measure. The relevant question for the proportional analysis on the other hand, was whether or not the L1 and L2 readers use and allocate their reading time (or saccades in the case of transition) on each reading measure differently. This further leads to interesting questions such as whether the readers used this reading time strategically and which group made better use of their reading time.

For the absolute analysis, the L2 readers with the longer average reading time made more shifts between text and picture areas, spent longer time re-reading text while re-inspecting pictures and re-inspecting pictures while re-reading text. They also spent a longer time looking back and forth between text portions and integrating them than the L1 readers did. This reading behaviour may not seem that surprising given that it was highly plausible that slower readers as opposed to faster readers would process a text more. However, since the slower group in this case was the L2 readers, it would be reasonable to expect them to spend additional reading time on text areas and on LF-Text Text in particular to compensate for the fact that they were reading in their less familiar language. Additionally, Schmidt-Weigand et al. (2010) demonstrated in their study that learners who had control over the pace of instruction spent extra instruction time on text rather than on pictures or alternating between text and pictures. Indeed, the absolute analysis revealed that the L2 readers did spend a longer time on LF-Text Text than the L1 readers. Surprisingly, this additional reading

time was not solely dedicated to text processing. The L2 readers also demonstrated a greater extent of text-picture integration by making more transitions between text and picture areas and spending a longer time re-reading text while re-inspecting pictures and re-inspecting pictures while re-reading text. It is therefore encouraging that the pictures were not ignored and the extra time was spent by engaging in what is considered as strategic reading behaviour (Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013).

For the proportional analysis, this main effect declined but remained nonetheless. Overall, the L2 readers still demonstrated the more strategic reading behaviour. They dedicated longer proportions of reading time on integrating text and pictures than their L1 counterparts. The main effect of language on first-pass processing also came out significant and interestingly, the L1 instead of the L2 readers spent longer proportions of time reading and inspecting the material upon first encounter. One would expect the L1 readers to spend shorter proportions of time on first-pass processing because reading comprehension ability is a resource during text-processing and especially during first-pass reading (Mason, Tornatora, et al., 2013). Hence, it can be presumed to facilitate the more skilled L1 readers to process the text. Additionally, longer first-pass fixation time on text areas usually indicates comprehension difficulty (Rayner et al., 2006), and it was plausible to predict that the L2 readers would read the text longer than the L1 readers for the first time because they would experience more difficulties with the text. Even though the L2 readers did not spend longer proportions of time on first-pass processing, they did however reread and reinspect the representations and integrate them for proportionately longer, possibly to compensate for the shorter time they spent on the representations during first-pass. Overall, it appears that the L1 and the L2 readers differed in the ways they read and processed the same multi-representational material, with the L2 readers demonstrating the more strategic and ‘effortful’ reading behaviour.

This difference could be due to several reasons. Firstly, the L1 readers had an advantage over the L2 readers since they were reading in their native language. This was evident in the L1 readers’ superior English reading comprehension scores, reading time and reading patterns (i.e. less and shorter fixations). The L2 readers' greater integrative processing may also be an indication of their less fluent L2 reading and serve as a compensatory strategy in L2 reading. However, this reading behaviour were nevertheless useful for learning and the extra time and effort put into such behaviour was not in vain.

Perhaps prior knowledge could also explain these differences since the L2 readers did outperform the L1 readers in the prior knowledge task. Given this advantage, the L2 readers may have more resources available to be spent at integrating text and pictures (Mason, Tornatora, et al., 2013). Supporting this, a significant correlation between prior knowledge and the time spent re-reading text during the re-inspection of pictures was found. Prior knowledge was also found to be marginally related with the amount of text-picture transitions. Both correlations however were rather weak; .26 and .25 respectively, and given that the main effect of language group on prior knowledge was far lower than its effect on English reading comprehension ability (effect sizes, d = .51 versus .89), it appears that prior knowledge had less of an impact on reading behaviour.

Even though the L1 readers did not integrate information for as long as the L2 readers did, they did spend a proportionately longer time at integrative than first-pass processing (47% as opposed to 30%). This suggests that they did not underestimate the complexity of the material, and this was further supported by the fact that they made similar amount of transitions between text and pictures as the L2 readers. The proportion of time they spent at integrative processing however was significantly shorter than that of the L2 readers. Since the look-from fixation times are essentially the duration of fixations on text and picture areas, and the L1 readers' average fixation durations on either text or picture areas are in general, significantly shorter than the L2 readers', these findings are therefore, after a closer scrutiny, sensible. Hence, the L1 readers were not underestimating the material, they did indeed spend a fair proportion amount of time processing, reprocessing and integrating the representations.

Another issue of interest was whether the L1 and the L2 readers also performed differently in the learning outcome test. The results did not find significant difference in the test scores, indicating that the two groups of readers performed similarly in the test although each displayed unique reading behaviour. This finding suggests that the L2 readers need to exert more effort and time in reading text as well as relating and integrating it with pictures to learn equally well as their L1 counterparts. Such strategic reading behaviour could potentially compensate for the L2 readers' English disadvantage as well.

The study also examined to what extent distinct reading patterns could be identified solely based on the readers' reading behaviour. To address this question, two sets of cluster analysis were therefore carried out; one using the absolute eye tracking measures and the other using the proportional measures. The two analyses identified different number of clusters, but in both analyses, the resulting clusters can be clearly distinguished in terms of the readers' integrative and first-pass processing. However, these were not the only variable predicting the cluster membership as English reading comprehension ability was also found to be a significant predictor, especially for the four clusters derived from the absolute cluster analysis. When the role of language group was scrutinised for both cluster analyses, it was observed that the distribution of the L1 and L2 readers across clusters were not random, with the high integrator clusters being dominated by the L2 readers.

Studies 1 and 2 found that cluster membership was predictive of learning outcome and irrespective of individual characteristics, readers in the high integrator clusters in both studies learned best from the multi-representational text than the lower integrator clusters. Study 3 also supported these findings as those in the lowest integrator clusters performed worst in the learning outcome test, whereas those in the higher integrator clusters performed best. First-pass processing was found to be equally important in predicting the learning outcome although its relationship with learning outcome performance appeared inversive. In particular, those in the higher first-pass clusters performed worse in the learning outcome test, whereas those in the lower first-pass clusters performed best. Again, this research found that making more integrative processing is associated with better learning. What is more interesting here is that those making this strategic reading behaviour are disproportionately the L2 readers who dominated the high integrator clusters in both analyses. These findings along with findings from Studies 1 and 2 therefore support the crucial role of integrating text and pictures when learning from multi-representational texts which have been emphasised again and again in the literature. This thesis has put forward that text-picture integration might be equally and perhaps more crucial when reading in one's second language.

6.4. CONCLUSION

Study 3 has thus found that the L1 and L2 readers read the same authentic multi-representational text in different ways. By examining both groups' eye movements, it

was found that the L2 readers' reading time was longer, and they displayed greater extent of text-picture integrative processing as well as text processing than their L1 counterparts. When variation in reading time was controlled, the L2 readers again demonstrated this reading behaviour, allocating more of their reading time than the L1 readers did on integrative processing. The L1 readers on the other hand allocated more of their reading time than the L2 readers, on first-pass processing. The extra time and effort the L2 readers put into such reading behaviour proved useful when their learning performance was compared to the L1 readers'. Even though the L2 readers did not outperform the other group, they did perform equally well, which was reassuring given their disadvantage in English. The L2 readers also dominated the high integrator clusters in both cluster analyses, the clusters that learned most successfully from the multi-representational text. The results of this study further support the importance of integrating text and pictures to learn successfully from multi-representational text. The role of integration might perhaps be even more crucial when learning is in one's L2 as it could potentially compensate for the L2 disadvantage and this way they could potentially learn equally well in their L1 and L2.

7. GENERAL DISCUSSION

7.1. INTRODUCTION

In this chapter, the answers to the thesis' four research questions are discussed in turn. Next, the limitations of the research and opportunities for future work are outlined followed by the pedagogical implications of the findings. The final section describes the contributions of this thesis and provides some concluding remarks.

7.2. RESEARCH QUESTION 1

RQ1: How do readers of English as a second language (i.e. L2 readers) read an authentic multi-representational material in English?

This RQ explored the reading behaviour of the L2 readers by examining their eye movements when presented with a double page of multi-representational material from a science textbook. Essentially, two issues were addressed; whether L2 readers displayed heavily text-based processing of the material and whether different patterns of reading behaviour can be distinguished during this reading. With regard to the former, fixation data indicate that the L2 readers processed the text areas longer and more frequently than they did the picture areas. On average, the L2 readers spent 72% of their learning time on the text areas and only 27% on the pictures. Similarly, 73% of their total fixations were on the text areas, whereas only 25% were on the pictures. Effect sizes were large, d = 5.38 for proportion of fixation duration and 5.53 for proportion of fixation frequency. Hence, these figures indicate that multi-representational reading in the L2 is text-dominant with the text areas attended to almost three times more than the picture areas were even though almost half the page was composed of pictures.

This tendency, as reviewed in Chapter 2, has been widely observed when multi-representational materials are read in the readers' L1. However, having to read such materials in L2 might be harder as the text is read in the weaker language. The pictures could therefore be an alternative information source for the L2 readers to gain more understanding of the text (Ainsworth, 2006; Lee & Mayer, 2015). The large effect size however refuted this argument and it appears that reading of multi-representational materials in the L2 is similarly text-dominant as in the L1. Granted

that this tendency has been observed among L2 readers previously (Pellicer-Sanchez et al., 2018; Tragant Mestres & Pellicer-Sánchez, 2019; Yusof et al., 2019). However, the materials used in these studies were narratives in which representational pictures only make concrete the characters, actions and scenes and so may play much less of a role in learning than visual representations used in more expository and domain-specific materials like the two-page spread used in this research. Pre-dominant processing on the text versus the pictures was therefore sensible in narrative texts. In contrast, the pictures used in this research play a wider range of functions in supporting learning and therefore might invite more attention in the part of learners. Compared to Pellicer-Sánchez et al.'s (2018) and Tragant Mestres and Pellicer-Sánchez's (2019) findings, there was indication that this was indeed the case as it appears that the L2 readers processed the pictures for a shorter proportion of time than the L2 readers in the present research. Nevertheless, in all three works, readers still displayed a text-dominant processing of multi-representational materials when reading in their L2, suggesting that this tendency in addition to being applicable to narratives can also be extended to expository as well as more domain-specific materials in the L2.

Another unique element of the present research was it used authentic material comprising verbal and multiple visual representations as the stimulus, rather than a simplified illustrated text containing a piece of text and a solitary picture. Readers could therefore be presumed to divide their attention more uniformly across the text and picture areas. Although, Hannus and Hyönä's (1999) study has used similar authentic multi-representational materials and found a text-dominant tendency among their participants, it primarily focused on child readers. They could have spent a much longer time on the text as their reading skills were still developing or they could have less experience navigating multiple sources of information in an expository setting. However, the large effect size found indicates that the text-dominant processing when reading authentic multi-representational materials was also demonstrated by adult readers and may therefore be extended to a different population (i.e. adult readers) reading in their L2.

