University of Bedfordshire

Title: Reading during an academic reading-into-writing task: an eye-tracking study

Name: Nicola Latimer

# Reading during an academic reading-into-writing task: an eye-tracking study

By Nicola Latimer

A thesis submitted to the University of Bedfordshire, in fulfilment of the requirements for the degree of Doctor of Philosophy

The Centre for Research in English Language Learning and Assessment, University of Bedfordshire

28 September 2018

I, Nicola Latimer, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have drawn on or cited the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis or any part of it is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Acknowledgements

It has taken me rather a long time to complete my thesis which has necessitated several changes to my supervisory team. This has had the rather fortuitous effect of allowing me to benefit from help and advice from a number of the esteemed and dedicated academics at CRELLA.

Firstly, I would like to express my sincere gratitude to Professor Cyril Weir. I have been extremely lucky to have had such a distinguished and revered supervisor. My thanks also go to Professor Lynda Taylor for her kind encouragement, wise advice and insightful comments. John Field and the late Stephen Bax have also been extremely generous with their time and expertise - thank you. However, my biggest thanks go to my patient, diligent and adept direct of studies, Dr Sathena Chan. I could not have asked for a better supervisor. To the whole team at CRELLA, an enormous thank you.

Of course, I could not have done it without the love and support of my wonderful family and friends. To my amazing mum and dad, who inspired in me a love of books, reading and education in general – my everlasting gratitude. To my loving husband John, who encouraged me to 'bite the bullet' and do my PhD - thank you. I could not and would not have done it without your support. To my wider family and friends, thank you for your unflagging love and support. I am truly blessed.

I dedicate my thesis to my sons, Ben and Daniel. You are the best 'things' I have ever done! I am honoured and proud to be your mum.

# Abstract

The study aimed to establish the types of reading university students engaged on an academic reading-into-writing task through a mixed-methods approach. To achieve this, eye-tracking technology was used to record 30 students’ eye-movements as they engaged in a one-hour computer-based academic reading into writing test task. After the test events, stimulated recall interviews and a cognitive process questionnaire were used to collect more comprehensive data on these students' academic reading processes. The study also investigated whether there were differences in the reading patterns of students with more experience of performing academic reading into writing tasks and those students with less experience. Differences in the way high and low scoring students tackled the task were also investigated.

30 participants (15 first year undergraduates and 15 third year undergraduates or postgraduates) were recruited from a range of UK universities. The participants were observed and their eye movements were recorded whilst they completed the reading into writing task. After the task, participants watched a replay of their reading and writing activity and were prompted to recollect their thought processes during the task. Finally, participants completed a short cognitive processing questionnaire.

This research made five key findings relating to university students' academic reading processes in the context of a reading into writing task. Participants spent almost half their time (47 per cent) fixating on their own emerging text and about a third of their time reading the source texts. The task instructions were relatively poorly attended to.

The fixations on the written source text were more homogeneous than fixations on the participant emerging work. Fixations on the written source texts reported a shorter mean

fixation duration and contained much higher rates of regression than for reading reported in the literature.

Careful reading accounted for less than 30 per cent of reading of the written source texts. Other forms of selective reading accounted for the remaining 70 per cent of reading. The predominance of selective reading appears to result from participants targeting their reading to spend more time on the more relevant sentences, although several factors seem to interact to determine total time on each sentence.

When differences between the Year 1 (Y1) and Year 3+ (Y3+) groups were examined, it emerged that the Y3+ group spent much more time fixating on their own work than the Y1 group. It also emerged that the Y3+ group engaged in more selective reading than the Y1 .The increased levels of selective reading may have contributed to the greater attention that the Y3+ group devoted to the more relevant sentences.

When the results for the five highest and five lowest scoring participants were compared it emerged that the low scoring participants were much less effective at focusing their attention on the most relevant sentences.

In short, these findings suggest that reading during an academic reading into writing task is different to the careful reading described in the literature. It demands a wide range of selective reading skills and strategies in addition to careful reading skills. Task representation can influence the way the writer interacts with the source text(s). The findings imply that development of selective reading skills, in conjunction with developing task representation skills, could help inexperienced students produce better written work earlier in their courses.

# Contents

University of Bedfordshire

Centre for Research
in English Language
Learning and
Assessment
(CRELLA)
Putteridge Bury
Campus Luton, LU2
8LE
United Kingdom
t +44 (0)1582 489159
www.beds.ac.uk/crella

26 April 2016

To whom it may concern,

**Re: Ethical approval for Nicola Latimer's research project**

Please accept this letter as confirmation that Nicola Latimer's research project titled "Investigating the Cognitive Processes in Academic Reading into Writing" has been scrutinised by CRELLA's ethics committee and has been approved.

Yours faithfully,

Dr Chihiro Inoue

Lecturer in Language Assessment
CRELLA, University of Bedfordshire

# List of tables

## List of figures

# 1 Introduction

## 1.1 Background: The changing face of UK university population

In recent decades UK universities have attracted an increasingly diverse range of students. Two factors have been key in broadening the range of students attending higher education in the UK.

Firstly, the UK's higher education institutions have long been considered world leading which has attracted international students from all over the world. It has been estimated the UK economy benefitted by about £8 billion pounds in 2008/09 due to this 'trade' in higher education (Conlon et al., 2011). The benefits to the UK are not limited to financial income. In 2013, the Careers Research and Advisory Centre published a report suggesting that there were other qualitative, but no less important, benefits to international students' attendance at UK universities (Mellors-Bourne et al., 2013). These benefits included increased trust between the UK and nations sending students to the UK and an increase in the positive understanding of the UK's culture and values.

Secondly, the UK government has sought to broaden the range of home students applying to and attending university. In 2010, the Conservative and Liberal Democrat coalition government issued a policy document stating that '(a)nyone with the ability who wants to go to university should have the chance to do so, whatever their economic or social background.' (Department for Business Innovation and Skills, 2010). Subsequently, 2013 saw a relaxation of the previous limit on the numbers of domestic students annually accepted onto university courses, followed by the removal of any limit the following year.

These twin factors have seen the numbers of both home and international students attending university expand. Many of the students starting university as a result of the expansion come from sectors of society that have traditionally been underrepresented. These changes are, many would argue, good for the UK economy and good for UK society as a whole, increasing social cohesion and opportunities for all. However, for universities charged with delivering courses it has meant adapting to accommodate and support students with a broader range of needs - both academic and more holistic needs. Much of the academic support required is in the form of developing students' levels of academic literacy. Academic literacy is essential for all students at all universities across the country however, universities with high levels of international students and universities recruiting large numbers of students in response to the widening participation agenda are likely to face substantial challenges.

This study was conducted, in part, at the University of Bedfordshire Luton campus and recruited approximately one third of its participants (total 30 participants) from the university. The University of Bedfordshire is one of the 'new universities', created in 1992 when former polytechnics and colleges of higher education were converted to universities as a result of the Further and Higher Education Act 1992. The university has a positive approach to diversity, aiming to improve access to higher education for groups such as those from neighbourhoods with low rates of participation in higher education and pupils eligible for free school meals. In 2011 the Office for National Statistics reported that around eight per cent of the population of England and Wales do not consider English to be their main language. However, in Luton, where the university is based, the figure is likely to be considerably higher. For example, according to The National Association for Language Development in the Curriculum website (2013) around 51 per cent of children in Luton's state primary schools did not have English as their first language. This suggests that the university is likely to have a higher than average number of home students with English as a second or additional language. The university also recruits a high proportion of non-UK students. According to The Complete University Guide (2018) 18 per cent of the student body in 2017-2018 came from outside the UK. Therefore, the university's student body is likely to be diverse and include students from a wide range of backgrounds.

An increase in the numbers of people attending university is to be welcomed but brings with it the challenge of offering the appropriate academic support to ensure the increasingly diverse range of students fully benefit from the opportunities that university offers.

## 1.2 Encouraging the development of academic literacy

Recruiting a broader range of students will be of little benefit to individual students or to society as a whole if those newly recruited groups do not thrive and succeed during their courses and in the world of work after leaving university.

The exact definition of academic literacy is much debated (Geisler, 1994) and even the way we approach defining academic literacy varies (Lea and Street, 2006) but inevitably, is considered essential to the success of any university student (Lea and Street, 1998). Academic literacy encompasses much more than simply the ability to read and write but reading and writing are at the heart of academic literacy. Reading and writing provide the foundation skills that enable students to access and participate in the academic world.

As students progress through their studies, they are expected to develop their reading and writing skills to a sophisticated level. They are expected to incorporate an understanding of the academic genre in which they find themselves, adhering to the stylistic conventions of their institution as well as acknowledging and critically evaluating the sources they read. Therefore, academic reading differs from many other types of reading; it is usually driven by the need to inform academic writing - sometimes referred to as 'reading into writing'

Several studies (Carson, 2001; Carson and Leki, 1993; Flower, 1990; Grabe, 2001, Johns, 1993; Leki and Carson, 1994, 1997; Lenski,1998) have suggested that reading into writing ability is essential for academic success not least because reading-to-write is one of the commonest forms of assessment in higher education (Bridgeman and Carson, 1983; Hale et al., 1995; Rosenfeld, Leung, and Oltman, 2001). Therefore, a thorough understanding of how reading into writing operates in an academic context would seem to be essential in devising interventions to support the development of students' academic literacy. At present there is a limited understanding of the reading involved in reading into writing and eye tracking offers the opportunity to investigate this further.

## 1.3 Aims of this study

Effective teaching of academic reading and writing skills must be underpinned by an in-depth understanding of how these skills function in practice. Such an in-depth understanding offers opportunities to improve identification of those with a skill deficit and develop more targeted interventions to aid the development of skills. Therefore, this study seeks to increase our understanding of reading as part of an academic reading into writing task by recording and analysing the reading activity of students whilst they engage in an academic reading into writing task.

There are several widely quoted interactive models of reading (e.g. Rumelhart, 1994, Just and Carpenter,1980, Kintsch and Van Dijk, 1978, Perfetti, 1999, Rayner and Pollatsek, 1989, Khalifa and Weir, 2009) and several researchers have made valuable contributions to our understanding of the process of reading into writing (Chan, 2013, Plakans and Gebril, 2012, Spivey and King, 1989, and Yu, 2008), there remains a gap in the literature in relation to how different types of reading are utilised in the process of reading to write.

Whilst not a model, Chan (2013) offers a detailed account of the cognitive processes involved in reading into writing including the role played by high-level reading skills in the academic context. The high-level reading skills in Chan's account draw on the Khalifa and Weir model (2009) to distinguish between different types of reading proposed by Khalifa and Weir which vary according to the reader's purpose for reading.

Using the cognitive processes outlined by Chan (2013) and Khalifa and Weir's (2009) model of reading to guide analysis of eye tracking data, the research aims to offer a detailed factual account of the reading which participants engaged in whilst completing a short reading into writing academic task. This study aims to use retrospective recall data and questionnaire data to triangulate and assist with interpretation of the eye tracking data.

## 1.4 Structure of thesis

This thesis is divided into six chapters. After the introduction chapter 2 reviews the literature relating to reading into writing and eye tracking reading before identifying the gaps in the literature that prompted the research questions . The methodology (3) describes the data collection procedure, the development of the research instruments and the procedures used to analyse the data. Chapter 4 sets of the findings in relation to each of the research questions. The discussion chapter 5, attempts to draw the findings together into a broader understanding of the way participants tackled the reading into writing task used in this study. Finally, chapter 6 suggests what the findings in this study might add to our understanding of the way students tackle academic reading into writing tasks more generally.

# 2 Literature Review

## 2.1 Introduction

This chapter is broadly divided into four sections. The first section (2.2) reviews the literature relating to academic literacy before considered the extent to which reading into writing tasks elicit the skills requisite for academic literacy.

The second section (2.3) considers what the models of reading and writing in the literature can add to our understanding of the reading into writing process. This section includes a discussion on cognitive processes because models of reading and writing are constructed in terms of the cognitive processes involved.

The third section (2.4) reviews the literature relating to the use of eye tracking to investigate the cognitive processes involved in reading.

The fourth and last part considers how eye tracking data might be interpreted in light of the literature (2.5) to offer insight into the cognitive processes of reading into writing.

The chapter concludes by outlining the research questions posed by this study (2.6).

## 2.2 The role of reading into writing in academic literacy

This section reviews the literature discussing the role of reading into writing in academic literacy. In order to consider this, we must first ask what is meant by 'academic literacy'. There is no universally agreed definition of academic literacy. It is not something that can be defined concisely in one or two lines, it is much discussed and debated. Therefore, some of the literature that usefully contributes to the discussion is now reviewed.

Weir and Chan (forthcoming) synthesises the abilities and skills that several leading UK universities cite as being essential reading and writing skills for undergraduate students. Those skills include identifying appropriate sources and applying different reading skills effectively (such as reading quickly and selectively, reading careful, reading critically). Also included are writing skills such as demonstrating a thorough understanding of the subject matter, presenting information in a balanced and unbiased manner, supporting views or opinions with evidence, accurately summarising and acknowledging sources.

Weir and Chan (forthcoming) argue that many of the skills or abilities outlined above require the student to add their own original thoughts or interpretation of the information, transforming the source information by adding another level of interpretation. This is what Bereiter and Scardamalia (1987) refer to as knowledge transformation. Knowledge transformation relies upon the generation of new or novel ideas based on reviewing the relationship between existing facts and ideas. In short, when writing, it is not enough to simply re-tell or summarise the facts; the writer must reflect and generate a novel response or approach beyond that which has been proposed in the text.

The skills demanded of students at tertiary level in the UK go beyond being able to read large quantities of expository texts and reporting upon their contents. Bereiter and Scardamalia (1997) term this type of reporting, without the generation of any novel ideas or approaches, knowledge telling. Much of the discussion surrounding academic skills refers to 'knowledge telling and knowledge transforming' as defined by Bereiter and Scardamalia's 1997 work. Bereiter and Scardamalia's work outlines two models, one of knowledge telling, one of knowledge transforming. These are discussed in turn below.