Another issue of interest was whether despite this text-dominant reading, the adult L2 readers would display distinct reading patterns as found in the L1 literature (Jian & Wu, 2015; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013). To address this, the clustering procedure used by Mason, Tornatora, et al. (2013) was adopted and seven eye tracking measures (both first-pass and integrative processing

measures) were input as the clustering variables. The cluster analysis identified two reading patterns and a further analysis revealed that the two can be distinguished from each other mainly in terms of integrative processing. More readers belonged to Cluster 1, and compared to those in Cluster 2, Cluster 1 readers made more transitions between text and picture areas (i.e. transitions), longer re-reading of text areas during the re-inspection of picture areas (i.e. LF-Pict Text) and vice versa (i.e. LF-Text Pict), and longer re-reading of a text area during the re-reading of another (LF-Text Text). They also inspected the picture areas for a shorter time upon first contact (i.e. FP-Pict). Since the two clusters essentially differed in terms of integrative processing, they were either labelled high or low integrators. Cluster 1 readers with higher means for this processing were therefore the high integrators, whereas Cluster 2 readers were 'low' integrators due to their lower means.

Instead of a two-cluster solution, Mason, Tornatora, et al. (2013) identified a three-cluster solution; high, intermediate and low integrators. Regardless, the two studies similarly found that the clusters significantly differed for the same five eye tracking measures. The same patterns were also observed, that is, the high integrators in both studies made more frequent transitions, and longer LF-Pict Text, LF-Text Pict, and LF-Text Text compared to the low integrators. However, the pattern for FP-Pict differed. In Mason, Tornatora, et al.'s (2013) study, the high integrators spent significantly more time inspecting the one picture in the illustrated text during initial reading, and the low integrators spent less time doing so. The high and low integrators in the present research, however, did the exact opposite with the high integrators spending a shorter time on FP-Pict and the low integrators spending more. This could be due to the different ways in which the two studies computed their second-pass measures, and essentially this occurred due to the different degrees of complexity in the materials used. In Mason's, a refixation on a particular AOI would be counted as second-pass only if all other seven AOIs have been attended to and exited. This is not the case for the present research, as any refixation on a particular AOI would be counted as second-pass after the AOI is exited once, and re-entered. There is no need for a participant to attend to all other 27 AOIs before a refixation is counted as a second-pass. Such a computation seems sensible with simplified materials like Mason's, as readers can be presumed to comprehend the overall content only if they have attended to all of the AOIs. It is not surprising therefore to find the child-readers to have a prolonged FP-Pict given that they could have made a number of refixations on the solitary picture, before they attended to all AOIs. Using this computation however would have been less sensible for this present research given its number of

AOIs. There would be very prolonged first-pass processing on some AOIs given that the readers might have attended to them multiple times before they completed a cycle (or all AOIs), and this would mean very short second-pass reading (or none at all) given that not all AOIs were re-entered (or entered at all). Given this complexity, a different computation was felt necessary and this might have contributed to the differential patterns in FP-Pict between the two studies. Having said that, the remaining measures, all of which are integrative measures, show identical patterns.

Another likely explanation would be that some of the pictures, and especially the Photo, contains verbal representations. These however were classified as picture AOIs. The overall fixations on these captioned pictures might have therefore included some fixations on the text, and not solely on the visual information.

To answer the RQ, readers of English as a second language read an authentic English multi-representational material in a text-dominant fashion. Yet, two reading patterns, i.e. high and low integrators can be distinguished especially with regard to the extent of integrative processing that occurred. These findings, although specific to the study sample (L2 sample with a variety of L1s) and material (a double page of multi-representational material from a science textbook), are nevertheless very similar to the ways in which multi-representational reading has been found to proceed in L1, therefore suggesting that existing models can be extended to reading of more authentic, expository materials containing multiple functional pictures by adult L2 readers.

7.3. RESEARCH QUESTION 2

RQ2: Is the reading behaviour of the L2 readers when studying an authentic multi-representational text in English related to individual characteristics in their prior knowledge, English reading comprehension ability and English vocabulary knowledge?

This RQ examines whether the grouping resulted from the cluster analysis can be explained by readers' pre-existing characteristics in prior knowledge and two English proficiency variables; English reading comprehension ability and English vocabulary knowledge. No significant relationships were found between any of the three measures of individual characteristics and the two emerging clusters. Similarly, there

were no significant relationships between any of the individual characteristic and eye tracking measures.

These non-significant findings are contrary to previous studies which have suggested the important roles of these variables on the processing of text and pictures. Moreover, given the complexity of the two-page spread and the fact that the readers were reading in their L2, these variables were expected to have more of an impact on eye movements. Text-picture integration, for instance, has been associated with prior knowledge as it equips readers with a schema which could be used to integrate representations (Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013; Scheiter et al., 2014). Mason, Tornatora, et al. (2013) found that low prior knowledge readers engaged in a less integrative processing within and across representations compared to readers with higher prior knowledge. A possible explanation behind this discrepancy might be that a reader of the two-page spread does not have to possess pre-existing knowledge about solid, liquid and gas properties in order to integrate the verbal and visual information in this material given that necessary information is already provided in the text-picture combination. Hence, it could be that, as suggested by Schnotz et al. (2017), text-picture integration (for this particular material, at least) is a domain-general competence, not requiring domain-specific prior knowledge.

Understandably, reading comprehension ability and vocabulary knowledge are important resources during the comprehension of any piece of text. They were predicted to account for some of the variance in how readers processed the two-page spread since it contains a substantial amount of text in the readers' L2. Systematic relationships between eye tracking measures such as number of fixations, total fixation time and FP-Text with comprehension difficulties have been recorded with higher values on these measures being indicative of comprehension issues (Rayner et al., 2006; Scheiter & Eitel, 2017). Mason, Tornatora, et al. (2013), for instance, has found less skilled readers spending a much longer FP-Text compared to the more skilled readers. This relationship was not evident in the present research. However, there was a marginally significant negative relationship between FP-Text and English vocabulary knowledge which suggests a tendency for readers with less versus more English vocabulary to read the text segments for a longer time during the initial reading. Although it was only a trend, this might mean that the text segments were difficult to the L2 readers at the local word level and contained less familiar English vocabulary. As such, FP-Text could be prolonged as readers took more time to identify such words. On the other hand, the non-significant relationship between

English reading comprehension ability and FP-Text might mean that the text components might be less demanding to comprehend at the more global sentence and paragraph levels. Hence, the readers on the whole might not need to spend much time on individual text segments during initial reading to understand its overall content. Another likely explanation is, there was not enough variation in the scores of both English proficiency measures to arrive at significant relationships, given that a UK university sample with a good overall command of English was used. Overall, these non-significant findings indicate that the grouping of participants, which was based on their reading behaviour, was not related to the three pre-existing characteristics of the readers; prior knowledge, English reading comprehension ability and English vocabulary knowledge.

7.4. RESEARCH QUESTION 3

RQ3: Is the reading behaviour of the L2 readers when studying an authentic multi-representational text in English related to the outcomes of learning?

For this RQ, the performances of each cluster on the learning outcome test were compared to examine whether the different ways in which the material was processed can be linked to different learning outcomes. The high integrators (63%) significantly outperformed the low integrators (53%) in the learning outcome test. This indicates that the eye movements characterised by more attempts to integrate verbal and visual information, longer integration of verbal and visual information during re-reading or re-inspection, and shorter inspection of pictures during first-pass were linked to better learning outcome. The opposite patterns as displayed by the low integrators, on the other hand, were linked to poorer outcome.

As reviewed in Chapter 2, many have argued for the importance of integrating across representations (Ainsworth, 2006; Mayer, 2005b; Renkl & Scheiter, 2015) and the positive relationship between integrative processing and learning outcomes has been widely observed in a number of L1 studies (Hannus & Hyönä, 1999; Johnson & Mayer, 2012; Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013; O'Keefe et al., 2014). When it comes to L2 reading however, the relationship might not be as straightforward. The L2 readers could have made extensive integrations not because they were trying to integrate information but because they were confused and had a hard time navigating the material. Since the high integrators in the present study also

had longer, not shorter, first-pass processing, this hypothesis was all the more plausible. However, we have seen from RQ2 that pre-existing characteristics; both English proficiency and prior knowledge could not explain why the L2 readers read in the way that they did. Finding that integrative processing was similarly associated with better learning outcome in the L2 thus lends further support to the idea that integration is key to learning successfully with multiple representations.

These findings correspond to Mason, Tornatora, et al.'s (2013) study which also found high integrators outperforming the low integrators. The strategic eye movement patterns displayed by the high integrators in Mason, Tornatora, et al.'s research correspond to the patterns displayed by the high integrators in this present research. The same was true for the low integrators in both works. Consequently, whether looking at children reading in L1 (as in Mason, Tornatora, et al., 2013) or adults reading in L2 as in this thesis they both were indicative of the pivotal role of integration when processing multi-representational materials for learning. Thus, integrative processing may not only be beneficial for child L1 readers to learn from simplified multi-representational materials, but may also support adult L2 readers to learn from more authentic materials containing a long piece of text accompanied by multiple functional pictures. To substantiate this, future work involving intervention studies could be performed to see if students can be taught integrative processing strategies and if so, whether or not their multi-representational learning outcomes improve. Several studies have observed that teaching L1 readers integrative processing resulted in them outperforming those in a control group who were not taught such strategies (Mason, Pluchino, & Tornatora, 2015, 2016).

7.5. RESEARCH QUESTION 4

RQ4: Are there differences in the reading behaviour and learning outcomes of the L1 and L2 readers when studying an authentic multi-representational text in English?

For this RQ, the reading behaviour and learning outcome data gathered from *the L2* and *the L1 studies* were combined and analysed. Comparison of individual characteristic variables indicates that the L1 versus the L2 readers, understandably, were better and more skilful at English reading comprehension ability, but this sample unexpectedly had a lower level of prior knowledge in the specific subject matter. They also read the multi-representational text more than 2 minutes faster than the L2

readers did. Due to the difference in reading time, absolute and proportional analyses were conducted to compare the two groups' reading behaviour. In the absolute analysis, L2 readers were found to demonstrate a greater text-picture integrative processing than the L1 readers as they made more frequent transitions between text and picture areas and spent a longer time re-inspecting picture areas during the re-reading of text areas and vice versa. They also spent more time re-reading a text area during the re-reading of another. When variation in reading time was controlled for in the proportional analysis, the difference in reading behaviour was only marginally significant but the effect size remained large. The L1 readers spent more of their reading time than the L2 readers did, processing both text and picture segments during first-pass. The L2 readers on the other hand, spent a greater proportion of their time than the L1 readers re-reading text areas during the re-inspection of pictures. They also spent more of their time on re-inspecting pictures during the re-reading of text areas and re-reading a text area during the re-reading of another but these when compared with the L1 readers' data were only marginally significant. Overall, both analyses found that the L2 readers displayed more strategic reading behaviour with respect to learning outcomes by displaying greater extent of text-picture integrative processing as well as text processing than their L1 counterparts. Despite overall differences in reading behaviour, the two groups performed similarly in the learning outcome test.

These findings therefore confirmed that the L2 and the L1 readers did not read and process the multi-representational material in identical fashion despite some similarities that were (indirectly) observed in Studies 1 and 2. Overall, it appears that the L2 readers in this study demonstrated more strategic and 'effortful' multi-representational reading than the L1 readers. What made these two groups read the same text differently? An obvious answer would be that the L1 readers read it in their L1 whilst the L2 readers, in their L2. The L1 readers therefore had a clear advantage over the L2 readers and this was reflected in their higher scores on the English reading comprehension ability test and their faster reading time. Moreover, the literature on bilingual text processing has indicated that people read less fluently in their L2 (Cop et al., 2015; Whitford & Titone, 2015). The L2 readers in this research did indeed display such reading behaviour by making more and longer fixations than the L1 readers did, resulting in a longer average reading time. This indicates that they were again, less fluent in reading English than the L1 readers. Interestingly, the two groups did not differ just on text processing (and re-processing, for that matter), as their text-picture integrative processing also differed, more so in the absolute analysis.

Their greater integrative processing may also be an indication of their less fluent L2 reading and serve as a compensatory strategy in L2 reading.

Alternatively, non-linguistic factors such as motivation, interest, and epistemic belief of knowledge may potentially explain why the L1 and L2 readers read the same text differently. Whatever the reason behind this difference may be, it is nevertheless encouraging that the pictures were not ignored and the extra time was spent by engaging in what is considered as strategic reading behaviour (Mason, Tornatora, et al., 2015; Mason, Tornatora, et al., 2013).

Although the effect of language group on reading behaviour was only marginally significant in the proportional analysis, the effect size remained rather large, and so were the effect sizes for integrative measures. Overall, the L2 readers still demonstrated the more strategic reading behaviour. They dedicated longer proportions of reading time on integrating text and pictures than their L1 counterparts. The difference between L1 and L2 readers on first-pass processing was significant and unexpected as the L1 not the L2 readers spent longer proportions of time reading and inspecting the material upon first encounter. It might be expected that the L1 readers would spend shorter proportions of time on first-pass processing because reading comprehension ability is a resource during text-processing and especially during first-pass reading (Mason, Tornatora, et al., 2013). Hence, it can be presumed to facilitate the more skilled L1 readers to process the text. Additionally, longer first-pass fixation time on text areas usually indicates comprehension difficulty (Rayner et al., 2006), and it was plausible to predict that the L2 readers would read the text longer than the L1 readers for the first time because they would experience more difficulties with the text. Even though the L2 readers did not spend longer proportions of time on first-pass processing, they did however reread and reinspect the representations and integrate them for proportionately longer, possibly to compensate for the shorter time they spent on the representations during first-pass.

Despite displaying different reading behaviour when reading the same multi-representational material, the L1 and L2 readers performed similarly in the learning outcome test. This suggests that in order to attain the same learning outcome as the L1 readers, the L2 readers needed to exert more effort and time. In addition to spending this extra effort and time to re-read and integrate verbal elements of the text, the L2 readers also needed to integrate them with pictorial elements to a greater extent than the L1 readers did. The findings suggest that they did not need to shift

their attention between text and pictures more frequently, but these transitions did need to be longer, especially when they are re-inspecting a picture, in an attempt to integrate its information with the verbal information that they have just reread. This reading behaviour could compensate for the L2 readers' English disadvantage and so they could learn equally well as the L1 readers. These findings along with findings from Studies 1 and 2 reiterate the crucial role of integrating text and pictures when learning from multi-representational texts which have been emphasised frequently in the literature. This thesis has put forward the argument that text-picture integration is equally and perhaps even more crucial when reading in one's L2, as it could possibly compensate for the L2 disadvantage and this way readers could potentially learn equally well in their L1 and L2.

RQ 4 also addressed whether distinct subgroups, irrespective of the readers' language group and individual characteristics could be identified on the basis of eye movement patterns alone. This required the L1 and L2 readers' data to be combined in a single cluster analysis. Two sets of cluster analysis were carried out; one using the absolute eye tracking measures and the other using the proportional measures. Although the two analyses identified different number of clusters, the resulting clusters can be clearly distinguished in terms of the readers' integrative and first-pass processing. With the absolute clusters, four clusters were identified; High Integrator/ Low FP, High Integrator/ High FP, Low Integrator/ Low FP and Low Integrator/ High FP. With the proportional clusters, three patterns were identified; High Integrator/ Low FP, Intermediate Integrator/ Intermediate FP and Low Integrator/ High FP. Prior knowledge was not found to be related to cluster membership, but English reading comprehension ability did, in particular for the four clusters derived from the absolute cluster analysis. Those in the Low Integrator/ Low FP cluster were marginally more significantly skilful in English reading comprehension than the High Integrator/ High FP readers. Given that the low integrator cluster was dominated by the L1 readers who were superior with respect to English reading comprehension ability, this finding was to be expected. In fact, in both cluster analyses, the distribution of the L1 and L2 readers across clusters were not random, with the high integrator clusters being dominated by the L2 readers. With regard to this matter, it might be necessary to pay a closer look at whether any one language group dominated the high integrator clusters as existing language or cultural practices might influence their L2 multi-representational reading. As discussed in the L2 study, although the Chinese readers dominated the high integrator cluster, they were also dominant in the sample as a whole and the number of other L1s in the cluster was almost completely in proportion.

Thus, no one language group was particularly dominant in this cluster or any other clusters for that matter. Differences therefore were unlikely to be solely due to existing cultural practices related to the L2 readers' native language. Future work however could control the proportions of the L2 sample's native languages more carefully so that each native language is represented more equally.

With regard to the link between reading behaviour and learning outcome, it was found that those in the lowest integrator clusters derived from both the absolute and proportion analyses performed worst, whereas those in the higher integrator clusters performed best. Additionally, first-pass processing was found to be equally important in predicting the learning outcome. Its relationship with learning outcome performance however appeared inverse, in contrast to the positive relationship between integrative processing and learning outcome performance. Those in the higher first-pass clusters derived from both the absolute and proportion analyses performed worse in the learning outcome test, whereas those in the lower first-pass clusters performed better. Again, this research found that making more integrative processing is associated with better learning. What is more interesting here is that those making this strategic reading behaviour are disproportionately the L2 readers who dominated the high integrator clusters in both analyses.

7.6. LIMITATIONS AND FUTURE WORK

This section outlines the limitations of the undertaken research and opportunities for future work as derived from the research findings and their discussion. The first limitation concerns the relatively modest sample sizes for each study. Although the overall sample sizes for each study were comparable with other studies', the present research involved group comparisons following cluster analyses. Having more participants therefore would have been more ideal in terms of increasing confidence in the results, especially with the four-cluster solution found in Study 3.

A second limitation also relates to sampling. Having to use an opportunity sample means that the researcher had to work within the constraints and limitations of the sample. For instance, having used a UK university sample meant that the L2 readers had a good command of English. Hence, the sample did not have much variability in terms of English proficiency. Secondly, the sample consists of small numbers of speakers from a variety of first languages. Having a sample with multiple L1s posed

a difficulty in terms of measuring the participants' L1 literacy, one of the major predictors of L2 reading. This would mean identifying and administering a literacy test in each of the languages, each of which needs to be comparable to the other. Thirdly, as seen in Study 3, there was an unexpected difference with regard to relevant prior domain knowledge between the two language groups, despite the significant effort made to exclude those with advanced science backgrounds. Having a larger and wider range of sample can potentially resolve these issues and generate opportunities for future work. For instance, sampling readers of a wider range of English proficiency could allow us to examine whether there would be more variety in the reading patterns exhibited by L2 readers. In addition, sampling more people from different L1 groups and even different reading conventions (e.g. right-to-left reading) could allow us to examine the effects of L1 on L2 multi-representational reading behaviour. As reviewed in Chapter 2, interactions between a bilingual's two languages occur during reading and comprehension. Sampling different L1 groups could help to uncover whether certain languages have conventions (linguistic or others) that cross over to implicate multi-representational reading in the L2 as well. This could potentially explain why in Study 3 the L2 rather than the L1 readers dominated the high integrator clusters.

The reading process in a given language and its efficiency can also be examined using brain imaging techniques. Psycholinguistic research, for instance, has immensely benefitted from techniques such as event-related potential (ERP) recordings and electroencephalography (EEG). ERP recordings, for one, have been used to examine the interrelations between bilinguals' two languages and even the processing of multimodal stimuli (see Kutas and Federmeier (2011) for a review). Since these techniques enable direct observation of brain activities during complex cognitive tasks and can be (and have been) used in a variety of domains, the field of multi-representational learning would also greatly benefit from their applications. Lee (2014), for instance, has used the EEG technique as an alternative measurement of cognitive load during a multimedia lesson.

Another alternative is to have a within-subject design, comparing a particular reader's multi-representational reading and learning in both her L1 and L2. This could indicate whether or not the reader demonstrates consistent reading behaviour when processing multi-representational materials in two languages. This approach can be seen in the work of Singer Trakhman, Alexander, and Silverman (2018) who compared undergraduates' L1 reading profiles across two mediums; print and online.

In this study, the same four reading profiles were found across the mediums but those who made up these different profiles differed to some extent. A majority of the undergraduates (i.e. 60%) swapped to a different reading profile across the mediums whilst the remaining 40% demonstrated consistency. It would be interesting to see how variable people are when they read in different languages and whether swapping also takes place across and as a result of languages. By using a counterbalanced within-subject design, we could pinpoint whether differences are attributable to language or other factors.

Future research could also complement eye tracking data with verbal protocols as a form of methodological triangulation. As argued in Chapter 3, although eye tracking data is valuable it can be ambiguous if cognitive processing were to be inferred from them. Verbal protocols can help disambiguate the interpretations as participants themselves can provide accounts into their own cognitive processing by verbalising their thoughts during reading retrospectively. Their eye movement recordings can be replayed to them to stimulate their recall. The researcher has piloted this approach and pilot participants were asked to verbalise their thought processes in English, their L2. Unfortunately, the participants reported this as very difficult as they also needed to remember what they were thinking at the time. If this were to be pursued, the validity of the data could be jeopardised. Future work could therefore ask participants to interact and verbalise their thought processes in their L1, something that was not possible in the present research given the participants spoke a variety of first languages that unfortunately were not part of the researcher's repertoire.

Another opportunity for future work would be to sample from more appropriate age-groups. In the present research, material intended for secondary school students was read by university students, and although science majors were excluded from the sample (i.e. the subject matter was not overlearned), it would be more ideal for future work to use stimuli that are age-appropriate, such as using an undergraduate-level Chemistry text to be read by Chemistry undergraduate participants. Additionally, using a secondary school material meant that the text difficulty level was not at a level appropriate for a university population. Future work could thus manipulate this text characteristic and examine whether reading behaviour would differ as a function of the levels of text difficulty.

Although using authentic materials was an important component of this research, it only examined reading and learning with one particular material. There are other

domains (e.g. History, Geography and other science-related domains) which typically make use of similar materials. Visual representations in these domains may have more variety in terms of structures and functions. Additionally, like much research of "pictures", the pictures in the study and especially the Photo, contained text that might have attracted more and longer fixations towards them and rather than inspecting the visual elements, the readers might have been reading the verbal information instead. Future work could explore this further by systematically manipulating not just the presence of text on pictures but also the structural and functional roles that a particular picture plays in relation to other representations.