**Knowledge Telling**

Bereiter and Scardamalia suggest that when we write in response to an assignment we rely on two things to help us generate our texts. Firstly, we consider what we know about the subject and they refer to this as content knowledge. Of course, the writer could then just 'pour' their content knowledge onto the page unchecked and in no particular order or structure; however, Bereiter and Scardamalia suggest that as children mature into adults they begin to monitor and shape that content knowledge. This consideration of the style and format that we think may be appropriate for the task is called discourse knowledge. Bereiter and Scardamalia suggest that this combination of content and discourse knowledge helps us generate our written response to the assignment. We then engage in a cycle of writing, reading what we have written (which may elicit more ideas from our memory) and editing and so on. We use information that we retrieve from our memory, possibly alongside information that we have read / heard and reorganise it into a coherent whole according to our discourse knowledge. This sounds like a perfectly reasonable way to tackle the assignment and a relatively unproblematic solution and in a limited way it is. However, this knowledge telling approach will not lead to the generation of any new or novel ideas. It is merely a device for telling others what we know and what we have read.

Writing tasks which ask students to write about what they already know (tasks which do not demand that students engage with source materials) are likely to result in knowledge telling. Tasks which ask students to draw on a single source of information are also likely to result knowledge telling. Only when students are set tasks which require them to integrate information from more than one source of information is knowledge transformation likely to emerge.

**Knowledge Transforming**

Bereiter and Scardamalia suggest that reflecting critically on the content within sources and the nature of their relations represents a higher order of skill. This skill relies on reflection about the relationship between sources to produce novel ideas or new approaches which develop the subject under consideration beyond the writer's existing knowledge and beyond what they have gleaned from reading the source materials. Bereiter and Scardamalia refer to this higher-level skill as knowledge transformation. Tasks which demand integration of multiple sources of information will not always elicit knowledge transformation, but they offer the challenge of knowledge transformation in a way that writing only or writing from single sources do not.

Bereiter and Scardamalia's (1987) models identify the cognitive processing that people go through when writing to 'tell' knowledge and writing to 'transform' knowledge. They identified clear differences between knowledge telling and knowledge transforming. Whilst knowledge telling is a more linear process, with content knowledge and discourse knowledge feeding into the process as it progresses, knowledge transformation is much more interactional, reflective, cyclical process. The analysis of the problem (the demands of the assignment or task) help shape the writers' goals. The formation of these goals leads to a constant interaction between consideration of what to include (content knowledge) but also how to ensure that the content is re-shaped and adjusted to suit the present writing goals (discourse knowledge). A cyclical process develops as writers reflect on their proposed content and how it might be transformed to meet the needs of the current task. Reflection on how to interpret content in light of the question may lead to a revision in understanding of what content is required. Once more content is added, this may lead to a refinement in how content is shaped to suit the task. This cyclical problem-solving process eventually leads to a deeper, more sophisticated understanding of how different content selected in answer to the task relates to each other and to the task set.

The researcher suggests that Bereiter and Scardamalia's concepts of knowledge transformation lie at the heart of what it means to be academically literate. Flower (1990) suggests that reading into writing tasks demand the knowledge transforming processes described by Bereiter and Scardamalia (1987).

Flower (1990) uses the phrase 'critical literacy' to describe the process of reading in order to write where the writing does not just report or retell the information gleaned but transforms it for a new purpose. Flower suggests that a critically literate person is able to read facts and 'turn facts into concepts, to turn concepts into a policy or a plan, and to see the issue and define the problem within a problematic situation'. (ibid p5)

Flower's definition of a deeper, more sophisticated literacy suggests that 'reading-to-write… is a litmus test that lets us distinguish between the receptive process of basic literacy and the testing/transforming process of critical literacy' Flower (1990: p5). The researcher suggests that Flower's description of critical literacy accurately represents the skills demanded of university students studying in the UK today.

The discussion above underlines that writing about what one already knows is unlikely to constitute knowledge transformation; it is more likely to represent a knowledge telling exercise. For knowledge transformation to take place, students will need to be challenged to interpret information from sources.

Weir and Chan (forthcoming) suggests that 'students' abilities to summarise or integrate ideas from different sources would appear to be a critical focus for assessing a student's academic literacy.'

Given the argument that reading into writing tasks which require the integration of information facilitate knowledge transformation in a way that writing-only tasks do not, it is perhaps not surprising that a number of studies have considered the predictive validity of reading into writing test tasks and concluded that they offer advantages over writing only test tasks.

Gebril (2010:113) concludes that the implications of his 2010 study:

*provides support for many of the testing programs that adopt an integrated approach to language test development. This type of information would also be helpful for university administrators when they make admission or placement decisions. In other words, having a composite score involving different writing tasks would provide a more accurate picture of a student's writing ability and also could be a good predictor of the success of this student in university classes.*

In their summary, Cumming, Kantor, Baba, Erdosy, Eouanzoui and James (2005) concluded that there were significant differences between the writing that examinees produced in independent essay tasks and the integrated reading-writing or listening-writing tasks. Cumming et al. suggested that their findings supported the inclusion of integrated tasks in Next Generation TOEFL.

Cumming et al. are not alone in concluding that reading into writing tasks have much to offer over independent writing tasks. Asencion-Delaney suggests that 'when reading to learn or to integrate, reader/writers construct elaborate models of the text structure and situation, enabling them to select information from the source text, evaluate it, and use it for writing purposes' (2008:p141). Grabe (2003) also suggests that in performing reading into writing tasks students need to make a series of complex decisions:

*1. How much information should be taken from the text; which information should be taken*

*2. How the information taken will fit with task and writer goals*

*3. How accurately the information should be represented when going from text source to student writing*

*4. What formal mechanisms should be used for transforming or using the textual information*

*Writing from multiple texts requires even more demanding planning, processing, and revising. The interpretation of task demands and the integration of textual information force critical decision making that requires much practice and consistent efforts to 'traverse the topical landscape' from multiple directions.* *Grabe (2003: p242-3).*

In an earlier study, Cumming, Kantor, Powers, Santos and Taylor (2000) suggested that there is a key difference between the demands made on undergraduates and graduates, emphasising that 'in many undergraduate contexts, the emphasis of writing is often on telling people about the knowledge one has, or is acquiring, rather than on using writing to create unique or novel knowledge, as might be expected in graduate studies' (ibid: p5). However, the researcher suggests it is debatable that undergraduate writing tasks require students simply to refer to source texts in a knowledge telling sense. If, when writing, students refer to more than one text, they are expected to comment on the relationship of those to texts either to each other or to a central theme. It is not enough to simply relate the contents without discussing their relevance. This, in accordance with Grabe's earlier citation, the researcher would argue represents at least partial knowledge transformation and requires a complex range of skills which extend beyond reading and summarising a text or series of texts.

Leki (1993:p9) also appears to highlight another factor which usefully contributes to this discussion. Leki points out that many studies which have investigated L2 reading and writing skills independently report aspects which are common to both.

*…these studies reveal that less skilled readers and writers both appear to attend to the same thing, to the text on the page rather than to the meaning potential of that text, to the forms of the letters and words rather than to the overarching connection between them.*

Leki goes on to point out that the more complex… (1993:p12) *ability to integrate or internalize new information in writing that undergirds the notions both of knowledge-transforming (Bereiter and Scardamalia, 1987) and of critical literacy (Flower et al., 1990) … may in fact be what we actually mean when we speak of comprehension of a text.*

The researcher suggests that this highlights a critical factor for assessing students. In order to cope successfully with the demands of higher education in the UK, students' reading and writing skills must be such that they do not absorb so much of the student's attention that they are unable to focus on the bigger picture, the connections which need to be made between the texts they are reading in order to inform and develop new hypotheses. Alamargot, Plane, Lambert and Chesnet (2010) suggest that gradual automation of the lower level processes is achieved as expertise develops. Alamargot et al. used eye and pen tracking to study the cognitive processes of candidates whilst they read the opening paragraphs of a story and then wrote a conclusion to the story. The candidates in their study varied from novice writers (school grade 7) through to expert writers. Alamargot et al. concluded that growing automation of first low-level and then high-level processes enables writers to speed up the composition process and move from a knowledge telling approach to a knowledge transforming approach, thereby giving greater attention to high level planning and coherence.

In assessing reading and writing separately it is much harder to ensure that this differential between language-based skills and critical literacy is established. Indeed, Leki and Carson's (1997) study of a US university EAP writing course found considerable benefits to students completing 'text-responsible' writing tasks, in other words tasks which necessitated reading into writing. This perhaps underlines the argument that reading to write, or writing based on reading texts, is more challenging than performing separate reading and writing tasks and that practice of such tasks is useful for prospective university candidates.

Several studies have advocated that academic writing skills need to be considered in the light of their wider context. For example, Lea and Street (2006:p368) propose that academic skills need to be considered from a much broader perspective that…

> *views the processes involved in acquiring appropriate and effective uses of literacy as more complex, dynamic, nuanced, situated, and involving both epistemological issues and social processes, including power relations among people, institutions, and social identities.*

Lea and Street's work focused on the nature of academic literacy and reviewed the problems students face in adapting to tertiary education in the UK and the type of interventions which may be helpful in facilitating the development of the skills required. They suggest that 'an approach using the academic literacies model foregrounds the variety and specificity of institutional practices, and students' struggles to make sense of these' Lea and Street (2006:p376). Lea and Street (2006:p368) define the academic literacies model to which they refer as follows.

> *An academic literacies perspective treats reading and writing as social practices that vary with context, culture, and genre (Barton & Hamilton, 1998; Street, 1984, 1995). The literacy practices of academic disciplines can be viewed as varied social practices associated with different communities. In addition, an academic literacies perspective also takes account of literacies not directly associated with subjects and disciplines, but with broader institutional discourses and genres. From the student point of view, a dominant feature of academic literacy practices is the requirement to switch their writing styles and genres between one setting and another, to deploy a repertoire of*

*literacy practices appropriate to each setting, and to handle the social meanings and identities that each evokes.*

The literacy described here by Lea and Street is more complex than the need to combine sources in an original way. It also suggests the need to adapt writing to meet socially constructed demands such as genre and subject specific contexts. If students are expected to juggle multiple levels of demands, adapting their writing to suit, then the case for students having automated lower level skills becomes even more critical.

Thus far the researcher has outlined what might be meant by academic literacy and, as discussed above, established that the concept of academic literacy is much deeper and wider than simply being able to read and write. Tasks which demand writing in response to a text and which are situated in an academic context are far more likely to demand the cognitive 'juggling act' which characterises most written academic assignments in UK higher education. Nevertheless, reading and writing provide the foundation skills enabling students to access and participate in the academic world and therefore the following sections will review how both reading and writing skills operate and how they may interact during reading into writing. Most of the literature regarding the processes of reading and writing refer to cognitive processes. Therefore, to begin, it would be useful to establish what is meant by a cognitive process.

## 2.3 Models of reading and writing that may help us understand reading into writing

Much of the literature on models of reading and models of writing discusses these skills in terms of cognitive processes. Therefore, before reviewing models of reading and writing there follows a discussion of what is meant by the term cognitive process.

### 2.3.1 What do we mean by cognitive processes?

Field (2004, p. 61) defines cognition as 'the use or handling of knowledge'. Field elaborates by referring to cognition as both a 'faculty' and a 'process'. With regards to this study we are interested in exploring the cognitive processes which take place whilst students engage in reading in order to write. This might be restated as exploring the way in which students process the knowledge they acquire from reading texts and integrate it with their existing knowledge to produce a written academic assignment.

Of course, studying cognitive processes is a much-debated area. After all, we cannot observe or measure with certainty what happens in the mind. We can observe behaviour that results from cognitive processes or we can measure brain activity in the form of electrical activity but neither of these tells us what 'thinking' is taking place behind the scenes. We are therefore limited to interpreting behaviour and drawing conclusions about the cognitive processes it represents and asking subjects to report what they are thinking, and this is inevitably problematic.

For the moment let us focus more closely on defining what cognitive processing is in relation to reading into writing. Field (2004) and Eysenck and Keane (2005) suggest that many of the ideas that are current in cognitive psychology arose from the information processing approach (Newell, Shaw and Simon, 1958) which was developed in the 1950s. This approach focuses on the flow of information through the mind as tasks are performed. As a result of this information processing approach many researchers have adopted the opinion that human behaviour is based on problem-solving. Much of the recent research into the nature of expertise has been carried out from this standpoint. Newell and Simon (1972) report that cognitive psychology has developed protocol analysis (or 'think aloud' as it is sometimes referred to) as a powerful tool with which to investigate the processes people use when tackling problem-solving tasks. This perhaps explains the widespread use of protocol analysis for studying reading into writing.

Turning our attention to the type of cognitive processes that readers / writers engage in, Field (2003) describes 'an information processing approach' in which any form of communication is broken down into a whole series of stages or sub-tasks which need to be completed in order to process the information. Field uses the example in Figure 1 to illustrate the kind of information processing that occurs when somebody asks you the way to the station, for example.