Another valuable extension to this work would be to take a more fine-grained approach to the analysis of visual representations as these were treated as a unitary construct, i.e. *pictures*. Despite this straightforward classification, we have seen from the structural and functional analyses of the reading material provided in Chapter 3, how complex an authentic multi-representational material of this nature could be. The visual representations in the reading material differ in their structures and functions and these (the functions, at least), were discussed by the readers themselves. For instance, the picture components of the rubbish mover photograph in the stimulus play a complementary role in relation to the text captions over the photograph. For an L2 reader, however, the picture components may play a constraining role instead. They could help the reader specify, for instance, what a piston is or what squashing actually means in that particular context. Analysing the visual representations in terms of their structural and functional properties could therefore shed more light into how learners are actually using (or not using) them and whether or not these are effective for learning. For example, it may be that integrating text and pictures is more important for representations that play a constraining rather than a complementary role, as in the case of the picture components and text captions mentioned above.

7.7. IMPLICATIONS

This thesis has demonstrated the important role of integrating information across multiple representations when learning from multi-representational materials in the L1 and the L2. To help learners develop this strategic processing, learners require adequate instructional support. Teachers can support learners if they understand on how these materials should be taught and good instructional practices could help their learners engage in this strategic processing. Past research (in L1) has already shown

that the more teachers believed that students should be given explicit strategies when learning with these materials, the more engagement was reported by their students (Schroeder et al., 2011). This relationship, furthermore, was mediated by teachers' perceived instructional practices. Students' self-reported engagement was predicted by how well teachers managed the classroom and how adaptive explanations and supports was provided during lessons that contained these materials. Indeed, the existing L1 literature has recorded the importance of providing learners direct and explicit instructions and strategies to process and learn these materials successfully (Cromley, Perez, et al., 2013; Peeck, 1993; Schwonke et al., 2009). There is a body of research that examines how to implement text-picture integrations more actively in the classroom such as eye movement modelling examples (EMME) that shows learners gaze replays of expert learners when processing a to-be-learned material (Mason, Pluchino, et al., 2015; Mason et al., 2016). Although such practice may not be available in all classrooms, the essence of EMME is to make the learners metacognitively aware of the need to integrate information across representations. The teachers could therefore model the best way to go about doing this so that learners can benefit from the direct and concrete example and proceed to emulating it. Additionally, the teachers can also take advantage of their learners' text-dominant tendency by using text to guide learners' attention to relevant pictures so that they can consequently be integrated with verbal information.

Finally, teachers in a classroom with learners from a variety of L1 backgrounds could note that their L1 and L2 learners do not learn at an identical speed. The L2 learners need considerably more time to learn the same contents and materials that their L1 peers are learning. Despite this, teachers could take note of the positive and reassuring finding that although their L2 learners could be facing an arduous task of learning new contents in the L2 and they may need to work harder at this compared to their L1 peers, they could at the end achieve the same positive outcomes as their peers. Hence, it is helpful to constantly remind and motivate the L2 learners that they can perform as well as their L1 peers when they put in the additional time and effort. As a teacher myself, I know I will endeavour to remember to do so going forward!

7.8. CONTRIBUTIONS AND CONCLUDING REMARKS

This thesis has contributed to the research on learning with multiple representations. It has demonstrated the ways in which reading and processing of multi-

representational materials takes place in the L2 and how this is related to the learning of the materials. The findings build upon what has already been observed and a step forward has been made towards understanding multi-representational learning in the L2. Hopefully, this thesis has made it clear that the L2 community also needs the same amount of attention when it comes to multi-representational learning as the L1 community. Perhaps this is the greater population that is in need of such research given the increasing number of people learning in their non-native language at present. Understandably, reading and learning in the L2 is a harder task compared to reading and learning in the L1.

To a lesser extent, this thesis has also contributed to the learning with authentic materials that combine verbal and multiple visual representations. The structural and functional analyses of the stimulus used in the thesis has demonstrated the complexity of this multi-representational system and how learners themselves perceived and combined representations based on their understanding and goals.

Overall this thesis has set out to demonstrate how reading and learning of an authentic multi-representational material took place in the L2. Similar to multi-representational reading in the L1, multi-representational reading in the L2 was also text-dominant. The ways in which L2 readers went about reading this material could also predict the outcomes of their learning. As reiterated and found in the L1 literature, the more and the longer the L2 readers integrated verbal and visual information, the more they learned from the material. In contrast, the less and the shorter they integrated verbal and visual information, the less they learned from the material. Although the L2 reading patterns appeared similar to L1 reading patterns, the integrative processing made by the L2 readers was more extensive and the L2 readers learned the material as well as their L1 peers. By and large, this thesis has demonstrated that successful learning with multi-representational materials in the L1 and the L2 is closely related to the extent of integrations made across representations. Hence, regardless of which language multi-representational learning takes place in, integration across representations should be a must in order to learn the content of these materials successfully. It appears therefore that the models of multi-representational learning that have been widely observed and promoted in the L1 apply here to the L2 as well and should similarly be encouraged. Hopefully, future work could uncover more insights and further increase our understanding of how this growing population of L2 learners interact and learn with multi-representational materials.

REFERENCES

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2), 131-152.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and instruction, 16*(3), 183-198. doi:http://dx.doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2014). The multiple representation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2d ed., pp. 464-486). New York: Cambridge University Press.

Ainsworth, S. (2018). Multi-modal, multi-source reading: A multi-representational reader's perspective. *Learning and Instruction, 57*, 71-75. doi:10.1016/j.learninstruc.2018.01.014

Ainsworth, S., & Lowe, R. (2012). Representational learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 2832-2835). Boston, MA: Springer US.

American Association for the Advancement of Science. (2016). Retrieved from http://assessment.aaas.org/topics/AM#/

Arndt, J., Schüler, A., & Scheiter, K. (2015). Text–picture integration: How delayed testing moderates recognition of pictorial information in multimedia learning. *Applied Cognitive Psychology, 29*(5), 702-712. doi: 10.1002/acp.3154

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A, 58*(6), 1065-1086. doi: 10.1080/02724980443000476

Bartholomé, T., & Bromme, R. (2009). Coherence formation when learning from text and pictures: What kind of support for whom? *Journal of Educational Psychology, 101*(2), 282-293. doi: 10.1037/a0014312

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography, 6*(4), 253-279. doi: 10.1093/ijl/6.4.253 %J International Journal of Lexicography

BERA. (2011). *Ethical guidelines for educational researchers.* London: British Educational Research Association

Bergen, B. & Chan Lau, T. T. (2012). Writing direction affects how people map space onto time. *Frontiers in Psychology,* 3(109). doi: 10.3389/fpsyg.2012.00109

Bernhardt, E. (1994). A content analysis of reading methods texts: What are we told about the nonnative speaker of english? *Journal of Reading Behavior, 26*(2), 159-189.

Bernhardt, E. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics, 25*, 133-150. doi: doi:10.1017/S0267190505000073

Bisson, M.-J., van Heuven, W., J. B., Conklin, K., & Tunney, R., J. (2013). Incidental acquisition of foreign language vocabulary through brief multi-modal exposure. *PLoS ONE, 8*(4), e60912. doi: 10.1371/journal.pone.0060912

Bisson, M.-J., Van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014a). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics, 35*(2), 399-418. doi: 10.1017/S0142716412000434

Bisson, M.-J., Van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2014b). The role of verbal and pictorial information in multi-modal incidental acquisition of foreign language vocabulary. *The Quarterly Journal of Experimental Psychology, 68*(7), 1-56. doi:10.1080/17470218.2014.979211

Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 38(1), 92-101. doi: 10.1037/a0024828

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge: Cambridge University Press.

Collewijn, H. (1999). Eye movement recording In R. H. S. Carpenter & J. G. Robson (Eds.), *Vision research: A practical guide to laboratory methods* (pp. 245-285). New York: Oxford University Press

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research, 32*(3), 453-467. doi: 10.1177/0267658316637401

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE, 10*(8), e0134008. doi: 10.1371/journal.pone.0134008

Creswell, J. W. (2014a). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Harlow: Pearson.

Creswell, J. W. (2014b). *Research design (international student edition): Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, California: SAGE Publications, Inc.

Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction, 26*, 45-58. doi: 10.1016/j.learninstruc.2013.01.003

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013). Improving students' diagram comprehension with classroom instruction. *The Journal of Experimental Education, 81*(4), 511-537. doi: 10.1080/00220973.2012.745465

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*(1), 59-74. doi: 10.1016/j.cedpsych.2009.10.002

de Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. d. Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9-29). New York: Oxford University Press

Department for Education. School pupils and their characteristics: January 2109. (2019). Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812539/Schools_Pupils_and_their_Characteristics_2019_Main_Text.pdf

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179-201). New York: Oxford University Press

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., & van Heuven, W. J. B. (2018). Visual word recognition in multilinguals. In S. Reuschemeyer & G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics* (Vol. 2, pp. 118-143). Oxford: Oxford University Press

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*(1), 78-103. doi: 10.1598/RRQ.38.1.4

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 663-679. doi: 10.1037/0278-7393.33.4.663

Eitel, A., & Scheiter, K. (2015). Picture or text first? Explaining sequence effects when learning with pictures and text. *Educational Psychology Review, 27*(1), 153-180. doi: 10.1007/s10648-014-9264-4

Ellis, N. C. (2005). Intoroduction to part i: Acquisition. In J. F. Kroll & A. M. B. d. Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 3-29). New York: Oxford University Press

The English Effect. (2013). Retrieved from https://www.britishcouncil.org/sites/default/files/english-effect-report-v2.pdf

EyeTracking, I. (2011). Eyetracking: Usability, visibility, optimization Retrieved 29 December 2015, from http://www.eyetracking.com/About-Us/What-Is-Eye-Tracking

Farley, A. P., Ramonda, K., & Liu, X. (2012). The concreteness effect and the bilingual lexicon: The impact of visual stimuli attachment on meaning recall of abstract l2 words. *Language Teaching Research, 16*(4), 449-466. doi: 10.1177/1362168812436910

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll* (4th ed.). London: SAGE.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233. doi: 10.1037/h0057532

Friedrich, T. E., & Elias, L. J. (2016). The write bias: The influence of native writing direction on aesthetic preference biases. *Psychology of Aesthetics, Creativity, and the Arts, 10*(2), 128-133. doi: 10.1037/aca0000055

Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S., & Boroditsky, L. (2011). How linguistic and cultural forces shape conceptions of time: English and Mandarin time in 3D. *Cognitive Science, 35,* 1305-1328. doi:10.1111/j.1551-6709.2011.01193.x

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, N.J.: Lawrence Erlbaum.

Get your document's readability and level statistics. (2018). Retrieved from https://support.office.com/en-us/article/get-your-document-s-readability-and-level-statistics-85b4969e-e80a-4777-8dd3-f7fc3c8b3fd2?ui=en-US&rs=en-US&ad=US#__toc342546557

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research, 42*(3), 237-288. doi: 10.3102/00346543042003237

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105. doi: 10.1080/09588221.2012.700315

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.

Graddol, D. (1997). The future of english? *A guide to forecasting the popularity of the English language in the 21st century* Retrieved from https://englishagenda.britishcouncil.org/sites/default/files/attachments/books-the-future-of-english.pdf

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. (7th ed.). New Jersey: Pearson Prentice Hall.

Hall, G. (2018). *The experiences of secondary school students with English as an additional language: Perceptions, priorities and pedagogy*. Retrieved from https://www.teachingenglish.org.uk/sites/teacheng/files/J154%20ELTRA_Secondary%20Students%20Eng%20as%202nd%20Lang%20Paper_A4_FINAL_web.pdf

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*(2), 95-123. doi: http://dx.doi.org/10.1006/ceps.1998.0987

Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 447-465). New York: Cambridge University Press

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*(6), 717-742. doi: http://dx.doi.org/10.1006/jmla.1993.1036

Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2013). Dynamics of mental model construction from text and graphics. *European Journal of Psychology of Education, 28*(4), 1105-1126. doi: 10.1007/s10212-012-0156-z

Höffler, T. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. *Educational Psychology Review, 22*, 245-269. doi:10.1007/s10648-010-9126-7

Höffler, T., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations – evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior, 27*(1), 209-216. doi: 10.1016/j.chb.2010.07.042

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2009). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology, 23*(9), 1215-1226. doi: 10.1002/acp.1525

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*(2), 172-176. doi: http://dx.doi.org/10.1016/j.learninstruc.2009.02.013

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Jared, D. (2015). Literacy and literacy development in bilinguals. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 165-182). New York: Oxford University Press.

Jian, Y.-C., & Wu, C.-J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text-diagram integration. *Journal of Science Education and Technology, 24*(1), 43-55. doi: 10.1007/s10956-014-9519-3

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied, 18*(2), 178.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=ovfta&AN=00006832-198007000-00001

Kang, H. (2014). Understanding online reading through the eyes of first and second language readers: An exploratory study. *Computers & Education, 73*, 1-8. doi: http://dx.doi.org/10.1016/j.compedu.2013.12.005

Kantowitz, B. H., Roediger, H. L. I., & Elmes, D. G. (2009). *Experimental psychology* (9th ed.). Belmont: Wadsworth, Cengage Learning.

Kazandjian, S., Cavézian, C., Zivotofsky, A. Z., & Chokron, S. (2010). Bisections in two languages: When number processing, spatial representation, and habitual

reading direction interact. *Neuropsychologia, 48*(14), 4031-4037. doi: http://dx.doi.org/10.1016/j.neuropsychologia.2010.10.020
Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.0521629861
Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. doi: 10.1037/0033-295X.85.5.363
Koda, K. (2004). *Insights into Second Language Reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.
Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning, 57*, 1-44. doi:10.1111/0023-8333.101997010-i1
Kozan, K., Erçetin, G., & Richardson, J. C. (2015). Input modality and working memory: Effects on second language text comprehension in a multimedia learning environment. *System, 55*, 63-73. doi: http://dx.doi.org/10.1016/j.system.2015.09.001
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the n400 component of the event-related brain potential (erp). *Annual Review of Psychology, 62*(1), 621-647. doi: 10.1146/annurev.psych.093008.131123
Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., . . . Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90-115. doi: http://dx.doi.org/10.1016/j.edurev.2013.10.001
Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65-100. doi: 10.1111/j.1551-6708.1987.tb00863.x
Lee, H. (2014). Measuring cognitive load with electroencephalography and self-report: Focus on the effect of english-medium learning for korean students. *Educational Psychology, 34*(7), 838-848. doi: 10.1080/01443410.2013.860217
Lee, H., & Mayer, R. E. (2015). Visual aids to learning in a second language: Adding redundant video to an audio lecture. *Applied Cognitive Psychology, 29*(3), 445-454. doi: 10.1002/acp.3123
Lee, H., & Mayer, R. E. (2018). Fostering learning from instructional video in a second language. *Applied Cognitive Psychology, 32*(5), 648-654. doi: 10.1002/acp.3436
Lee, V. R. (2010). Adaptations and continuities in the use and design of visual representations in us middle school science textbooks. *International Journal of Science Education, 32*(8), 1099-1126. doi: 10.1080/09500690903253916
Lemhöfer, K., & Broersma, M. (2012). Introducing lextale: A quick and valid lexical test for advanced learners of english. *Behavior Research Methods, 44*(2), 325-343. doi: 10.3758/s13428-011-0146-0
Levesley, M., Johnson, P., & Gray, S. (2008). *Exploring science: How science works Year 7*. Essex: Pearson Education
Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows, & Houghton, H. A. (Eds.), *The psychology of illustration: I. Basic research* (pp. 51-85). New York: Springer
Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in chinese reading. *Journal of Experimental Psychology: General, 143*(2), 895.
Lin, J. J. H., Lee, Y.-H., Wang, D.-Y., & Lin, S. S. J. (2016). Reading subtitles and taking enotes while learning scientific materials in a multimedia environment: Cognitive load perspectives on EFL students. *Educational Technology & Society, 19*(4), 47-58.

Liu, Y., Jang, B. G., & Roy-Campbell, Z. (2018). Optimum input mode in the modality and redundancy principles for university ESL students' multimedia learning. *Computers & Education, 127*, 190-200. doi: 10.1016/j.compedu.2018.08.025

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6-14. doi: 10.1016/S1364-6613(99)01418-7

Lohse, G. L., Biolsi, K., Walker, N., & Rueter, H. H. (1994). A classification of visual representations. *Communications of the ACM, 37*(12), 36-49.

Lotto, L., & De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*(1), 31-69. doi: 10.1111/1467-9922.00032

Macdonald, M. C., & Hsiao, Y. (2018). Sentence comprehension. In A. E. Hernandez, E. M. Fernandez & N. Aznar-Bese (Eds.), *The oxford handbook of psycholinguistics* (Vol. 2), pp. 171- 196. Oxford: Oxford University Press

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology, 41*, 172-187. doi: https://doi.org/10.1016/j.cedpsych.2015.01.004

Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *47*(6), 1083-1095. doi: 10.1111/bjet.12271

Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*(3), 356-384. doi: 10.1080/00220973.2012.727885

Mason, L., Scheiter, K., & Tornatora, M. C. (2017). Using eye movements to model the sequence of text–picture processing for multimedia comprehension. *Journal of Computer Assisted Learning, 33*(5), 443-460. doi:10.1111/jcal.12191

Mason, L., Tornatora, M., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing, 28*(6), 851-872. doi: 10.1007/s11145-015-9552-5

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*(1), 95-109. doi: http://dx.doi.org/10.1016/j.compedu.2012.07.011

Mayer, R. E. (2005a). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2005b). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31-48). New York: Cambridge University Press.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology, 28*(5), 653-660. doi: 10.1002/acp.3050

McNamara, D. S., & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). New York: Elsevier

McNeil, L. (2011). Investigating the contributions of background knowledge and reading comprehension strategies to l2 reading comprehension: An exploratory study. *Reading and Writing, 24*(8), 883-902. doi: 10.1007/s11145-010-9230-6

McNeil, L. (2012). Extending the compensatory model of second language reading. *System, 40*(1), 64-76. doi: http://dx.doi.org/10.1016/j.system.2012.01.011
Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for l2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720-739. doi: 10.1016/j.system.2013.07.013
Myers, A., & Hansen, C. (2011). *Experimental psychology (International edition)* (7 ed.). Belmont: Wadsworth Cengage Learning.
Nation, I. S. P. (2001). *Learning vocabulary in another language.* New York: Cambridge University Press.
Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59-82.
Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory, 3*(5), 485-497. doi: 10.1037/0278-7393.3.5.485
O'Brien, E. J., & Cook, A. E. (2015). Models of discourse comprehension. In A. Pollatsek & R. Treiman (Eds.), *The oxford handbook of reading* (pp. 217-231). New York: Oxford University Press
O'Keefe, P. A., Letourneau, S. M., Homer, B. D., Schwartz, R. N., & Plass, J. L. (2014). Learning from multiple representations: An examination of fixation patterns in a science simulation. *Computers in Human Behavior, 35*, 234-242. doi: http://dx.doi.org/10.1016/j.chb.2014.02.040
Oller, J. W., & Tullius, J. R. (1973). Reading skills of non-native speakers of english. *IRAL: International Review of Applied Linguistics in Language Teaching, 11*(1), 69.
Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*(1), 110-117. doi: https://doi.org/10.1016/j.chb.2009.09.001
Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction, 19*(3), 228-242. doi: http://dx.doi.org/10.1016/j.learninstruc.2008.04.003
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie, 45*(3), 255-287. doi: 10.1037/h0084295
Palmer, S. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259-303). Hillsdale, NJ: LEA.
Park, E., Cho, M., & Ki, C. S. (2009). Correct use of repeated measures analysis of variance. *The Korean Journal of Laboratory Medicine, 29*(1), 1-9. doi: 10.3343/kjlm.2009.29.1.1
Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction, 3*(3), 227-238. doi: 10.1016/0959-4752(93)90006-L
Pellicer-Sanchez, A., Tragant, E., Conklin, K., Rodgers, M., Llanes, A., & Serrano, R. (2018). L2 reading and reading-while-listening in multimodal learning conditions: An eye-tracking study. *ELT Research Papers 18.01*.
Pérez, A., Hansen, L., & Bajo, T. (2018). The nature of first and second language processing: The role of cognitive control and l2 proficiency during text-level comprehension. *Bilingualism*, 1-19. doi: 10.1017/S1366728918000846
Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Maryland: Rowman & Littlefield Publishers, Inc.
Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25-36

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior, 19*(2), 221-243. doi: https://doi.org/10.1016/S0747-5632(02)00015-8

Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 467-488). New York: Cambridge University Press

Rasch, D., & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science, 46*(2), 175-208

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing stem learning. *Educational Psychology Review, 29*(4), 717-761. doi: 10.1007/s10648-016-9365-3

Rau, M. A., Michaelis, J. E., & Fay, N. (2015). Connection making between multiple graphical representations: A multi-methods approach for domain-specific grounding of an intelligent tutoring system for chemistry. *Computers & Education, 82*, 460-485. doi: 10.1016/j.compedu.2014.12.009

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422. doi: 10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. doi: 10.1080/17470210902816461

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*(3), 241-255. doi: 10.1207/s1532799xssr1003_3

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*(3), 219-226. doi: 10.1037/1076-898X.7.3.219

Renkl, A., & Scheiter, K. (2015). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*, 1-23. doi: 10.1007/s10648-015-9340-4

Robson, C. (2011). *Real world research* (3rd ed.). Sussex: John Wiley & Sons Ltd.

Robson, C., & McCartan, K. (2016). *Real World Research: A Resource for Users of Social Research Methods in Applied Settings* (4th ed.). Chichester: Wiley.

Rueda, R. (2011). Cultural perspectives in reading: Theory and research. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4), pp. 84-103. New York: Routledge.