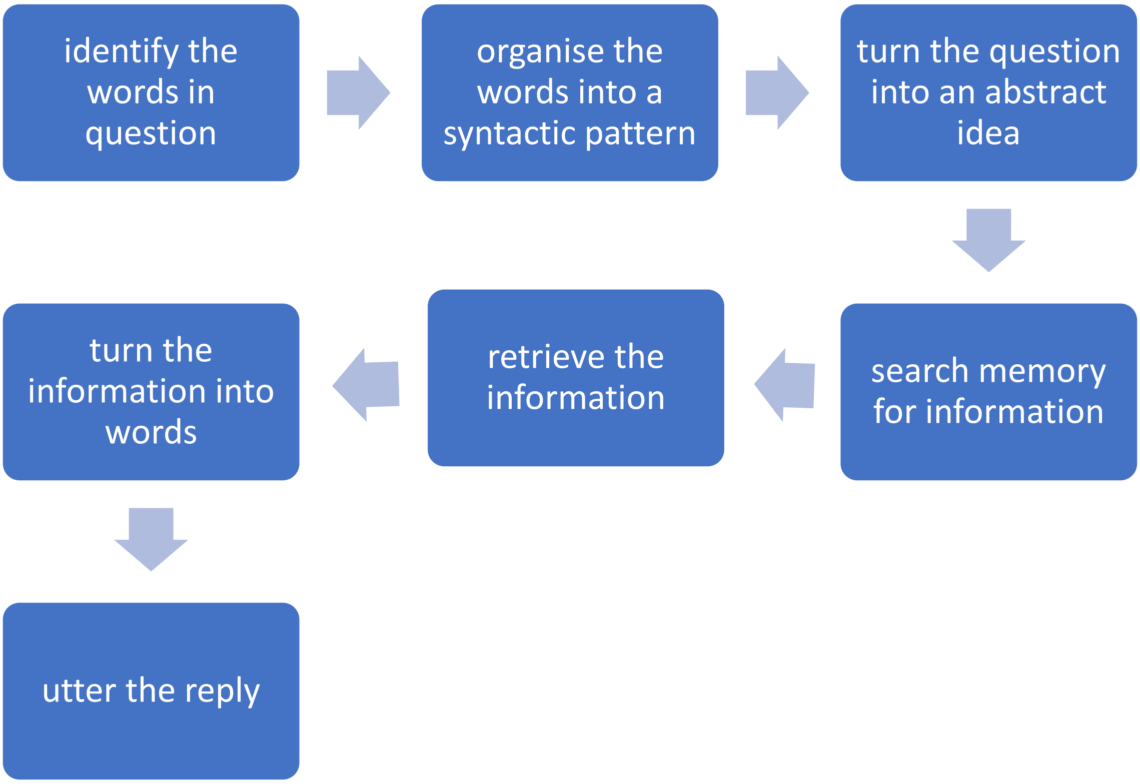



*Figure 1 Cognitive stages in listening and responding to a question Field (2003:p17)*

However, Field goes on to suggest that this is a simplification of the way communication is processed. Field suggests that this kind of 'bottom-up' serial processing of information is not what really happens. Instead Field suggests that the listener or reader is simultaneously using the context (the bigger picture) to analyse the information in a 'top-down' way. For example, when a stranger approaches in the street, we are already thinking back to previous such encounters and beginning to predict what the stranger may want or say.

Therefore, we can begin to see that cognitive processes happen at a range of levels, some of which are more conscious and effortful, and others occur at a level which escapes our notice on a daily basis. In other words, some **low-level processes happen so quickly and automatically that in everyday life we are unaware of them**. Only when a break-down in communication occurs do we occasionally have cause to review them; for example, if we mishear a word which causes a sentence to be syntactically wrong or semantically unlikely, then we may be fleetingly aware of our re-interpretation of the sounds into a more plausible word. When reading, low level processes like word recognition are automated in adult readers / competent readers to the extent that for the most part readers are unaware of the process. Likewise, in writing, low level processes like spelling are usually automated for adult writers, most of the time.

**Higher level processes are more conscious and effortful**. In Field's (2003) example in Figure 1, the person giving directions will probably be aware of actively considering where the station is in relation to their current location. In reading, higher level processes might include working out which ideas are key / central to text. In writing, deciding how to order the ideas within the text would be a high-level process.

The level of consciousness associated with a process affects how we can record it / evidence it. Higher level cognitive processes which are conscious have the ability to be accurately reported by participants using techniques such as protocol analysis or think aloud methods of research. For other lower level processes, the sub-conscious nature of them puts them largely beyond the reach of accurate self-reporting. However, as will be discussed later, some methodologies such as the study of eye-movements during reading, can be used to infer some of the lower level processes.

Having outlined what is meant by ‘a cognitive process’, we will now go on to review the models of reading and writing available in the literature and consider how their cognitive processes may interact during reading into writing.

### 2.3.2 Models of writing

Flower and Hayes’ (1981) model of the writing process represents a seminal work that has been widely quoted and reviewed. Perhaps most importantly for this study, it acknowledges reading in the process of writing. The model is shown in Figure 2 and alongside the model Hayes and Flower (1981:p366) present four key points on which their theory of writing is based. The four key points are:

> 1. *The process of writing is best understood as a set of distinctive thinking processes which writers orchestrate or organize during the act of composing.*
> 2. *These processes have a hierarchical, highly embedded organization in which any given process can be embedded within any other.*
> 3. *The act of composing itself is a goal-directed thinking process, guided by the writer's own growing network of goals.*
> 4. *Writers create their own goals in two key ways: by generating both high-level goals and supporting sub-goals which embody the writer's developing sense of purpose, and then, at times, by changing major goals or even establishing entirely new ones based on what has been learned in the act of writing.*

In the explanation that accompanies this model Flower and Hayes are clear that the revision process is a cyclical one in which the writer will repeatedly review the text so far. This reviewing is naturally in the form of reading the text produced so far.

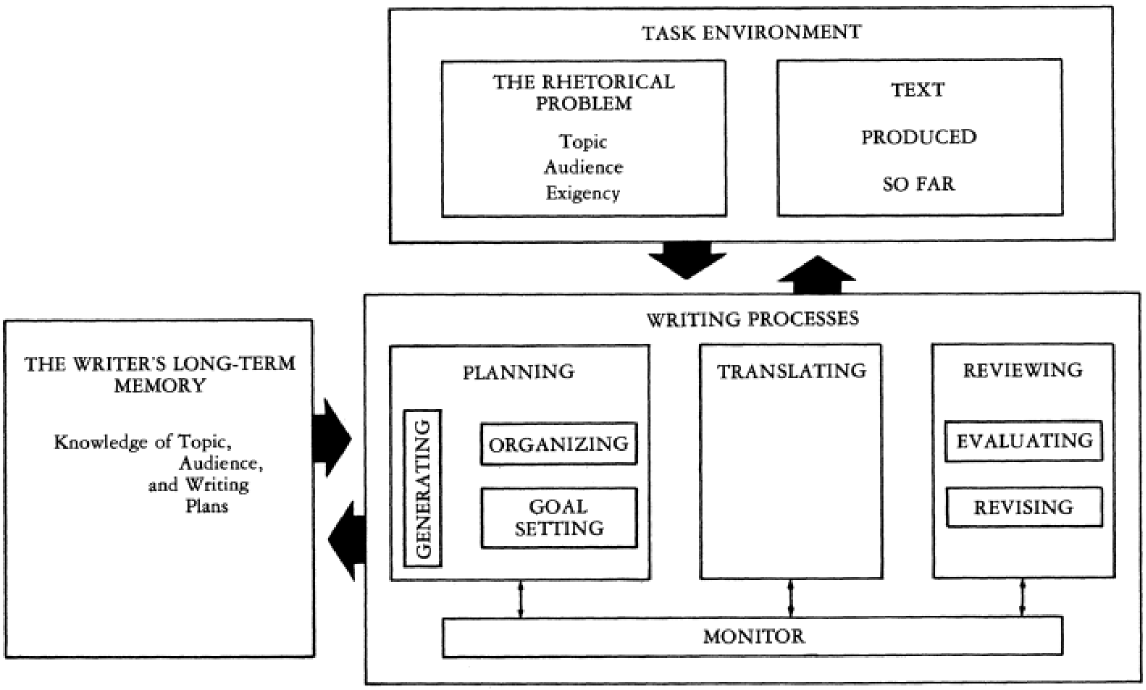



*Figure 2 Hayes and Flower's 1980 Model of writing*

Hayes and Flower's 1980 model attempts to represent what Hayes and Flower see as a complex process in which 'writers are constantly, instant by instant, orchestrating a battery of cognitive processes as they integrate planning, remembering, writing, and rewriting' Flower and Hayes (1981 p387). They are at pains to point out that the processes described 'have a hierarchical, highly embedded organization in which any given process can be embedded within any other.' (ibid p366.)

Essentially Flower and Hayes identify two key areas which exert influence on the writing process. These are the Task Environment (incorporating the rhetorical problem and the text produced so far) and the writer's long-term memory (incorporating the writer's knowledge of the topic, assumptions about the audience and writing plans). Both of these areas interact with the process of writing as it proceeds.

The writing process itself is divided into three sub-processes: planning, translating and reviewing, all of which interact with the fourth sub-process, monitoring. Whilst the diagrammatic model provides a relatively clear and simple picture of the process Hayes and Flower are clear that, in fact, the reality is a complex one. Their use of the term 'orchestrating' perhaps casts the writer as a conductor, directing different sections of the orchestra; however, it could be added that there is no score to follow! Instead the conductor is doubling as composer, attempting to compose the score as he goes, repeatedly testing out new chords and phrases, adding in small sections of familiar ones and repeatedly replaying the emerging score to check if it captures the elusive musical motif he feels is playing just out of earshot.

Hayes and Flower's model (1980) is useful when reflecting on the cognitive processes of reading into writing. If this model were to be adapted to reflect reading into writing, in addition to the writer's long-term memory, the source or reading texts would need to be represented. A revised model might see readers drawing on the texts for information which then interacts with their memory, prompting new ideas. Hayes and Flower (ibid.) account for the role of the writers emerging text in their model, but any model of reading into writing would have to consider whether the role of the emerging text is the same in writing only and reading into writing tasks, or whether the role of the emerging texts changes.

This model was later updated by Hayes (1996) and the revised model shown in Figure 3 places much greater emphasis on reading, citing reading as a central process in writing.

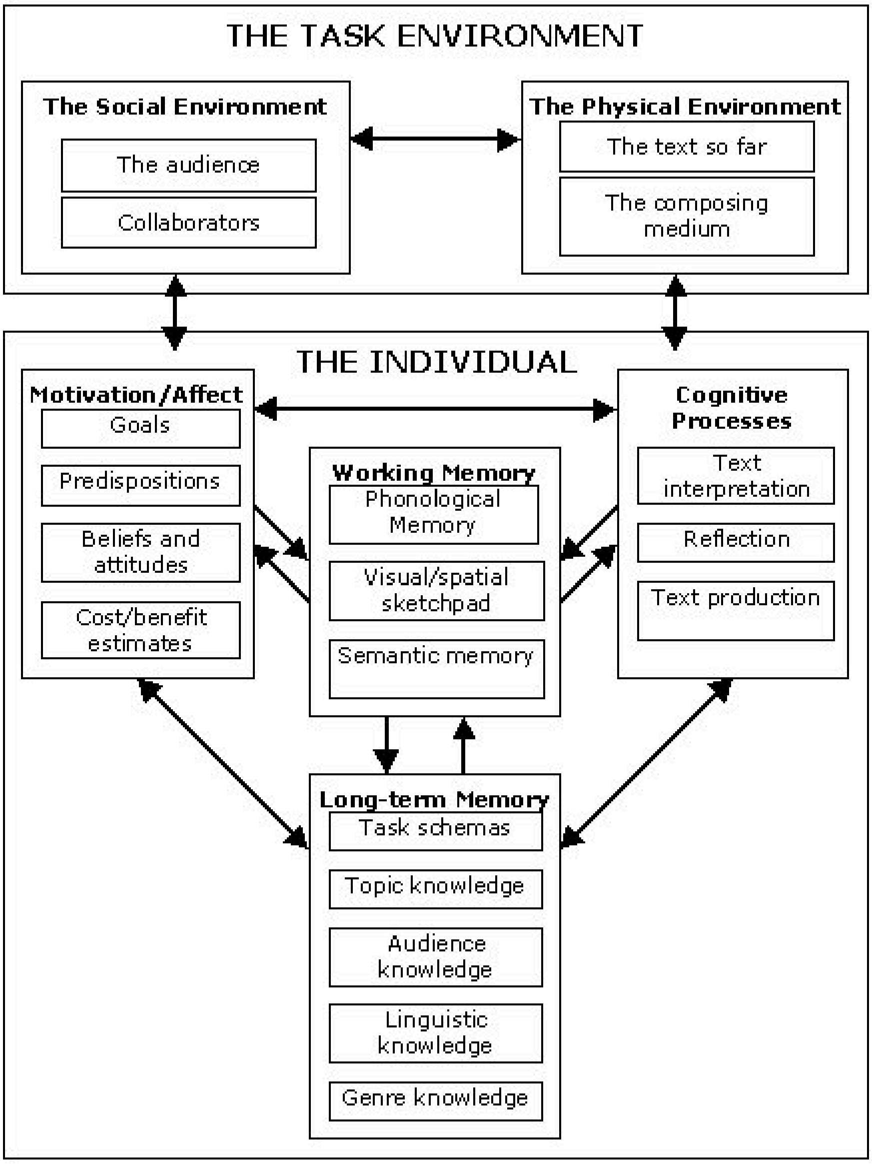



*Figure 3 Hayes new model of writing Hayes (1996:4)*

Reading is referred to repeatedly in the chapter that accompanies Hayes' new model and Hayes identifies three distinct roles for reading.

**Reading for comprehension;** the first change regarding reading is the inclusion of source texts in the writing process. Hayes notes that the social environment incorporates 'other texts that the writer may read while writing' (Hayes 1996:4).

**Reading to define the writing task;** Hayes cites reading to understand the task as being central to a successful outcome.

**Reading to revise;** more significantly Hayes recognizes the central role of reading what has been written so far and 'the text so far' is identified as a crucial factor in the task environment.

In terms of cognitive processes, instead of using 'revision' (as in the earlier model) Hayes uses the term 'text interpretation' and 'reflection' in the later model and goes on to say 'Text interpretation is a function that creates internal representations from linguistic and graphical inputs. Cognitive processes that accomplish this function include reading, listening and scanning graphics.' (Hayes 1996:4)

Hayes' (ibid.) revised model emphasises the central role of reading in the writing process. According to Hayes' model the way in which the writer reads, and the cognitive processes that reading triggers, would seem to impact very heavily on the writer's final written product. The researcher suggests that whilst Hayes makes a valuable contribution to understanding the role of reading in the writing process, it is an area that would benefit from further investigation.

When reviewing Hayes' new model, Wengelin, Leijten and Van Waes (2010) suggest that reading during writing can serve several purposes.

***Reading for revision***: Wengelin et al. suggest that here readers are interested in reading to check for comprehension and 'the identification of various text characteristics, such as spelling errors, poor lexical choice, or poor organization' (2010:736). Wengelin et al. (ibid) go on to state that ***Reading for evaluation*** 'could also involve considerations of whether the text conforms to the structures of the given genre and achieves the writers' communicative and rhetorical goals.' Finally ***Reading to facilitate*** may offer writers ''opportunities for improvements that do not stem from problems'' (Hayes, 1996: 15). This accords with Johansson, Johansson, Wengelin and Holmqvist's (2008) suggestion that reading the text so far can assist with improving the quality of the text and generating new ideas.

The researcher suggests that the reading to evaluate and reading to facilitate suggested by Johansson et al. (2008) may contribute to, or play a part in, knowledge transformation proposed by Bereiter and Scardamalia (1987).

Chenoweth and Hayes (2001) propose a model for writing which could also be interpreted as incorporating the reading process. They suggest that the process of writing requires the coordination of three levels: control level, process level and resource level.

The resource level includes internal 'resources' – knowledge and skills which enable the writer to retrieve relevant information from long term memory, hold partially composed sentences in working memory and apply sub-skills such as letter/word recognition or letter formation/spelling which subconsciously contribute to the reading and writing processes.

The process level includes both internal and external components. External components include the text composed so far, task materials (such as the question prompt, reading / input texts, feedback and the writers notes) and resources such as dictionaries, spell checkers and style guides. Although not shown on the model, Chenoweth and Hayes suggest that this external section also incorporates the perceived 'audience' for the text as well as representing more generally the social and physical elements incorporated by Hayes' 1986 'task environment'. The internal processes are responsible for: generating the abstract ideas to be written about (proposer), converting the ideas to linguistic form (translator), evaluating the language used (both that already written down and that about to be), and the transcriber which converts linguistic strings into words on paper / screen.

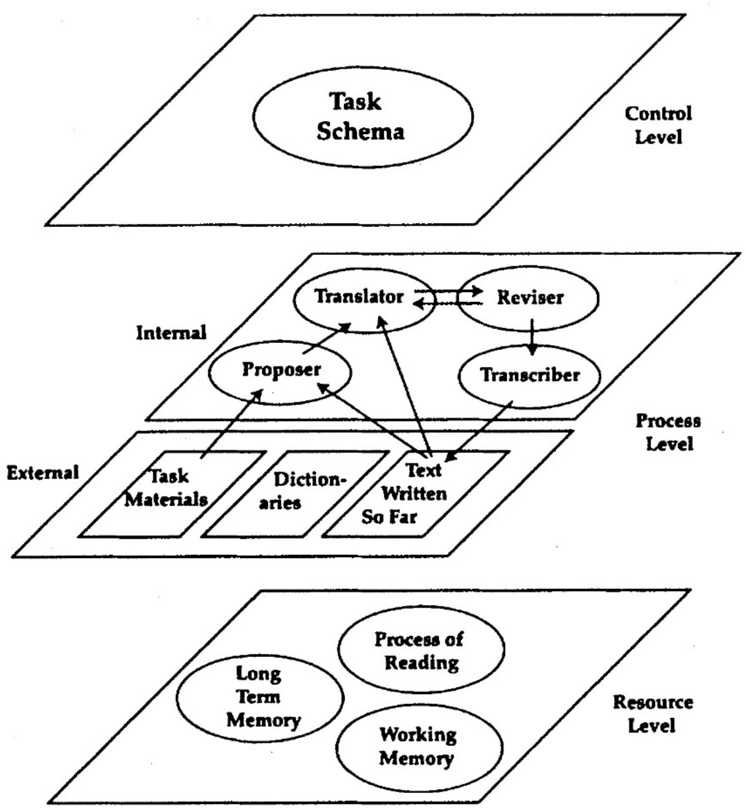



*Figure 4 Chenoweth and Hayes 2001 Model of writing (2001:84)*

The control level acts as the conductor, reviewing the task goals and activating / directing the processes in the process level. Chenoweth and Hayes stress that the pattern of interactions between processes in the process level will be different from one person to another and from one task to another depending on the writer's perception of the task schema.

The researcher finds the Chenoweth and Hayes model to be logical and convincing; however, the researcher suggests that when considering academic writing in particular, external influences are rather under represented. For example, there is little mention of the audience and the writer's consideration of the reader. In addition, the control level is rather vague and does not really show how the decisions made there interact with external influences, experience and working memory. Once again, this model lacks detail in terms of how specific reading processes interact with the task, the text so far and task materials to account for the cognitive processes involved in reading into writing.

Many different models of writing agree that working memory plays a vital role in the process. Some argue that working memory is limited and the availability of working memory accounts, in part, for the differences between expert and non-expert writers (e.g., Fayol, 1999; Kellogg, 1987; McCutchen, 1996; Olive and Kellogg, 2002; Swanson and Berninger, 1996). Non-expert writers, who have not yet mastered some lower level writing processes to the extent that they are automated, have less working memory to devote to macro writing goals. Torrence and Galbraith (2006:p70) suggest that for non-expert writers…

> *Devoting resources to these low-level processes leaves less capacity for syntactic processing, content retrieval, rhetorical-structuring and so forth (e.g., Fayol, 1999). Hence, novice writers produce shorter and less complex sentences and texts compared with those of writers who have achieved greater levels of orthographic and grapho-motor automaticity.*

In other words, having mastered basic skills such as spelling, writing or typing and composing grammatically accurate sentences in addition to having a good range of vocabulary at their disposal, more accomplished writers are able to devote their attention to higher level skills such as coherence. This concept may prove very relevant to reading into writing where, the researcher would argue, even greater demands are placed on working memory as students attempt to integrate the two skills.

### 2.3.3 Models of reading

Whilst the models of writing discussed above all give reading a place in the writing process, none of them discuss how different cognitive processes interact during the reading process to achieve comprehension. Grabe and Stoller (2011:25) and Rayner et al. (2012:21) suggest that metaphorical models of reading aimed at explaining how reading comprehension takes place can be divided into bottom-up models, top-down models and interactive models. As with Field's (2003) model of listening discussed in section 2.3.1, interactive models suggest that both low level processes (such as word recognition in the example in section 2.3.1) and high-level processes (such as 'consideration of previous experiences' in Field's example) work simultaneously to enable comprehension of the text on the page.

The account of comprehension provided by Perfetti and Adlof (2012) also suggests that the processes that are responsible for comprehension of a text occur at multiple levels of language (orthographic, word recognition, parsing of sentences and so on) and those from word recognition upwards interact with the reader's general knowledge to achieve comprehension.

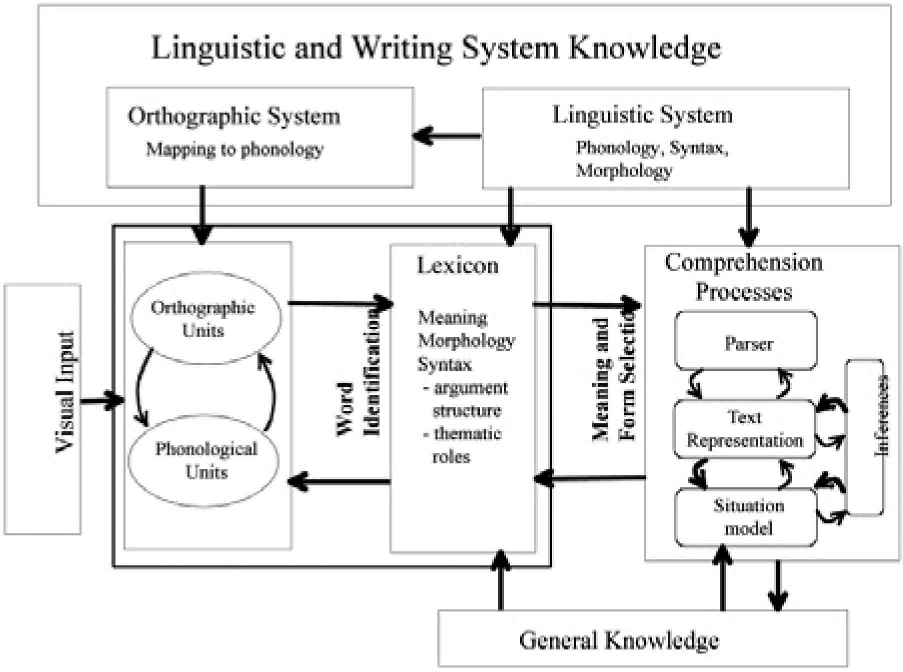



*Figure 5 Perfetti and Adlof (2012) model of reading comprehension*

In their discussion of their model Perfetti and Adlof (2012) suggest that skilled readers engage in three processes which less skilled readers do not. These processes are inferencing, comprehension monitoring and the use of strategies. The researcher suggests that the highly repetitive use of monitoring for comprehension and the routine use of strategies could lead to skilled readers automating not only the low-level cognitive processes discussed earlier (such as word recognition), but they may also begin to automate some of the processes that might be thought of as higher-level processes too. This would seem to be an aspect worth of investigation.

Stanovich (1980) also proposes an interactional model of reading. In Stanovich's model when deficits in lower level skills such as word recognition cause a delay in the reading process, higher level processes such as predicting based on context have time to offer the reader assistance in resolving the 'word'. However, this increased reliance on high-level processes draws on working memory resources and depletes the availability of working memory to assist with other high-level processes such as making connections between propositions in a text.

Stanovich's explanation is, for the researcher, an extremely plausible explanation of the cognitive processes at work during reading and one which helps to account for individual differences in reading ability. The central role of working memory in facilitating processes has been cited by all the models of both reading and writing discussed above (Chenoweth and Hayes, 2001; Hayes, 1996; Hayes and Flower, 1980; Perfetti and Adlof, 2012). If, as Stanovich suggests, low level processes are weak and need to draw heavily on working memory, then there is unlikely to be sufficient working memory left to assist with higher level processes. This concept could prove central to any model of reading into writing as the number of processes being orchestrated by the reader / writer would seem to be much higher than in reading or writing in isolation.

Rayner et al. (2012) also suggest that interactive models of reading are necessary to account for a reader's ability to arrive at comprehension. Rayner suggests that not only do high and low-level processes operate simultaneously but that at times, processes compete for priority as they are integrated to achieve comprehension. For example, a literal understanding resulting from the parsing of a phrase such as 'she had itchy feet' may be out competed by an idiomatic interpretation of meaning if it was followed by 'and so decided to book a holiday'.

The models discussed above provide a broad understanding of how the processes of reading and writing might fit together to represent the processes of reading into writing. However, in places, they are quite theoretical, devoting less attention describing the way some elements of the models are operationalised when reader (and writers) engage in a task.

A model which is more relevant to this study because it includes an account of how the reader's conscious goals influence the application of reading processes in an academic context is supplied by Khalifa and Weir (2009) (Figure 6). During an academic reading into writing task, the reader is reading for a clear purpose: to supply the content for their essay or assignment. Therefore, the role of reader's goals would seem to be central to the way that reading processes might be operationalised during the task.

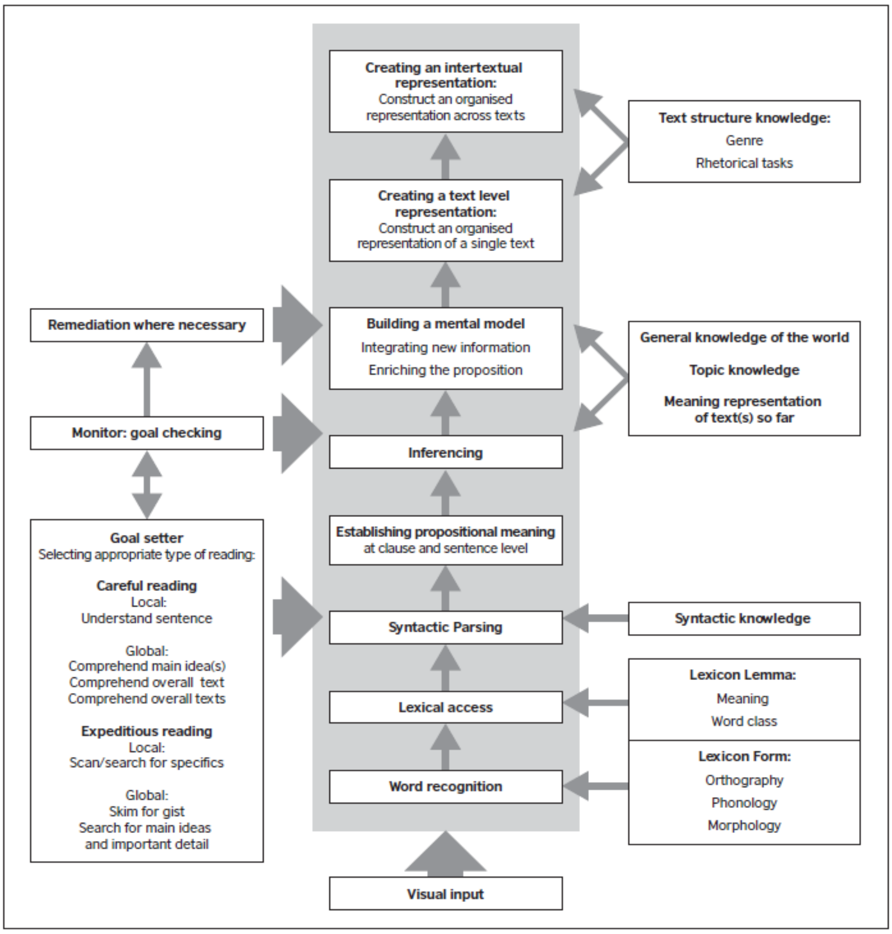



*Figure 6 Khalifa and Weir's model of reading (Khalifa and Weir 2009:43)*

Khalifa and Weir's model of reading suggests that the 'goal setter' (in the left-hand column in Figure 6) determines how the reader will engage with the text. During the section on cognitive processes (2.3.1) the discussion centred around low-level and high-level processes. High-level processes are conscious and effortful whilst low-level processes are automated and largely subconscious. Khalifa and Weir (2009) suggest that readers make decisions about how to read, utilising **different types** of **high-level reading processes** to access the information they need to meet their reading goals.

For example, the reader may not elect to pursue a slow and careful parsing of every sentence in a bid to achieve total comprehension if their goal is merely to assess whether the text is relevant to their current assignment.