Russell, D. (2016, 2 May 2016). Nottinghamshire school with 39 different languages rated 'good' by ofsted, *Nottingham Post*. Retrieved from http://www.nottinghampost.com/Nottinghamshire-school-39-different-languages/story-29204130-detail/story.html

Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction, 36*(Supplement C), 11-26. doi: https://doi.org/10.1016/j.learninstruc.2014.11.002

Scheiter, K., & Eitel, A. (2017). The use of eye tracking as a research and instructional tool in multimedia learning. In C. A. Was, F. J. Sansosti & B. J. Morris (Eds.), *Eye-tracking technology applications in educational research* (pp. 143-164). Hershey, PA: IGI Global

Scheiter, K., Schüler, A., & Eitel, A. (2017). Learning from multimedia: Cognitive processes and instructional support. In S. S. & C. U. (Eds.), *The psychology of digital learning* (pp. 1-19). Cham: Springer

Scheiter, K., Schüler, A., Gerjets, P., Huk, T., & Hesse, F. W. (2014). Extending multimedia research: How do prerequisite knowledge and reading

comprehension affect learning from text and pictures. *Computers in Human Behavior, 31*, 73-84. doi: http://dx.doi.org/10.1016/j.chb.2013.09.022

Scheiter, K., & van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Applied Cognitive Psychology, 23*(9), 1209-1214. doi: 10.1002/acp.1524

Scheiter, K., Wiebe, E., & Holsanova, J. (2009). Theoretical and instructional aspects of learning with visualisations. In R. Z. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 67-88). Hershey, PA: Information Science Reference. doi:http://dx.doi.org/10.1016/j.learninstruc.2009.02.011

Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system- and self-paced instruction in multimedia learning. *Learning and Instruction, 20*(2), 100-110.

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363. doi: 10.1177/1362168808089921

Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101-120. doi: 10.1023/A:1013136727916

Schnotz, W., Wagner, I., Ullrich, M., Horz, H., & McElvany, N. (2017). Development of students' text-picture integration and reading competence across grades 5–7 in a three-tier secondary school system: A longitudinal study. *Contemporary Educational Psychology, 51*(Supplement C), 152-169. doi: https://doi.org/10.1016/j.cedpsych.2017.06.003

Schroeder, S., Richter, T., McElvany, N., Hachfeld, A., Baumert, J., Schnotz, W., . . . Ullrich, M. (2011). Teachers' beliefs, instructional behaviors, and students' engagement in learning from texts with instructional pictures. *Learning and Instruction, 21*(3), 403-415. doi: https://doi.org/10.1016/j.learninstruc.2010.06.001

Schüler, A. (2017). Investigating gaze behavior during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction, 49*, 218-231. doi: https://doi.org/10.1016/j.learninstruc.2017.03.001

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology, 23*(9), 1227-1243. doi: 10.1002/acp.1526

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. d. Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). New York: Oxford University Press.

Şendurur, E., Doğusoy, B., & Yondemir Çalişkan, N. (2018). Investigation of non-native learners' informal learning processes from cognitive-load theory perspective. *Interactive Learning Environments*, 1-12. doi: 10.1080/10494820.2018.1517096

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*(2), 227-237. doi: https://doi.org/10.1016/S0959-4752(02)00022-1

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs: For generalized causal inference*. Belmont: Wadsworth

Sinatra, G. M., & Broughton, S. H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly, 46*(4), 374–393.

Singer Trakhman, L. M., Alexander, P. A., & Silverman, A. B. (2018). Profiling reading in print and digital mediums. *Learning and Instruction, 57*, 5-17. doi: 10.1016/j.learninstruc.2018.04.001

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. *Reading Psychology, 31*(3), 301-325. doi: 10.1080/02702710903256502

Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: Using if-then plans to facilitate cognitive processing. *Learning and Instruction, 35*, 1-15. doi: 10.1016/j.learninstruc.2014.09.002

Staub, A. (2015). Reading sentences: Syntactic parsing and semantic interpretation In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 202-216). New York: Oxford University Press.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction, 29*(1), 123-145. doi: 10.1080/07370008.2010.507318

Tokowicz, N., & Perfetti, C. A. (2005). Introduction to part ii: Comprehension. In J. F. Kroll & A. M. B. d. Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 173-177). New York: Oxford University Press

Tolentino, L. C., & Tokowicz, N. (2011). Across languages, space, and time: A review of the role of cross-language similarity in l2 (morpho)syntactic processing as revealed by fmri and erp methods. *Studies in Second Language Acquisition, 33*(1), 91-125. doi: 10.1017/S0272263110000549

Tragant Mestres, E., & Pellicer-Sánchez, A. (2019). Young EFL learners' processing of multimodal input: Examining learners' eye movements. *System, 80*, 212-223. doi: 10.1016/j.system.2018.12.002

van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*(1), 88-107. doi: 10.1016/j.jml.2010.08.006

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction, 16*(3), 199-212. doi: 10.1016/j.learninstruc.2006.03.007

Van Der Zee, T., Admiraal, W., Paas, F., Saab, N., & Giesbers, B. (2017). Effects of subtitles, complexity, and language proficiency on learning from online education videos. *Journal of Media Psychology: Theories, Methods, and Applications, 29*(1), 18-30. doi: 10.1027/1864-1105/a000208

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press

van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*(2), 95-99. doi: 10.1016/j.learninstruc.2009.02.009

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458-483. doi: 10.1006/jmla.1998.2584

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English.

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition, 137*, 40-46. doi: https://doi.org/10.1016/j.cognition.2014.12.008

Verhallen, M. J. A. J., & Bus, A. G. (2011). Young second language learners' visual attention to illustrations in storybooks. *Journal of Early Childhood Literacy, 11*(4), 480-500. doi: 10.1177/1468798411416785

Verhoeven, L. (2011). Second language reading acquisition. In M. L. Kamil, P. D. Pearson, E. B. Moje & P. P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4) pp. 661-683. New York: Routledge

Warren, P., Boers, F., Grimshaw, G., & Siyanova-Chanturia, A. (2018). The effect of gloss type on learners' intake of new words during reading. *Studies in Second Language Acquisition, 40*(4), 883-906. doi: 10.1017/S0272263118000177

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. [journal article]. *Psychonomic Bulletin & Review, 19*(1), 73-80. doi: 10.3758/s13423-011-0179-5

Whitford, V., & Titone, D. (2015). Second-language experience modulates eye movements during first- and second-language sentence reading: Evidence from a gaze-contingent moving window paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 1118-1129. doi: 10.1037/xlm0000093

Wiederholt, J. L., & Blalock, G. (2000a). *Gray silent reading tests: Examiner's manual*. Texas: Pro-Ed

Wiederholt, J. L., & Blalock, G. (2000b). *Gray silent reading tests: Gsrt*. Texas: Pro-Ed.

Won, M., Yoon, H., & Treagust, D. F. (2014). Students' learning strategies with multiple representations: Explanations of the human breathing mechanism. *Science Education, 98*(5), 840-866. doi: 10.1002/sce.21128

Wynder, M. (2018). Visualising accounting concepts: Insights from cognitive load theory for English as a second language students. *Accounting Education, 27*(6), 590-612. doi: 10.1080/09639284.2017.1361847

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning, 24*(1), 39-58. doi: 10.1080/09588221.2010.523285

Yusof, S. M., Lazim, Z. M., Salehuddin, K., & Shahimin, M. M. (2019). Graphic novels: Understanding how fifth graders read literary text through eye movement analysis. *Kritika Kultura, 33/34*, 388-427. doi: http://dx.doi.org/10.13185/3076

APPENDICES

Appendix 1: Research Stimulus

7Gc Bits and pieces

What is particle theory?

HowScienceWorks

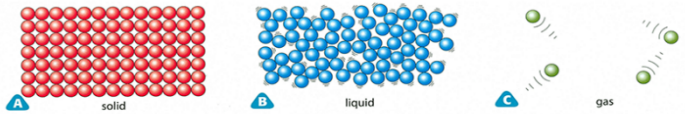
There is a theory that can explain the different properties of solids, liquids and gases. This is called the **particle theory** or particle model. It says that all things are made of tiny pieces called **particles** that are constantly moving. Solids, liquids and gases have different arrangements of particles, which give them their special properties.

1 What are all materials made from?

Solids have particles that are very close together. The particles can vibrate (jiggle around) but are fixed in their places so they can't move past each other. That's why solids keep their shapes.

The particles in a **liquid** are close to one another but can move past each other. That's why liquids flow.

In **gases**, the particles are far apart and can move anywhere by themselves. That's why gases spread all over the place and fill any container they are in.



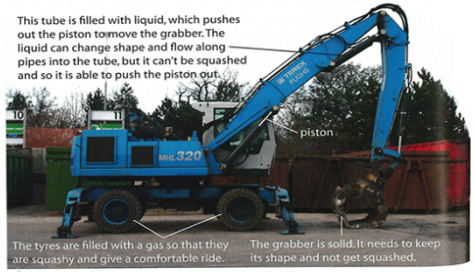
2 What decides whether something is a solid, a liquid or a gas?

Explaining the properties

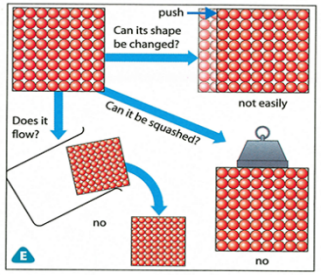
This machine is used to move rubbish around at the recycling centre.

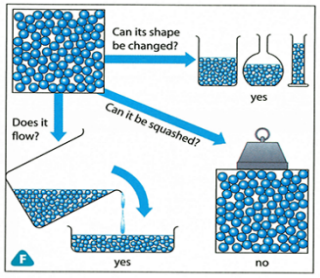
Things can only be squashed if the particles get closer together. In a solid, the particles are already very close together. This makes it very difficult for the volume of a solid to be made smaller.

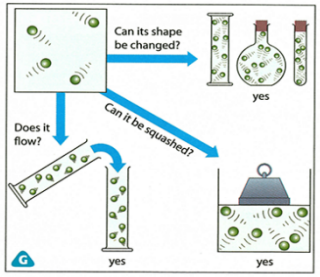
3 Explain why the grabber of the rubbish-mover is made of a solid.



D *A recycling centre rubbish mover. This machine is used to move and squash rubbish at a recycling centre.*







Solids have a fixed shape because the particles are held tightly together by forces, called **bonds**. These bonds are very strong and stop the particles from moving around. The particles are fixed in place and can only vibrate. This stops the solid flowing.

4 What holds the particles of a solid together?

Liquids cannot be squashed because the particles in a liquid are already close together. However, the particles can move past each other quite easily. The bonds in a liquid are weak enough to let the particles move past each other but strong enough to hold them close together.

5 Explain why a liquid can move the shovel of the rubbish-mover.

Particles in a gas are very far apart and move very quickly in all directions. The particles are able to move all over the place because there are no bonds between them. Therefore, a gas does not have a fixed volume or shape.

When we squash a gas, the particles are moved closer together. The more the gas is squashed, the closer together the particles will get.

6 Why would a gas not be suitable to fill the tube that moves the piston in the rubbish-mover?

7 Why can gas particles spread out to fill any space?

8 Describe how the particles move in solids, liquids and gases.

I CAN...

- recall how the particle theory says the particles in solids, liquids and gases are arranged. H S W
- describe how materials have different properties depending on how their particles are arranged. H S W

Appendix 2: Shortlisted Research Stimulus 1

7Ge Pressure problems

What causes pressure in gases?

The methane gas produced in landfill sites can be dangerous. Gases flow very easily and so the methane can flow all through a landfill site and collect in pockets. Sparks can cause the gas to explode or the build-up of gas in a pocket can cause the ground above it to burst or crack. So, it is important to get the methane out of landfill sites.



A These cracks were caused by gas collecting inside the landfill.

1 Give two reasons why it is important to get rid of the gases inside a landfill site. H S W

When a gas builds up in a closed-in space it creates a 'pushing' known as **pressure**. Pressure is caused by particles. The moving particles bump into each other and hit the sides of their container. The forces of the particles hitting the sides cause pressure. If more particles hit the sides, there is more pressure.

Think about a balloon. As it is blown up, more and more gas particles are put inside. The pressure gets bigger because there are more particles hitting the sides of the balloon. Eventually the pressure becomes so great that the balloon bursts.

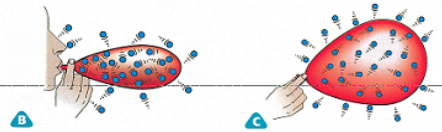
2 What causes pressure in a container of gas? H S W

3 Explain how the cracks in photo A were formed. H S W

Air pressure

The air particles around us put pressure on us. This is called **air pressure**. We can't feel this pressure because we are so used to it. However, we can see its effects.

When you blow up a balloon, more particles are blown into it. This increases the pressure. This makes the rubber stretch and the balloon gets bigger. When the pressure inside it is the same as the air pressure outside it, the balloon stops expanding.



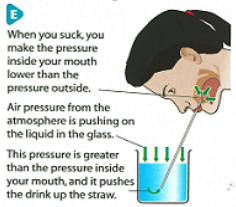
B C

4 Why does the air put pressure on us? H S W

Photo D shows a geyser that is only active in the mornings. At night, the air is very cold. In cold air the particles are close together and create more pressure. This pressure stops gases escaping from the geyser. In the morning, the air warms up and the air pressure is reduced. The gases suddenly escape, pushing boiling water up out of the ground too. The geyser lasts for about half an hour before all the boiling water has been used up.

If it were not for air pressure you could not suck things through a straw.

E



5 Explain how air pressure helps you to drink through a straw. H S W



D The geysers at El Tatio, Chile, are the highest in the world.

If you suck all the air out of something, you get a **vacuum**. A vacuum is a space where there are no particles at all, not even air particles.

Photo F shows two Magdeburg hemispheres. If the air is sucked out of these Magdeburg hemispheres, they are very difficult to pull apart. The pressure of the air on the outside of the hemispheres pushes them together.



F

H S W

What do you think will happen if all the air is sucked out of this can?

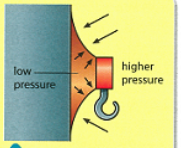
- Explain your prediction using ideas about particles.



G

6 Why will the hook in drawing H stay fixed to the wall? H S W

7 Explain why Magdeburg hemispheres stay together when there is a vacuum inside. H S W



H

I CAN...

- explain how gases cause pressure. H S W
- recall some effects of gas pressure.

Appendix 3: Shortlisted Research Stimulus 2

7Jb Using electricity

How do we use electricity?

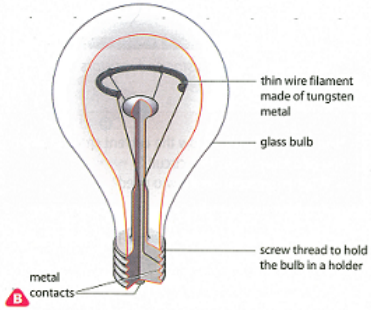
When a current flows through a wire, the wire may get hot. The heat produced by electricity can be useful.

Electricity is used to heat the blade of the knife in photo A, so it can cut through materials such as plastic or rubber.

If a piece of metal gets hot enough it gives out light as well as heat. Light bulbs have a piece of wire, called the filament, inside the glass. The filament has a high resistance. When a current flows through the filament, it gets so hot that it glows.



A

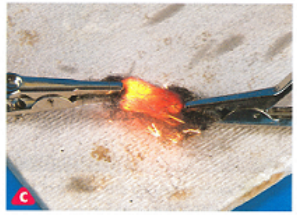


B

How could you find out whether a long wire or a short wire has the highest resistance?
- What equipment would you need?
- How would you stay safe?

Thomas Edison (1847–1931) is often given credit for inventing the light bulb. However, several other people had worked on developing light bulbs, and patented their ideas. A patent gives the person the right to make money out of their invention. Joseph Swan (1828–1914) had an earlier patent on light bulbs. Eventually the two men formed a joint company to sell the bulbs.

1 What happens to the temperature of a wire when a current flows through it?
2 Write a list of all the appliances you can think of that heat up when electricity flows through them.


C

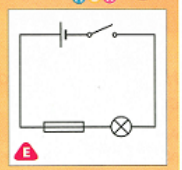
If a piece of metal gets too hot, it may melt. The steel wool in photo C conducts electricity. If the current is increased, the steel wool glows. If the current is increased even more, the steel wool will melt.

A **fuse** is a piece of wire that is designed to melt if the current gets too big.



D *A fuse and the symbol for a fuse.*

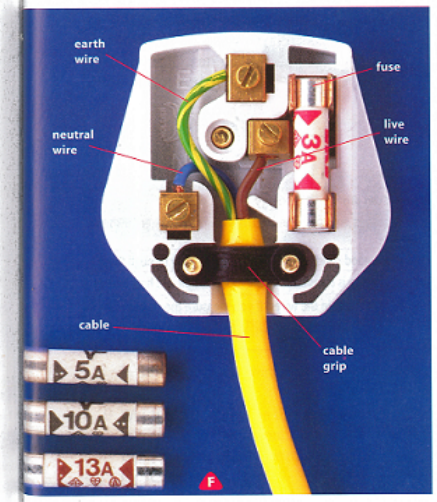
How would you find out how much current can flow through a piece of fuse wire before it melts?
- What apparatus would you need?
- What circuit would you use?


E

For most pieces of electrical equipment, the fuse is in the plug. If an appliance, like an iron, goes wrong a lot of electricity may flow through the wires. This can make the wires in the walls of the house very hot and make them catch fire. The plug has a fuse in it to stop this happening. The high current makes the fuse melt, so no more current can flow. It is important to have the correct fuse fitted in a plug. All appliances have a label on them to show what fuse should be used.

Equipment	Fuse
kettle	13 A
iron	13 A
TV	5 A
music centre	5 A
DVD player	3 A
central heating pump	3 A

G *Fuses used in some common appliances.*



F

3 How does a fuse stop the current flowing if the current is too big?
4 How can a fuse protect electrical equipment?
5 Which pieces of equipment in table G use the most current? Explain how you worked out your answer.
6 What do you think would happen if you put a:
a 3 A fuse in a plug used for a kettle
b 13 A fuse in a plug used for a TV?

I CAN...
- recall that some wires get hot when a current flows through them.
- explain what a fuse does.
- choose the correct fuse for an appliance.

Appendix 4: Prior Knowledge Task

Circle the correct response for each question.

1. In which state of matter are the particles spaced farthest apart?

   A. *A gas*

   B. A liquid

   C. A solid

   D. All are equal.

2. Why is ice harder than liquid water?

   A. The particles of ice are not moving.

   B. *The particles of ice are linked more tightly together.*

   C. The particles of ice are harder than the particles of liquid water.

   D. The particles of ice are made of solid atoms, and the particles of liquid water are made of liquid atoms.

3. In a cup of liquid water, when would the water particles stop moving?

   A. The particles would stop moving if the liquid water in the cup became a solid.

   B. The particles would stop moving if the liquid water in the cup became a gas.

   C. The particles would stop moving if the liquid water in the cup became still.

   D. *The particles would not stop moving in the cup of liquid water.*

4. In which state of matter is the connection between the particles the strongest?

   A. A gas

   B. A liquid

   C. *A solid*

   D. All are equal.

5. Which statement describes the particles of a gas?

   A. The particles are soft.

   B. The particles do not move.

   C. *The particles are far apart from one another.*

   D. The particles are often in contact with one another.

6. Why does liquid water take the shape of a cup it is poured into, but solid ice cubes do not?

   A. Because the particles of liquid water are softer than the particles of solid ice

   B. Because the particles of liquid water are smaller than the particles of solid ice

   C. Because the particles of liquid water are moving but the particles of solid ice are not

   D. *Because the particles of liquid water can easily move past one another but the particles of solid ice cannot*

7. The windows of your school are made of glass. Which of the following statements describes the motion of the particles that make up the glass?

   A. The particles of the glass are never moving.

   B. *The particles of the glass are always moving.*

   C. The particles of the glass move only when the sun warms the window.

   D. The particles of the glass move only when the window is being opened or closed.

8. Why can gases be squashed more easily than solids?

   A. Because the particles of gases are softer than the particles of solids

   B. Because the particles of gases weigh less than the particles of solids

   C. Because the particles of gases move faster than the particles of solids

   D. *Because the particles of gases are farther apart than the particles of solids*

9. Which statement describes the location of the particles of a gas in a sealed container?

A. The particles are packed closely throughout the container.

B. *The particles are spread far apart throughout the container.*

C. Almost all of the particles are at the top of the container.

D. Almost all of the particles are at the bottom of the container.

10. Why does liquid candle wax flow but solid candle wax does not?

A. Because the particles of liquid candle wax are softer than the particles of solid candle wax

B. Because the particles of liquid candle wax weigh less than the particles of solid candle wax

C. Because the particles of liquid candle wax are moving but the particles of solid candle wax are not

D. *Because the particles of liquid candle wax can easily move past one another but the particles of solid candle wax cannot*

Appendix 5: Learning Outcome Task

Answer the following questions.

1. What are all materials made from?

2. What decides whether something is a solid, a liquid or a gas?

3. Explain why the grabber of the rubbish-mover is made of a solid.

4. What holds the particles of a solid together?

5. Explain why a liquid can move the shovel of the rubbish-mover.

6. Why would a gas not be suitable to fill the tube that moves the piston in the rubbish-mover?

7. Why can gas particles spread out to fill any space?

8. Describe how the particles move in solids, liquids and gases. You can use drawings and words to describe the movements.