Khalifa and Weir suggest that, based on Urquhart and Weir (1998), reading can take the form of careful or expeditious reading. Careful reading 'is intended to extract complete meanings from the presented material' (Khalifa and Weir, 2009:46) and reflects the type of careful reading described by Rayner et al. (2012). The physical characteristics of this type of reading are described in more detail in section 2.4.1 but essentially it involves a slow, serial progress through the text, resolving any difficulties or misunderstandings as they occur. This type of careful reading can take place at local or global level. In the case of careful local reading the reader wishes to extract complete meaning from a single sentence. If the reader continues reading across several sentences, linking together propositions from more than one sentence, the reading is said to be global.

Khalifa and Weir (2009) acknowledge that, at times, careful reading may be more intense when readers are attempting to understand how macro and micro propositions link together. Cohen and Upton (2006:p7) suggest that reading to learn requires readers to…

> *…recognize the organization and purpose of a text, to distinguish major from minor ideas and essential from nonessential information, to conceptualize and organize text information into a mental framework, and to understand rhetorical functions such as cause-effect relationships, compare-contrast relationships, arguments, and so on…*

Grabe and Stoller (2011:p7) also suggest that reading to learn is more demanding than reading to comprehend and is likely to be slower as readers engage in rereading of the text and attempt to remember information, draw inferences and make connections to their existing knowledge. Grabe and Stoller (2011) also propose that reading to integrate information in order to write or critique a text forces the reader to make additional decisions about how different sources support or contradict each other in order to develop a new rhetorical framework incorporating ideas from a variety of sources. Rouet (2006) suggests in the process of integrating information from multiple sources experts consider the validity of the source, corroborating information across sources in addition to relating information in the texts to their prior knowledge.  The researcher suggests that reading as part of reading to write, as with reading to learn, is likely to be an intense, demanding form of careful reading as described by Khalifa and Weir.

Khalifa and Weir suggest that expeditious reading is quick, selective and efficient, guided by the reader's goals. Urquhart and Weir (1998) suggest that expeditious reading can take three forms: skimming, search reading and scanning. Each of which are now discussed in turn.

Skimming, Urquhart and Weir suggest, is when the reader attempts to build a broad understanding (a macro-structure) of the text by reading very selectively, reading the minimum amount of information possible. Grabe and Stoller (2011:p7) suggest that skim reading involves…

> *a combination of strategies for guessing where important information might be located in the text, and then using basic reading comprehension skills on those segments of the text until a general idea is formed.*

It incorporates reading for gist (skimming), where the reader wishes to quickly establish what the main idea or theme of the text is. Skimming or gist reading is necessarily a form of global reading as it must encompass several ideas or propositions distributed across the wider text. Duggan and Payne (2011, also see section 2.8.1) used eye tracking to monitor skim reading. They concluded that when reading for gist, readers start reading a section or paragraph carefully but as the quantity of information about the key point or proposition begins to fall the reader will abandon careful reading and advance to the next page, section or paragraph and start again.

Khalifa and Weir (2009), based on Urquhart and Weir (1998), also describe search reading. They argue that in search reading…

> *the reader is sampling the text, which can be words, topic sentences or important paragraphs, to extract information on a predetermined topic. Khalifa and Weir (2009:57)*

The key difference between reading for gist and search reading, is the reader's purpose or motivation. In reading for gist, the reader is perhaps aiming to sample the text guided by the structure of the text, whereas in search reading the information being sought is on a predetermined topic. Search reading differs from scanning where the reader is trying to locate an exact word or phrase. In search reading the reader will need to consider whether words under consideration have a semantic link to the theme of their search. Search reading can be both local and global as the reader's search could relate to a single proposition, or a theme that incorporates several propositions.

Khalifa and Weir suggest that scanning should always be considered local. Not because the scan for the word is confined to a single sentence but because the item sought (a single word or phrase) operates at a local level.

Grabe and Stoller (2011: 7) also describe a form of reading that they call search reading. However, in their 'reading to search for simple information' the reader is scanning the text for a specific piece of information. This seems to correlate to Khalifa and Weir's (2009) scanning in which the reader does not engage with the meaning of the text but rather engages in a word or text matching exercise. Once the 'match' is located, the reader may then engage in careful reading around the site of the match. To be clear, this research adopts Khalifa and Weir's term scanning to describe this type of word matching behaviour and uses the term search reading to apply to Khalifa and Weir's purposeful hunt for information on a predetermined topic. Having discussed the left-hand column on the Khalifa and Weir model (Figure 6), the central column of processes is now considered.

The central column breaks reading processes down into a series of processes, which start at the bottom with low level processes such as word recognition and lexical access through to high level processes such as building a mental model. The right-hand column lists the type of knowledge that the reader will need to consider for different processes.

Several recent studies have adopted the framework as a basis for their analysis of reading processes; for example, Bax (2013), Brunfaut and McCray (2015) and Wu (2011). For the researcher this model represents a logical and convincing model that is capable of accounting for the types of reading that experienced academic readers utilise when interacting with written sources of information. The only aspect that the model lacks is any reference to the role of working memory in the process. As discussed, earlier in this section, the automation (or lack of automation) of low-level skills is likely to impact on the amount of working memory available to engage in more strategic reading activities and therefore the researcher suggests that some reference to the role of working memory in the model is desirable.

The detailed description of types of high-level reading processes proposed by Khalifa and Weir offers suggests that it may be possible to differentiate the different types of reading utilised by readers in response to their reading goals. Therefore, the researcher proposes to adopt aspects of Khalifa and Weir's different types of reading to categorise the reading behaviour in this study. Having considered the models of writing and models of reading, we review the research suggesting the extent to which reading and writing skills are interrelated.

### 2.3.4 Reading into writing as a separate skill

This section considers whether reading into writing ability reflects a combination of reading skill and writing skill or whether it represents a third, separate skill. Consideration of this matter affects the approach that the researcher adopts when conducting research.

Some researchers have suggested that reading and writing share cognitive processes and both rely, at least in part, on the same knowledge; therefore, better readers make better writers. For example, Stotsky (1983) investigated the correlation between reading and writing ability in participants' first language and concluded that 'better writers tend to be better readers… better writers read more than poor writers and better readers tend to write with greater syntactic maturity than poor readers.' Stotsky (1983:636)

Whilst these findings could be interpreted as evidence of shared knowledge and processing they are far from conclusive. Stotsky's (ibid.) review of research seemed to suggest that development in one skill could be transferred to the other but only when students were directed to reflect on the shared knowledge and processes; the transfer from read to writing (or vice versa) could not be said to be automatic.

Whilst others have endorsed the idea of shared processes. Spivey and King (1989:p7) concluded that…

*general reading ability and success at synthesizing overlap to a great extent, and suggest that success at synthesis may be related to cognitive factors commonly associated with comprehension, such as sensitivity to text structure*

Fitzgerald and Shanahan's (2000:p40) review of reading-writing relations also suggests that readers and writers make use of considerable amounts of shared knowledge which they classified as follows grouped into the following categories:

***Categories of Knowledge That Readers and Writers Use***

*Metaknowledge (Pragmatics)*

*Domain knowledge about substance and content (prior knowledge, content knowledge gained while reading and writing)*

*Knowledge about universal text attributes*

*Procedural knowledge and skill to negotiate reading and writing*

Fitzgerald and Shanahan (ibid) conclude that:

*Various forms of research have supported the theoretical contention that reading and writing rely on analogous mental processes and isomorphic knowledge. However, the total amount of shared variance among a number of reading and writing indicators has never been documented to be more than about .50. Consequently, it is also important to acknowledge the separability of reading and writing.*

The concept of shared or common processes has been supported by other studies. Parodi (2007) conducted tests to assess the extent to which microstructural, macro structural and super structural levels of comprehension and production rely on shared processes and shared knowledge-based strategies. Parodi concluded that his study supplied quantitative data to support the hypothesis of shared basic strategies for reading and writing. However, he went on to suggest that more research of a qualitative nature needs to be conducted to further define the extent and nature of the commonality.

Although the literature above suggests that reading and writing share many basic processes several studies contest the assumption that the skill of reading into writing is simply an amalgamation of an individual's reading skill and an individual's writing skill. Instead they propose that reading into writing is a skill in its own right.

In her 2008 study Asencion-Delaney 'explored the extent to which the reading-to-write construct is the sum of one's reading ability and writing abilities or an independent construct' (2008:140). Asencion-Delaney concluded that reading-to-write was a unique ability only weakly related to reading comprehension skill and completely separate from the ability to write without referring to source material.

Asencion-Delaney's conclusion accords with the work of Cumming, Kantor, Baba, Erdosy, Eouanzoui and James (2005) who found that there were significant differences between texts written for independent writing tasks and reading to write tasks. Cumming et al.'s study compared the writing produced by 36 L2 candidates when writing independent essay tasks, reading into writing tasks and writing in response to a listening. The results indicated that there were significant differences between the discourse that examinees wrote for the independent essays and the integrated reading-writing or listening-writing tasks in respect to:

- Lexical sophistication (in terms of word length and different words produced),
- Syntactic complexity (in terms of words per T-unit and clauses per T-unit),
- Argument structure (in terms of propositions, claims, data, warrants, and oppositions),
- Voice in source evidence (in terms of specifying the self or other sources as evidence), and
- Message in source evidence (in terms of proportions of declarations, paraphrases, and summaries).

Cumming et al (2005:p32) conclude that

*Examinees tended, in the integrated tasks compared to the independent essay, to write briefer compositions, to use longer words, to use a wider range of words, to write longer clauses and more clauses, to write less argumentatively oriented texts, to indicate sources of information other than oneself, and to paraphrase, repeat verbatim, or summarize source information more than to make declarations based on personal knowledge.*

It is evident that theories based on shared processes have their failings, not least because communicative and functional aspects / influences are not taken into account sufficiently and these may play a strong role in reading into writing. Therefore, it seems reasonable to conclude that whilst reading and writing may share some basic processes these explanations lack a level of sophistication that helpfully distinguishes between writing without reference to a source and reading into writing. This conclusion inevitably leads us to ask how reading into writing can be defined in terms of cognitive processes.

### 2.3.5 The cognitive processes of reading into writing

Some researchers have made attempts to elaborate upon the reading into writing skill and describe the exact processes involved in writing based on reading. Seveg-Miller (2007) uses the term 'discourse synthesis' to refer to writing from sources. Seveg-Miller goes on to point out that discourse synthesis is more cognitively demanding than summarising a single text. Kintsch and van Dijk's (1978) work on summary writing has been widely quoted and proposes that summary writing involves three major operations.

1. Deleting redundant ideas
2. Substituting a series of ideas with one more general / overarching idea
3. Selecting or constructing a macro proposition for the summary

Segev-Miller (2007:p232) argues that

*The discourse synthesis task is similar to the summary, but it is cognitively more* *demanding: when synthesizing students are required to construct their own macroproposition, or rather 'superproposition' (henceforth the Macroproposition), from different or even sometimes contradictory propositions and macropropositions of multiple source texts, and to organize these in a previously non-existent conceptual structure. Conceptual restructuring or transforming, therefore, requires a higher order intertextual processing (Flower, 1989), and the production of personal and creative perspectives on the part of students (Schumacher & Gradwohl, 1991).*

Segev-Miller used weekly interviews, think-aloud protocols and journals to follow 12 undergraduates studying to be primary school teachers over two semesters as they completed their literature reviews. From this data Segev-Miller produced an in-depth taxonomy of the strategies used by her participants whilst writing their literature reviews. The taxonomy included over 90 strategies, too many to list here, which were divided into planning, evaluating and executing strategies.

Whilst this taxonomy is detailed and comprehensive, it is lacking in other key aspects. Firstly, it is not always evident how some of the strategies manifest themselves. In other words, it is difficult to know exactly what writers engaged in one of the strategies will actually be doing or thinking and thus how it might be identified or measured. Secondly, it contains so many strategies that it seems unlikely that every writer will use all of them in every discourse synthesis task. This therefore begs the question as to which strategies will be used when and why. For the researcher these issues are central. If we are to assist students to improve the quality of their written academic assignments (and subsequently their longer term academic outcomes) then a clear understanding of the process of academic reading into writing is essential to develop interventions likely to bring about an improvement.

Any model of reading into writing must necessarily be complex. Grabe (2001) suggests that, at a minimum, any theory of reading into writing needs to account for a theory of reading, a theory of writing and a theory of learning which shows how the two skills interact. Other factors, Grabe suggests, are likely to make the picture more complex still. These factors include such issues as:

- Directional influence e.g. reading to improve writing, writing to improve reading and combined reading and writing to improve learning.
- The wide variety of contexts in which studies have been carried e.g. primary / secondary education, first and second language learners and levels of proficiency.
- Motivational factors.

Grabe goes on to say that therefore, in addition to the theories of reading, writing and learning, some theoretical assumptions must also be made about theories of language, language processing, motivation and affective factors, the influence of social context and finally the theory of the role of background knowledge in reading and writing.

Stein (1989) used think-aloud methodology to identify four cognitive processes that 32 out of the 36 writers in her study engaged in when reading to inform writing whilst completing an academic university task. These processes where defined as:

**Monitoring** – Students check their understanding of what they are reading and how it relates to the demands of the task.

**Elaborating** – students consider the new knowledge they are gathering from the text in light of their existing knowledge and combine them to create new ideas and reflect critically on the ideas introduced by the text.

**Structuring** – this is when students begin to reorganise the ideas they have read, maybe by categorising them or prioritising them or looking for links / themes which may not have been explicit in the text.

**Planning** – this stage is when students decide on the overall structure of their text by referring to organising ideas. Experienced writers do much more planning than inexperienced writers.

However, the most complete account of the processes of reading into writing in the literature is suggested by Chan (2013). The work of Chan makes a significant contribution to research by reviewing a range of models of writing as well as key models of reading in order to suggest the following breakdown of cognitive processes at work during reading into writing in an academic context. See Table 1.

This study adopts Chan’s (2013) framework as the basis for this study. Therefore, each of Chan's proposed processes will be discussed in turn below.