Appendix 6: Scoring Rubrics for Learning Outcome Task

Q	Question	Answer	Full Points
1	What are all materials made from?	Particles [1pt]	1
2	What decides whether something is a solid, a liquid or a gas?	Arrangement of their particles [1pt] *Acceptable answer* - How close their particles are [1pt]	1
3	Explain why the grabber of the rubbish-mover is made of a solid.	Solid can keep its shape and can't get squashed. [2pt] **Breakdown of points** (i) Solid can keep its shape. [1pt] *Acceptable answer for (i)* - Solid has a fixed shape. [1pt] - Solid is not malleable. [1pt] (ii) Solid can't get squashed/ compressed. [1pt] *Acceptable answer for (ii)* - Solid is firm/ strong. [1pt]	2
4	What holds the particles of a solid together?	Strong bonds **Breakdown of points** (i) Strong [1pt] (ii) bonds [1pt] *Acceptable answer for (ii)* - connections [1pt]	2

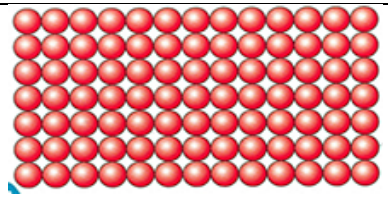
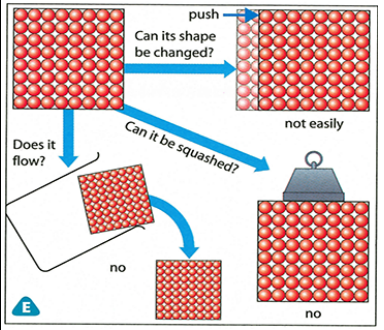
5	Explain why a liquid can move the shovel of the rubbish-mover.	A liquid can change its shape, flow (into the tube of the rubbish-mover), but it can't get squashed, so it can (push out the piston that) move the shovel. [4pt] **Breakdown of points** (i) A liquid can change its shape [1pt] (ii) flow (into the tube of the rubbish-mover) [1pt] *Acceptable answer for (i) and (ii) combined* - Liquid particles can move past each other [1pt] *Acceptable answer for (ii)* - Liquid can move along the tube - Liquid can move easily from one container to another [1pt] (iii) but it can't get squashed [1pt] (iv) so it can (push out the piston that) move the shovel. [1pt]	4
6	Why would a gas not be suitable to fill the tube that moves the piston in the rubbish-mover?	Gases do not have bonds between particles, and so can be squashed rather than fill the tube (that moves the piston). [3pt] **Breakdown of points** (i) Gases do not have bonds between particles [1pt] *Acceptable answer for (i)* - Gas particles are far apart [1pt] - Gas particles move freely [1pt] (ii) and so can be squashed [1pt] (iii) rather than fill the tube (that moves the piston) [1pt]	3
7	Why can gas particles spread out to fill any space?	They have no bonds. [1pt]	1
			TOTAL POINTS (Questions 1 to 7) = 14

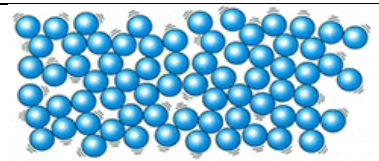
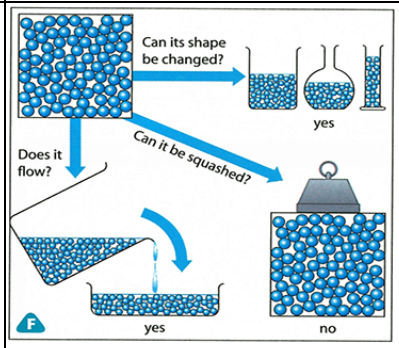
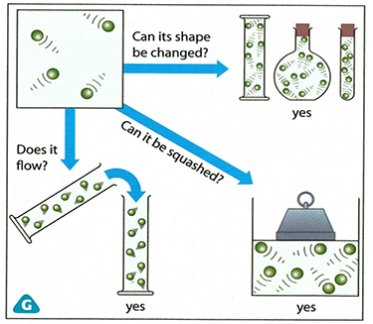
Q	Question	Answer	Full Points
8	Describe how the particles move in solids, liquids and gases. You can use drawings and words to describe the movements.	**<u>TEXT</u>** **SOLIDS** Solid particles are fixed in their places and only vibrate/ jiggle around. [2pt] **Breakdown of points** (i) Solid particles are fixed in their places [1pt] *Acceptable answer for (i)* - Solid particles can't move past each other. [1pt] (ii) Solid particles only vibrate/ jiggle around [1pt]	2
		LIQUIDS Liquid particles move past each other, easily [2pt] **Breakdown of points** (i) Liquid particles move past each other [1pt] (ii) easily [1pt]	2

		GASES Gas particles move anywhere, quickly, and can be moved closer [3pt] **Breakdown of points** (i) Gas particles move anywhere [1pt] *Acceptable answer for (i)* - Gas particles spread all over the place [1pt] (ii) quickly [1pt] (iii) and can be moved closer [1pt]	3
			TOTAL POINTS (text descriptions for Question 8) = 7

Answers to Question 8 can be marked in 2 ways: i) Text and drawing for each matter are marked as separate representations, each containing different information. As such, see Coding (I)

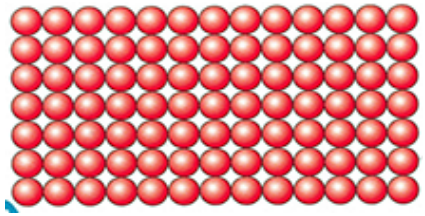
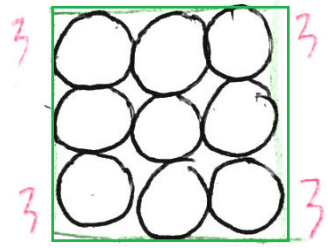
Coding (I)

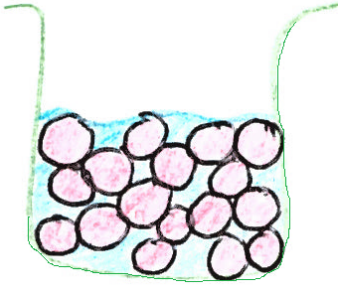
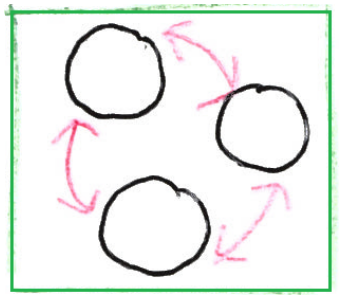
Q8	Drawing Section				
	Drawing	**(A)** **How points are given**	**Point**	**Answer options**	**Note**
Solids		-all particles are in contact [1pt] -particles arranged neatly [1pt] -very little space in between particles [1pt]	3		Points are given even if only 1 matter's property (shape, flow or squash) is drawn. The particles drawn to illustrate the property need to adhere to (A).

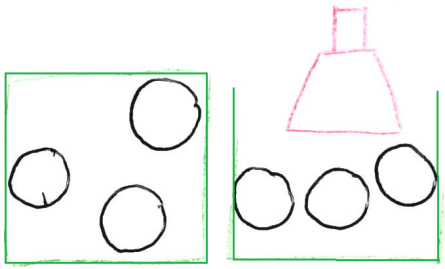
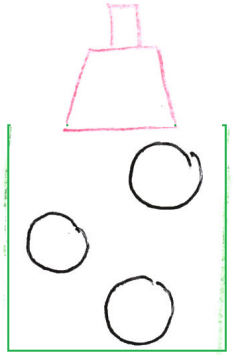
	Drawing	(A) How points are given	Point	Answer options	Note
Liquids		-most particles are in contact [1pt] -space between particles, both small and wide) [1pt] -lines/ arrows next to particles indicating movement [1pt]	3		Points are given even if only 1 property (shape, flow or squash) is drawn. The particles drawn to illustrate the property need to adhere to (A), **but lines indicating movement are optional** (since movement is shown through the matter's property)
Gases		-particles are far apart [1pt] -particles are not in contact [1pt] -lines next to particles indicating movement [1pt]	3		Points are given even if only 1 property (shape, flow or squash) is drawn. The particles drawn to illustrate the property need to adhere to (A), **but lines indicating movement are optional** (since movement is shown through the matter's property)
TOTAL POINTS (all drawings for Question 8) = 9					

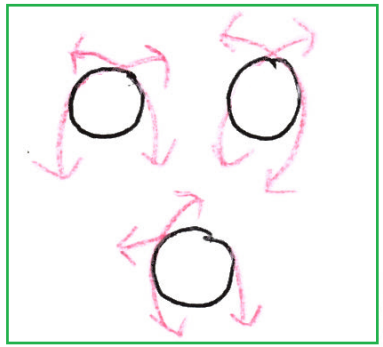
ii) If participants do not get full text and drawing points for each matter, use Coding (II), whereby text and drawing for each matter are marked as complementary representations; what one representation can infer about the other.

Coding (II)

Matter	Sample answer	Note	POINTS	
			Text	Drawing
Solids	Solids have a fixed shape, cannot flow nor can be squashed. 	Drawing indicates that solid particles are 'fixed in their places'. Therefore, even though the participant did not state this feature in their text, the text point should be given anyway.	**1** - for 'fixed in their places' [1pt]	**3** -all particles are in contact [1pt] -particles arranged neatly [1pt] -very little space in between particles [1pt]
	Solids have a fixed shape, cannot flow nor can be squashed. 	Drawing indicates that solid particles are: - 'fixed in their places', and - 'vibrate' as indicated by the number-3-like line outside the green box. Therefore, even though the participant did not state these features in their text, the text points should be given anyway.	**2** - for 'fixed in their places' [1pt] - for 'vibrate' [1pt]	**3** -all particles are in contact [1pt] -particles arranged neatly [1pt] -very little space in between particles [1pt]

Matter	Sample answer	Note	POINTS	
			Text	Drawing
Liquids	Liquids can <u>flow</u> but cannot be squashed. 	Drawing does not contain lines that signify movement of the particles. In such cases, check participants' textual description for any indication that the particles could move, such as 'flow' or 'take shape of container' or 'move past each other'. Since sample answer contains the word 'flow', this can be used as a substitute for lines that indicate movement, therefore the drawing point should be given anyway.	**0**	**3** -most particles are in contact [1pt] -small and wide spaces between particles [1pt] *Plus* - for 'flow' as a substitute for lines signifying movement [1pt]
	Liquids can <u>flow</u> but cannot be squashed. 	Arrows indicate that the particles: (i) 'can move past each other'. Therefore, even though the participant did not state this feature in their text, the text point should be given anyway. (ii) can move, so the arrows can be used as a substitute for 'lines' signifying movement. Therefore, the drawing point for 'lines' signifying movement should be given anyway.	**1** - for 'can move past each other' [1pt]	**1** - for 'arrows' as a substitute for lines signifying movement [1pt]

Matter	Sample answer	Note	POINTS	
			Text	Drawing
Gases	Gases can flow and can be squashed.  Drawing 1 Drawing 2	Drawings indicate that the particles: (i) can move Therefore, the drawing point for 'lines' signifying movement should be given anyway. (ii) 'can be moved closer'* Therefore, even though the participant did not state this feature in their text, the text point should be given anyway. * text point should only be given if the distance between particles in Drawing 2 is closer than in Drawing 1	**1** - for 'can be moved closer' [1pt]*	**3** -particles are far apart [1pt] -particles are not in contact [1pt] *Plus* - for two drawings that illustrate the matter's property as a substitute for lines signifying movement [1pt]
	Gases can flow and can be squashed. 	Drawing indicates that the particles: (i) can move Therefore, the drawing point for 'lines' signifying movement should be given anyway.	**0**	**3** -particles are far apart [1pt] -particles are not in contact [1pt] *Plus* - illustration of the matter's property as a substitute for lines signifying movement [1pt]

Matter	Sample answer	Note	POINTS	
			Text	Drawing
Gases	Gases can flow and can be squashed. 	Arrows indicate that the particles: (i) 'can move all over'. Therefore, even though the participant did not state this feature in their text, the text point should be given anyway. (ii) can move, so the arrows can be used as a substitute for 'lines' signifying movement. Therefore, the drawing point for 'lines' signifying movement should be given anyway.	**1** - for 'can move all over' [1pt]	**3** -particles are far apart [1pt] -particles are not in contact [1pt] *Plus* - for 'arrows' as a substitute for lines signifying movement [1pt]