*Table 1 Chan's (2013:73) cognitive processes in reading into writing*

| Cognitive processes | Working definitions |
|---|---|
| Task representation | Create an initial understanding of the task (e.g. the overall purpose of the test/assignment, structure of the test, time constraints, scoring criteria, word length, topic, genre and intended reader, rhetorical functions to perform) |
| Macro-planning | Plan for writing goals, content and organisation of the text, etc. Identify major constraints (genre, readership, language resources, etc.) |
| Higher-level reading | Careful reading to create textual and / or intertextual representations)<br>Search reading (e.g. select ideas which are relevant to the task context to put in the new text from the source texts based on a set of criteria perceived as appropriate) |
| Organising | Organise the ideas to put in the next text (e.g. prioritize ideas in terms of relevance or importance, re-order, re-combine, delete, categorise, create new structure, etc.) |
| Connecting and generating | Generate links between ideas or new meaning by connecting ideas/discourse features provided in the source texts with their own knowledge. |
| Micro-planning | Plan for the part of the text that is about to be produced |
| Translating | Translate abstract ideas into linguistic forms |
| Monitoring and Revising | Higher-level: meaning and coherence<br>Lower-level: accuracy or range of grammar, vocabulary and sentence structure, plagiarism |

#### 2.3.5.1 Task representation

Flower (1990:35) defines task representation as ‘an interpretive process that translates the rhetorical situation - as the writer reads it - into the act of composing.’

Plakans (2010) breaks task representation down into initial task representation, topic determination, genre identification and source text use. However, Chan (2013) suggests a more detailed breakdown drawing on the work of Flower et al. (1990), Grabe and Kaplan (1996), Ruiz-Funes (2001) Plakans (2008, 2010) and Scardamalia and Paris (1985). Chan suggests that task representation determines the overall ‘shape’ and ‘feel’ of the participant's answer. Some of the key task representation decisions made by the participant include:

**What to include**- whether to include only ideas from the source text or whether to add previous knowledge, comments or opinions. Whether to include all the ideas from the text or be selective.

**What type of format to use**- whether there is a standard format or style they are expected to follow.

**How to organise the ideas**- should there be an overarching idea / theme, should the ideas from the source text be summarised and presented in the order they appear in the source text or reordered for the writer's purpose? Do participants even consider that they have any choice in this matter or should they follow a predetermined order?

**The use of strategies by the participant when answering the question** – So, for example, does the writer use the source texts as a springboard to generate their own ideas or to illicit their own response? Does the writer read looking for themes / links? Does the writer tackle the task in a particular order?

**What other goals need to be met** - What other 'goals' influence the participant's approach? Does the writer want to get finished as quickly as possible? Does the writer want to impress the reader, or does he / she want to learn or improve their skills by doing the task?

Kantz (1990:76) states that 'task representations matter because they affect the written product'. However, Kantz's study found that the link between task interpretation and the finished product is not simple. Although students' initial thoughts and opening paragraphs seem to indicate one approach their final products did not always match the approach outlined. This led to confusion and differences of opinion amongst the lecturers grading the papers. This finding was echoed by Segev-Miller (2007:244). When following students working on a literature review over several months Segev-Miller concluded ...

*analysis of the processes underlying the products in the present study indicated that* *the connections between the subjects' task representations and the structures of their products were not that simple or direct.*

Whilst many students started with complex task representations which looked likely to result in knowledge transformation some students reverted to a knowledge telling approach when they encountered difficulties. Conversely, the initial knowledge-telling representations of some of the subjects during the reading process evolved to produce knowledge-transformation in the finished text.

Smeets and Solé (2008) compared post-graduate students' thoughts about task interpretation on a synthesis course work task with their finished texts. The task was completed in class and students' attitudes were surveyed after they had read the task instructions but before they started reading and writing. Students were asked to report their task representation attitude by selecting the statement which best described their understanding of the task from a choice of six statements. Half the statements described a knowledge telling approach and half described a knowledge transforming approach. Additionally, some of the statements from each of the knowledge telling / transforming groups described a text -by-text approach whilst others described an intertextual approach. This creates four categories: knowledge transforming with intertextual approach; knowledge transforming with a text-by-text approach; knowledge telling with intertextual approach; and knowledge telling with a text-by-text approach. Some categories contained two statements whilst others only contained one. Finished products were scored according to the amount of elaboration that had taken place and the extent to which a new macro-proposition had been generated.

Smeets and Solé concluded that task representation did influence the quality of the finished product. These findings do not necessarily contradict those of Segev-Miller and Kantz as Smeets and Solé's study represented a snap-shot in time. Perhaps given an extended period of time, some of those with a knowledge telling approach may have gone on to develop a knowledge transforming approach. Viewing the results of all three pieces of research, one might conclude that developing an understanding that knowledge transformation is a task requirement can take time and is unlikely to happen on tasks completed in a single session. Additionally, even when the task representation includes an intention to include knowledge transformation, the difficulty of achieving knowledge transformation may mean that the finished product fails to meet the aspirations of the task representation.

Allen (2004) carried out a case study of a Japanese student working in English on reading to write course work tasks over one semester. The student was a third-year undergraduate and Allen focused on task representation and integration of source materials. Allen concluded that 'additional training may be needed to assist students to develop the ability to represent tasks and integrate source text material in a way that most effectively supports their text.' (2004:87)

In accordance with Kantz, Smeets and Solé, and Allen, the researcher suggests that task representations evolve as the writer's work emerges. Therefore, it seems reasonable to conclude that task representation cannot be deemed to be something that happens once at the outset of a task. Given enough time, writers revisit task representation continually, sometimes leading to an improvement in their finished product, sometimes abandoning their ideal task representation in the face of difficulties. Therefore, task representation is a complex and sometimes time-consuming process and one which does not guarantee the quality of the finished product.

### 2.3.5.2 Macro-planning

Field (2004) suggests that, drawing upon world-knowledge, the writer decides upon the goals of writing including consideration of the genre and level of formality.

Flower (1990) suggests that many students' planning for writing in reading to write tasks is heavily reliant on one of two strategies, both of which have their basis in reading the texts. Flower suggests that using a 'gist and list' approach (1990:235) the student identifies the main idea which links the texts and adopts it as the macro plan for their text. Flower identifies 'True, Important, I Agree strategy -or TIA for short' (ibid) as another common approach. The student uses agreement (or lack of it) as the macro-plan for their writing. However, Flower points out that both these approaches result in knowledge telling, not transformation. In order to transform knowledge Flower suggested that students need to go beyond the aforementioned strategies and engage in constructive planning. Flower's constructive planning resulted in students evaluating what the lecturers were looking for, consciously trying to work out a new approach or angle of their own and actively monitoring their progress against their plan. The researcher suggests that the acceptance of a 'gist and list' or 'TIA'. or the rejection of them in favour of a novel approach equates to Chan's (2013) Macro- Planning stage.

Segev-Miller (2007), echoing Flower's findings, found that both successful and unsuccessful students constructed a macro proposition when engaged in discourse synthesis. However, Segev-Miller suggests that whilst successful students pursued the cognitively demanding path of developing their own, original macro proposition less successful students resorted to the less demanding tactic of adopting propositions or a macro proposition from one of the texts and using it as their own macro proposition.

Kellogg (2006) emphasises that skilled writers are able to shape their text to accommodate the way the reader is likely to interpret the text. Skilled writers do not produce a text to suit themselves, they produce a text that will suit the intended readership. This acute awareness of audience is reflected in the goals of the writer and affects the writer's decisions regarding genre and formality.

Shaw and Weir (2007) point to 'the relevance and adequacy of content to the task set, the appropriateness of the language used for the topic' as well as the effect on the reader as being indicators of the presence or absence of macro-planning.

There appears to be some overlap between task representation and macro planning in that both refer to goal setting and consideration of the intended reader. However, task representation seems to focus on interpreting the question rubric whilst macro planning focuses on how the writer plans to adapt their writing to meet the perceived demands of the task.

#### 2.3.5.3 Different types of reading

Although it is easy to observe when, and for how long, someone engaged in a reading into writing task reads the source texts, Chan suggests that this information is far less important than which reading processes have been used. Chan's framework specifically identifies higher level reading processes as being required during reading into writing tasks.

Chan proposes that the model of reading outlined by Khalifa and Weir (2009) is useful when considering reading as part of reading into writing. Khalifa and Weir's model of reading (see section 2.3) breaks reading into eight sub processes. These range from low level processes such as word recognition through to high level processes such as creating an intertextual interpretation. Khalifa and Weir's model (2009:43) lists the sub processes as follows:

**creating an intertextual interpretation***

**creating a text level representation***

**building a mental model***

**inferencing***

**establishing propositional meaning**

**syntactic parsing**

**lexical access**

**word recognition** *(* Higher level processes)*

Khalifa and Weir's (2009) description of reading identifies two very different types of reading: careful and expeditious, both of which can be applied at global and local levels as discussed in section 2.3

Chan reports that '(c)areful reading involves comprehension of every part of the whole text while expeditious reading means processing texts selectively, quickly and efficiently to access desired information from a text'. Khalifa and Weir (ibid.) suggest that comprehension of a text involves not just decoding sentences to understand the meaning of individual sentences, the *micro-structures*, but also understanding the relationship between the main ideas contained in the text, the *macro-structure* of the text.

Khalifa and Weir suggest that **careful global** reading is used to handle the majority of the text with readers reading each sentence slowly and sequentially, building up a detailed understanding of the text. **Careful local** reading occurs when, as with careful global reading, the reader is attempting to access a complete understanding of the meaning, but the bout of reading is confined to a single sentence.

With Khalifa and Weir's description of **expeditious global** reading the purpose is to skim read the content to get the gist or to search read for the main ideas. The reader will be quick, selective and regularly skip chunks of text, indicating strategic reading. **Expeditious local** reading is also quick and selective but here the reader is searching to locate a single word, fact or figure.

According to Khalifa and Weir's descriptions of reading it is possible to surmise that **careful global** reading and **expeditious global** reading both represent higher level reading skills as both are likely to contribute to the reader arriving at an understanding of the whole text(s) and how the text(s) relate to the writing task. It is difficult to see how Khalifa and Weir's locally focused reading activities can be seen to represent higher level reading skills. Expeditious local reading in particular, which relates to searching for a single piece of information, seems unlikely to contribute to overall understanding of the text. The only exception is perhaps the occurrence of inferencing when careful local reading is used to process sentences containing ambiguity. The researcher would argue that careful local reading may be used by a reader whilst re-reading an ambiguous sentence in order to try and resolve meaning.

Grabe and Stoller (2011) suggest that fluent readers initially form a **text model of reading comprehension**. That text model is the basis for a more elaborate interpretation of the text which Grabe and Stoller refer to as the **situation model of reader interpretation**. The creation of these two models relies upon two high level reading processes.

Lower level reading processes generate clause level meaning units which are linked together in the form of a network using higher level processes. The ideas and themes which allow new meaning units to be added to the network are reinforced and become central to the text model. Ideas which are not referred to again or provide no link to new ideas / themes are likely to be erased or forgotten. Thus, the text model becomes the reader's internal summary of the main ideas. Grabe and Stoller report that the coordination of ideas, which facilitates the building of the text model, is the most fundamental of the two higher level reading processes.

As the reader is building the text model they are simultaneously elaborating on this text model to build a more sophisticated situational model. The situational model is influenced by the reader’s goals, feelings and back ground expectations. Grabe and Stoller suggest that in the process of forming this situational model the fluent reader engages in the other high-level process of elaboration. This elaboration relies on making inferences, drawing on background knowledge, monitoring comprehension, forming attitudes about the text and author, adjusting goals and critically evaluating the information being read.

Grabe and Stoller suggest that these dual models of the text are what allows the reader to interpret a text according to what the reader thinks the writer is trying to say <u>and</u> according to the reader’s own purpose for reading. The writer suggests that creating this situational model could be seen as central to the connecting and generating process identified by Chan and described further in the next section.

#### 2.3.5.4 Connecting and generating

Chan (2013:67) proposes that ‘(c)onnecting is a process in which writers generate links between ideas or new meaning by connecting ideas in the source texts with their own knowledge’ based on the work of Spivey (1984,1990,1997).

Kintsch and van Dijk (1978) argue that for texts to be comprehensible they require both a microstructure and a macrostructure. Individual propositions (sentences / phrases) need to be semantically linked to those around them in order to create coherence at a local level. Texts which contain more than one idea also require a macrostructure or global organising idea to link the various microstructures together in order to remain coherent. Rather like Grabe and Stoller's text model discussed in the previous section, Kintsch and van Dijk propose that readers review the ideas from the text, deleting ideas which are not central to understanding the text, generalising or grouping together similar ideas, and gradually linking all the ideas together in a linear or hierarchical pattern called a text base. The construction of the text base relies on making connections between explicit information stated in the text and some implicit knowledge from the reader's memory.

However, for the purposes of this study we are particularly interested in how multiple texts are connected together in the mind of the reader.

Rouet (2006) suggests that when readers have to integrate multiple sources of information they are confronted by three additional problems compared to when working from a single text. Firstly, there is no coherence between the texts. There have not been written or designed to fit neatly together into one text base or macrostructure. Secondly, there may be differences, discrepancies or even contradictions between the texts. Lastly, when writing in response to a complex task there may be no correspondence between the content of the documents and the task, unlike comprehension questions which are designed to 'fit' the text they accompany. Rouet suggests that, because of these problems, those engaged in reading into writing have the additional burden of evaluating the worth of each source, selecting the ideas which are relevant and shaping them into a coherent whole.

Rouet (2006) suggests that when comprehending multiple documents readers construct a *documents model.* This model is made up of a *source model* and a *situations model.*

**The source model** contains information typically found in references and relates to where the texts came from, who wrote them and when. In addition, information about the type of document, the setting and context of the document are included. All of these factors are used by experienced readers to make judgements about the reliability of the source. Not all of the information in the source model is explicit, much of it will be based on the reader's prior knowledge about the author's motivation for writing, the prevailing culture or attitudes at the time and in the place where the document was written.

**The situations model** contains information about the contents of the various documents, including details such as main ideas and propositions.

Links are made not only linking source details to contents (who said what) but also between contents (whether the contents of one document support or contradict another) and between sources (this source is more reliable than that one). The source to source links and understanding why one writer may be motivated to propose one interpretation of events and another writer proposes a quite different interpretation is important in allowing readers to reconcile conflicting accounts / arguments.

Perhaps the complex web of links proposed by Rouet goes some way to explaining the findings of Britt and Sommer (2004). Britt and Sommer proposed the Restructuring Hypothesis. This hypothesis suggests that when students were given a task of constructing 'a well-structured initial representation of a text prior to reading a subsequent text' this task 'will aid in the between-text integration process'. This concurs with the work of Gil, Braten, Vidal-Abarca and Strømsø (2010) who concluded that writing summary essays led to better integration and understanding than writing argument essays.

Indeed, Cerdan and Vidal-Abarca (2008) concluded that the reading behaviour of students answering intra-text questions (where answers could be found within a single text) differed significantly from students asked to write an essay which demanded integration of information across multiple texts. They also concluded that, whilst cognitively more demanding, the integration of information across texts led to deeper understanding of the information contained in the texts.

Goldman (2004) suggests that for expert readers generating intertextual links is part of normal reading behaviour. Goldman suggests that experts tend to 'make cross text comparisons to corroborate information, pay attention to the source of the research' and engage in strategic reading behaviour.

Therefore, when readers become writers attempting to meet the demands of a task, they have to engage in many more cognitively demanding processes than simply constructing a text base. The additional burden of evaluating sources, making links between texts, resolving contradictions and deliberating what each text offers in terms of meeting the demands of the task, places heavy demands on the working memory and cognitive resources of the reader / writer. Generation of a 'superordinate' macro structure, that draws upon all the sources in response to the task, demands knowledge transformation. In the process of generating a new 'superordinate' macrostructure the reader / writer will need to decide how to order and organise their ideas. This will be discussed next.

#### 2.3.5.5 Organising

Spivey and King's (1989:11) definition of discourse synthesis includes three essential sub-processes, of which organising is the first.

> **Organising** - writers think about the overall structure of their writing and the structure of the source texts.
>
> **Selecting** - writers read and select ideas from the source texts.

**Connecting** - both linking ideas from the source texts with their own ideas and linking together ideas contained in their writing.

Based on the work of Spivey and King, Plakans (2009) suggests that organising, selecting, and connecting abilities should be considered as the cornerstones of the construct of academic writing.

Flower (1990) in her chapter Negotiating Academic Discourse suggests that for knowledge transformation to take place students need to move beyond listing the points in the source texts or giving their opinions about the ideas in the source text. Instead she suggests that students must view the source text through the lens of the task / question rubric in order to shed new light or give an alternative interpretation to what is said in the texts. In this sense, reorganising the ideas from the source text and prioritising seem to be critical to knowledge transformation.

Alamargot and Chanquoy (2001) appear to concur with this view when in Chapter 1, section 3.2.2 (no page number as digital edition of book) they suggest that organising is a problem-solving operation that allows writers to create 'new semantic relationships between activated knowledge' and is one of the central characteristics of Bereiter and Scardamalia's Knowledge Transformation.

Plakans (2009) lists arranging essay content, identifying rhetorical structures and summarising source texts as actions which contribute to organising. Several writers have suggested that good organising ability can be linked to the quality of the finished product. Kellogg (1994) concludes that when writers produce a more thorough plan, in the form of an outline, it enables writers to organise their ideas better prior to writing and results in a better final product. Chai (2006) also found that students whose writing plans showed a greater degree of organisation of idea units tended to produce better essays.

Deane et al. (2008) proposed that sophisticated organisation of ideas could be linked to familiarity with a topic domain. They suggested that familiarity with topic domains often included familiarity with organising models which enabled students to reduce the demand on their working memories.

Swales and Lindemann's (2002) findings appear to concur with the hypothesis proposed by Dean et al. Swales and Lindemann found that when students were asked to represent their ideas for writing as diagrams, the structure of the diagrams varied according to the student's discipline.

Parodi (2007) suggests that a lack of automaticity in lower level reading and writing skills leaves little working memory available for students to retain ideas across paragraphs and thus both perceive (when reading) and produce (when writing) a list of ideas rather than organising ideas into a more complex macro structure.

Chan bases her definition of organising on the work of Field (2004) and Shaw and Weir (2007). Field (2004:329) suggests that organising is a phase where the writer ''provisionally organises the ideas, still in abstract form, in relation to the text as a whole and in relation to each other'. Shaw and Weir (200:38) suggest that organising is used to 'determine which [ideas] are central to the goals of the text and which are of secondary importance'.

The researcher suggests that writers who devote time and pay attention to a more sophisticated organisation of their ideas are more likely to produce a product that represents knowledge transformation rather than knowledge telling.

#### 2.3.5.6 Micro-planning

Field (2004:329) suggests that decisions about the content of the next sentence or paragraph are influenced by the macro-plan, the text so far (and to what extent it has met the goals outlined in the macro-plan) and 'whether an individual piece of information is or is not shared with the reader' based on what has been said in the text so far or based on shared common knowledge. Shaw and Weir (2007) suggest that Field's model has advantages over the model proposed by Grabe and Kaplan (1996) in which the text so far has little impact.

It should be stressed that in Field's Micro-planning stage the propositions are still in the form of abstract ideas and have not yet been converted into linguistic form. The conversion to linguistic form takes place during the Translation phase which is discussed in the next section.

#### 2.3.5.7 Translating

Field (2004:329) describes translating as the process by which '(t)he propositional content that has been assembled undergoes a process of conversion from abstract to linguistic form'. Field suggests that the translating process results in the writer having a phonological version of the sentence in their head so that as the first part of the sentence is being written the writer knows how they will complete the sentence. The reality is that this process is so fleeting and subconscious that it would be very difficult to elicit any information from participants about the process as it usually escapes our notice. Therefore, this study will not attempt to investigate the translation process.

### 2.3.5.8 Monitoring and revising

Chan (2013) suggests that monitoring is when the writer checks the quality of the text produced. When the writer is dissatisfied with the text they embark on an episode of revision. Chan sub-divides monitoring and revising into high-level and low-level monitoring and revising. Low-level revisions target spelling, accuracy of grammar, vocabulary and sentence structure, whereas high-level revisions target meaning and coherence. The differences between Chan's (2013) high and low-level monitoring seem to reflect the differences between 'revising' and 'editing' outlined by Field (2003). Field (2003:117) suggests that revisions are made 'to the form of the text (for example, spelling corrections), while editing may involve rethinking decisions made at the formulation stage'.

The Chenoweth and Hayes model of writing (2001) breaks writing into four components one of which is termed the **reviser**. The reviser is responsible for monitoring both emerging language which has not yet been transcribed and modifying existing text. When problems are identified, changes are initiated.

Hayes, Flower, Schriver, Stratman, Carey (1987) break the revision process down into a series of sub-processes. In each of these sub-processes they explain the differences between expert and novice writers. In terms of **task definition** of what revision is or what revision entails, experts perceive revision as a whole text task. In the **evaluation** stage, expert writers consider their plans / notes as suitable targets for revision. **Problem representation** is portrayed as a scale with detection at one end and diagnosis at the other. Problem representation refers to the writer's ability to understand and analyse the nature of the problem. **Detection** is when a writer spots there is a problem but has little information regarding how to remedy it. Problem detection is persistently difficult for novice writers but particularly at global level. At the **diagnosis** end of the scale, not only does the writer spot the problem but they have a clear procedure for remediating the text. Expert writers are more likely to accurately diagnose problems and have more elaborate plans for resolving them. Finally, expert writers are more efficient in their selection of **strategies** (ignore, delay, search for example) than novice writers.

Kellogg (1996), Field (2004) and Shaw and Weir (2007) suggest that monitoring and revising are extremely cognitively demanding. This idea was singled out for attention by Quinlan, Loncke, Leijten and Van Waes (2012) in their study investigating the role of the monitor. Quinlan et al. presented participants with incomplete sentences which contained an error and asked participants to correct the error and complete the sentence. The working memory of participants was put under pressure by asking participants to listen to words which they had to remember and incorporate into the sentence. Cognitive load was altered by increasing the number of words that participants were asked to listen to and remember (between one and three words.) Quinlan concluded that greater cognitive load (i.e. remembering more words limited the availability of working memory) reduced the ability of participants to successfully complete the task.

Whilst it might seem reasonable to conclude that when writing in a second language the demands placed on working memory might lead to more local level revisions at the expense of global revisions, the work of Stevenson, Schoonen and Glopper (2006) suggest that this is not the case. In their work comparing L1 and L2 revising they concluded that the proportion of high-level revisions compared to low level revisions did not alter significantly in the L2 condition compared to the L1 condition. This could imply that expert L1 writers transfer the processes and strategies which they use in L1 revision to their L2 writing.

Whilst according to Chenoweth and Hayes (2001) revision can be seen as a process that can occur during the translation process, before the emerging language has been typed or transcribed, such behaviours cannot be observed and their fleeting, temporary nature (much like the translation process) make them difficult for participants to recall. For these reasons, revision prior to transcription will be excluded from this study.

In conclusion the researcher suggests that reading into writing for academic tasks is an extremely complex process involving many different cognitive processes, some of which are more conscious on the part of the student than others. It is beyond the scope of a single study to investigate all these processes in depth; however, the way students interact with the source texts is central to the process. This is also an area where eye tracking methodology offers the opportunity to generate quantitative data which could then be interpreted in light of more qualitative methods such as think-aloud protocols. Therefore section 2.4 reviews the literature more generally in relation to eye tracking reading before 2.5 discusses in detail how Khalifa and Weir's model of reading could be used as a framework to interpret eye tracking data.

## 2.4 Eye tracking reading

The research into eye-movements in reading broadly falls into two categories. Research that attempts to throw light on the fine detail of how the brain controls the movements of the eye during reading (for example the SWIFT model by Engbert, Nuthmann, Richter and Kliegl, 2005; E-Z Reader by Reichle, Pollatsek, Fisher and Rayner, 1998) and research that uses eye tracking as an indicator of the cognitive processes underlying reading.

The need for the brain to execute such precise control of the eye is due to the limited area in the human eye which can decipher visual input with great accuracy. Holmqvist et al. (2011) explain that when light enters the front of the eye, via the pupil, after passing through the lens it is projected to an area at the back of the eyeball called the retina. The light detecting cells in the retina are not evenly distributed. Instead, in a tiny area called the fovea the light detecting cells are extremely densely packed. Only the two degrees of the visual field which fall on the fovea are clearly focused. The remainder of the visual field, which falls on areas of the retina where the cells are more sparsely distributed, is slightly blurred or out of focus. Thus, when reading, to be able to recognise a word, the eye must move to ensure that the light bouncing off the word enters the eye and falls directly on the fovea. The two degree highly sensitive field of vision equates to an area approximately the size of a thumb print at arm's length. These metrics are useful when deciding on the font size and line spacing to be used when presenting on-screen reading material to participants. The text needs to be sized and spaced so that it can be clearly distinguished when participants' focus moves from one line of text to the line above or below.

The research into the fine detail of how the brain controls reading eye-movements is not significant for this study. Of much greater significance is the research that uses eye tracking as an indicator of the cognitive processes underlying reading and which will now be discussed.

When focusing on the cognitive processes of reading, many studies have used eye tracking to study reading at sentence level and below, for example, syntactic parsing, lexical access and word recognition. These studies are of value to this study because they provide detailed descriptions of the eye-movement characteristics of careful reading. Studies which have used eye tracking to explore the higher-level processes involved in global text processing are few. Such studies are also useful because they describe more holistic patterns of eye-movements that characterise different types of reading. The next section will start by reviewing the literature related to characterising careful reading followed by a section which discusses the few studies relating to eye tracking global text processing.

### 2.4.1 Eye-movement characteristics of different types of reading

As mentioned above the overwhelming majority of research which has used eye tracking to study the cognitive processes of reading, much of it summarised in Rayner et al. (2012:p377) , has studied the type of reading involved in the…

> *careful processing of written material…in terms of someone carefully reading a text book or a newspaper article or a novel (which you must read carefully in order to pay attention to the plot).*

This type of research has tended to limit description of reading to sentence level, assuming that the reading of every sentence will function in much the same way as the previous, or indeed the next sentence. This large body of literature provides an extremely valuable starting point, offering the detailed description of careful reading. The following paragraphs describe the literature on careful reading and explain how the literature might guide the categorisation and labelling of fixations.

When we read, whilst we may think that our eyes glide smoothly along the line of text that is not the case. In fact, the eyes make a series of jumps along the line interspersed with pauses (Holmqvist et al., 2011; Rayner, Juhasz and Pollatsek; 2005; Rayner et al., 2012). The jumps are called saccades and the pauses are called fixations. During fixations, which last typically last about 250 milliseconds, although the range can vary from just 50 milliseconds to 550 milliseconds (Rayner, Juhasz and Pollatsek, 2005:81), the eye remains relatively still. The eye is not absolutely still but rather trembling, making very tiny movements (micro-saccades) to keep the focus of the eye at the same location. The eye then moves extremely quickly during the saccade to the next fixation. Typically, saccades last just 40 milliseconds (Rayner et al., ibid). Wolverton and Zola (1983) demonstrated that during saccades no visual information is registered; the eye is moving too quickly. On average, reading saccades move forward 8-

character spaces (Holmqvist et al., 2011; Rayner, Juhasz and Pollatsek; 2005; Rayner et al., 2012). Not every word is fixated, with high frequency words and predictable words more likely to be skipped (Blanchard, Pollatsek and Rayner; 1989; Brysbaert and Vitu, 1998).

For the purposes of analysis in this study, the researcher aims to distinguish between careful reading and selective types of reading. One of the few eye-movement studies relating to selective forms of reading was conducted by Duggan and Payne (2011) and related to skim reading. Duggan and Payne (2011) propose that readers engage in 'satisficing' when they skim read. This involves reading the start of a page, section or paragraph carefully, but as the quantity of information about the key point or proposition begins to fall (the reader detects a diminishing rate of information in return for reading effort), the reader will abandon careful reading and advance to the next page, section or paragraph and start again. In other words, the decision of when to move to the next section is determined by whether the current section of text is satisfying the reader's need for relevant information. Duggan and Payne prompted participants to engage in skimming by giving their participants a task which required them to read more text than they could carefully read in the time allowed. They used analysis of the distribution of fixations across a document rather than patterns formed by fixations as the basis of their research.

Some researchers have attempted to use eye-movement patterns to detect reading and have published reading detection algorithms which identify gaze activity that suggests reading. However, it should be understood that it is not possible to identify the cognitive processes undertaken by participants from eye tracking data. It is only possible to state that eye-movements fitting a certain pattern *suggest* that reading is taking place. The following section reviews the literature on algorithms for detecting reading from eye tracking data.

### 2.4.2 Algorithms for detecting reading from eye tracking data

The reading detection algorithms published to date use one of three different methods to identify patterns. Some, Campbell and Maglio (2001), consider the direction and distance travelled between fixations (sometimes referred to as saccadic amplitude) whilst others, Kollmorgen and Holmqvist (2009) and Simola, Salojarvi and Kojo (2008) use hidden Markvok models to analyse eye tracking data generated by reading to detect data patterns. The research data is then examined for similar patterns of data in order to label sections of data which exhibit the same characteristics. A third type, Biedert et al. (2012) uses the speed of the movement between fixations to indicate the type of reading (based on the fact that the eye moves much faster over long, scanning type saccades).

The researcher experimented with those algorithms where the fine detail of their workings could be ascertained from the literature and found that the approach used by Campbell and Maglio (2001) could be successfully adapted to suit the demands of this project. The following sections describe the characteristics of careful reading that are described in the literature and consequently were incorporated into the researcher's own method of detecting careful reading.

To distinguish between careful reading and other types of selective reading it was not enough to consider whether fixations represented a move forward through the text or represented a regression. The researcher needed to distinguish between forward moving fixations which might represent careful reading and fixations which might represent other types of selective reading behaviour. The researcher decided to label fixations which occurred on the same line and at a distance of between 1 and 16 characters to the right of the previous fixation as short forward fixations. This decision was based on two factors. Firstly 16 character represents double the average forward saccade reported in the literature and secondly, this

16-character limit accounts for 95 per cent of the saccades reported by Rayner et al. (2012:95) of eight college-age readers.

Fixations which move from near the end of one line, to the beginning of the line below are referred to as Return Sweeps (Rayner et al. 2012:91). Although different from a short forward in terms of the physical movement, Return Sweeps are, in essence, a short forward movement from the end of one line to the beginning of the next. Fixations which represented a move to the line below in conjunction with a long to the left (more than 50 characters which represented over half a line) were labelled as Return Sweeps.

The only remaining type of forward moving fixations (once we have excluded short forward and Return Sweeps) the researcher labelled as long forward. Simola et al. (2008:5) reported that when participants changed from 'rauding' (which they define as normal reading in which the reader is looking at each consecutive word of a text to comprehend the content) to 'skimming' the length of forward saccades increases. For this reason, long (more than 16 characters) forward moving saccades are unlikely to form part of careful reading.

During reading the eyes do not move relentlessly forward through the text. On average 10-15 per cent of fixations are a return to an area of text already read (Holmqvist et al., 2011; Rayner, Juhasz and Pollatsek, 2005; Rayner et al., 2012). These backward movements are called regressions. Rayner et al. (2005) suggest that short regressions within the current sentence represent word recognition problems, whilst longer regressions back to previous sentences are likely to represent comprehension difficulties. Rayner (2009) reports that such short, sentence level regressions account for the majority of regressions. Holmqvist et al. (2011) showed that the number of regressions made decreased as a function of improved reading skill. For the purpose of analysis, any fixation which moved to an earlier part of the text, whether within the sentence or to an earlier sentence was labelled as a regression. Although the literature reported above suggests that regressive fixations account for 10-15 per cent of all fixations, the researcher was surprised to note a much higher proportion of regressive fixations in the data collected. This is discussed later in chapter four.

Unsurprisingly, the literature on reading fixations reports that better readers have, on average, shorter fixations and make longer saccades than less skilled readers (Rayner, 1998). Word frequency has also been shown to influence fixation duration with fixation duration increasing as words become less familiar. Inhoff and Rayner (1986) showed that even after controlling for word length (infrequent words tend to contain more characters than frequent words), infrequent words were fixated for longer than frequently occurring words.

More recent research has moved on to consider how reading skill and other factors such as word frequency and predictability may interact. Ashby, Rayner and Clifton (2005) conducted research into the effects of both word frequency and word predictability in relation to reading skill. They concluded that the low predictability of a word does interact with reading skill, extending the fixation times of average readers disproportionately compared to the increase in fixation times of skilled readers. The same research was less clear cut in relation to how skill and word frequency interact. However, Kuperman and Van Dyke (2011) demonstrated that differences in fixation durations between better readers and poorer readers remained constant regardless of word frequency.

Whilst careful reading may be wide-spread, it may be only one of several types of reading that are required to accomplish academic reading into writing tasks. Chan (2013, see section 2.3.5 of this chapter for a full discussion of Chan's work) suggests that during the meaning and discourse construction phase, participants engage in high-level reading processes. Khalifa and Weir (2009) suggest that skilled readers engage in different types of reading. Khalifa and Weir (ibid) differentiate between careful and expeditious reading (see section 2.3.5.3 of this chapter for a full discussion). Therefore, this research wishes to not only identify when reading is likely to be taking place, but whether it fits the patterns of careful reading suggested in the literature, or whether the eye-movements of the participants suggest that some other type of reading (selective reading) is taking place.

Rayner et al. (2012:377) acknowledge that the type of reading engaged in is likely to vary according to the reader's purpose for reading. Rayner limits the discussion of other types of reading to skimming: ‘the type of reading activity in which you skim over the text without really deeply comprehending it’ (Rayner et al., 2012:377). However, Rayner et al. do not elaborate on the type of eye-movements which might be characteristic of skim reading.

The skimming discussed by Rayner et al. would seem to fit into what Khalifa and Weir (2009) describe as expeditious reading. Khalifa and Weir describe expeditious reading as ‘quick, selective and efficient reading to access desired information in a text’ (2009:46). Khalifa and Weir suggest that expeditious reading involves targeted reading that does not aim to extract a complete understanding of the text. Instead expeditious reading incorporates skimming, scanning and searching in which, the writer suggests, the emphasis is on the reader’s purpose for reading, rather than an attempt to comprehend everything that the writer wishes to communicate. In the absence of literature defining the eye-movements that characterise expeditious or selective reading, section 2.5 includes a discussion of how the types of reading suggested by Khalifa and Weir might be characterised by various types of eye-movements in reading. However, before that the researcher wishes to briefly review the literature relating to eye tracking studies which have focused on text level processing, rather than individual cognitive processes within reading.

### 2.4.3 Eye tracking studies in global text processing

Studies which have used eye tracking to study global text processing are largely limited to the work of Hyönä and Lorch and colleagues. Hyönä, Lorch and Rinck (2003:314) define global text processes as 'those processes that that identify and represent relationships between pieces of text information that span relatively long distances in a text.' Hyönä et al. (2002) cite comprehension of the way topics and sub-topics in a hierarchical expository text relate to one another as an example of global text processing.

Investigation of global text comprehension focuses on differences in the way sentences are read according to the type or importance of the sentence. Hyönä (1994) and Lorch, Lorch and Matthews (1985) concluded that topic sentences are processed more slowly than subsequent sentences elaborating on a new theme. This effect is reduced when the new topic is easily related to the previous topic. Hyönä et al. (2002) suggest that some readers exhibit a strategic return to topic sentences and headings after finishing a paragraph.

Another early study on global reading reported interesting findings regarding the incidence of regressions. The work of Vauras, Hyönä and Niemi (1992), which focused on readers' ability to accurately recall coherently and incoherently structured texts, recorded a greater incidence of regressions on structurally incoherent text than on structurally coherent text, whilst differences in rates of forward moving fixations were not significant. This suggests that high rates of regression occur when the reader is having difficulty constructing an understanding of the paragraph or text.

Hyönä, Lorch and Kaakinen (2002) have also focused on the different types of global reading strategies employed by students. In the study, eye tracking data was analysed to reveal the reading strategies of 48 students reading two multi-topic expository texts for a summarisation task. Hyönä et al. (2002) analysed every sentence according to the number and duration of fixations in each of the following four categories:

- Forward fixations during the first reading of the sentence
- Regressive fixations during the first reading of the sentence i.e. fixations which re-inspected a word that had already been read
- Fixations which returned to a sentence after the reader had moved past it (These were termed 'lookbacks' rather than regressions)
- Fixations which occurred after a regressive saccade out of a sentence

Using cluster analysis, they identified four distinct reading profiles: fast linear readers, slow linear readers, non-selective reviewers and topic structure processors. Hyönä et al. (2002) stated that both categories of linear readers made very few 'lookbacks' (regressions outside the currently read sentence) to earlier sections of the text; in contrast, topic structure readers and non-selective readers both engaged in high numbers of revisits to earlier parts of the text. However, what differentiated the topic structure processors from the non-selective reviewers was that the topic structure processors demonstrated much greater selectivity in their revisits, returning to sentences containing strategically important information (topic sentences and summarising sentences).

Hyönä et al. (ibid) conclude by pointing out that whilst evidence from the literature would suggest that selective topic processing strategies would be the most effective approach to reading for summarising, less than 20 per cent of their student participants employed this approach.

Hyönä, Lorch and Rinck (2003) point out that many eye tracking experiments present sentences one at a time or prevent returns to earlier parts of a text. This prevents investigation of global text processing. The researcher is particularly interested in investigating the global reading patterns of participants; therefore, the design of this study allows participants to move freely back and forth through both the texts. No limit was placed on the number of times participants revisited any part of the task rubric or the source texts. The details of how this was achieved are explained in section 3.6.5.

The literature reported above suggests that a range of eye tracking measures can be used to gain insight into the reading processes of participants. For example, patterns of fixations lasting between 50-500 milliseconds moving forward approximately eight characters on average, with extremely short time delays between them (forward saccades in careful reading are reported to generally last just 40 milliseconds) would suggest careful reading. Evidence of more global text processing strategies can be gleaned from the order in which sentences are read and from the total number and duration of fixations on each sentence (after adjusting for sentence length). This study therefore proposes to gather eye tracking data and analyse it in terms of both the patterns of individual fixations indicative of careful reading and the wider patterns fixations in terms of their allocation to sentences.

## 2.5 Interpreting different types of reading through eye tracking data

As discussed in section 2.3, Khalifa and Weir suggest that the reader's goal or purpose for reading (goal setter) guides the type of reading (careful / expeditious, local / global) that will be used. This model recognises that there are several different types of reading that can be utilised by the reader to achieve their reading goals and that these goals may also dictate how much of the text will need to be processed. As stated towards the end of section 2.3.3, this study proposes to use the types of reading suggested by Khalifa and Weir (2009) to assist with the classification of reading. The following paragraphs explain how the physical properties of eye-movements in reading could be interpreted, in light of the Khalifa and Weir model, to suggest the types of reading that participants engage in as they process through a reading into writing task. Eye-movements are classified, and types of reading inferred by use of an algorithm which considers the properties of each fixation and larger patterns formed by successive fixations.

As discussed earlier in section 2.4.2, detecting reading from eye tracking data the researcher concluded that the Campbell and Maglio (2001) could be best adapted to suit the purposes of this researcher project. To explain, algorithms to detect reading from eye tracking data rely either on examining the pattern of fixations over a set number (window) of fixations or, as in the Campbell and Maglio approach, defining events that indicate when reading starts and using an accumulating set of fixations to base calculations on. The disadvantage of the moving window technique is that you are unable to detect reading until sufficient fixations have taken place (if your window size is 20 fixations, then 20 fixations must elapse before you can make your first calculation). With the 'episodic' approach you need only three fixations before a decision on reading behaviour can be made. This makes detection of reading faster and more flexible for tasks such as reading into writing where reading is likely to be conducted in a piece-meal fashion for large parts of the task (for instance when reading occurs during bouts of writing).

The research sought primarily to distinguish *careful* reading as described by Khalifa and Weir, from other types of *selective* reading. The researcher also wished to try and distinguish between Khalifa and Weir's local and global categories if possible. (See section 2.3.5.3 in this chapter for a more detailed explanation for Khalifa and Weir's four types of reading). However, the technical limitations of the algorithm prevented the researcher from categorising eye-movement behaviour to align precisely with the types of reading identified by Khalifa and Weir's model. Therefore, this section explains where the researcher's classifications and Khalifa and Weir's classifications align exactly and where they differed.

First the case of careful reading, both local and global, will be discussed. At the beginning of every episode the reading classification was reset to 'selective' reading. To begin classifying fixations as part of careful reading, a minimum of three short forward moving fixations needed to occur. On the third short forward moving fixation the classification changes from selective to careful reading. However, an additional constraint was imposed. short forward moving fixations must also be occurring at a ratio of 3:1 in relation to regressions. Only when both these criteria were met would fixations be classified as belonging to careful reading. Long jumps forward through the text were not permitted as part of careful reading and therefore the advent of a long jump forward would cause that episode of careful reading to end, resetting the classification to 'selective'.

To determine if the reading was local or global the algorithm also monitored whether reading in an episode remained within a single sentence. Khalifa and Weir suggest that the reader's goal in careful local reading is to arrive at sentence level understanding. This would therefore seem to limit the reading processes used in word recognition, lexical access, syntactic parsing and establishing propositional meaning. Khalifa and Weir suggest that some inferencing may be required but that careful local reading does not involve integrating individual propositions into a large meaning representation. Therefore, as long as the fixations classified as careful remained within a single sentence, the label *Careful Local* reading was applied. If the succession of careful fixations continued into a second adjacent sentence the categorisation would update to *Careful Global* reading. Khalifa and Weir suggest that careful global reading occurs when readers link individual propositions to build up a text level understanding of the text. Therefore, careful global reading also neatly aligns. As the methodical sequential parsing of one sentence is followed by the next sentence, the reader must necessarily integrate the new information, drawing inferences and making connections with the reader's knowledge of the world to in order to establish a broader text level understanding.  The researcher suggests therefore, that the eye-movement patterns identified as careful global reading reflect this cognitive process.

The link between Khalifa and Weir's Expeditious global reading and Expeditious local reading and the researcher's selective global and selective local (as the difference in names suggests) is less clear cut and do not align exactly. In Khalifa and Weir's model expeditious reading incorporates skimming to establish gist, search reading to find information on a predetermined topic or scanning to find an exact word or figure. The global and local classifications are applied slightly differently in the case of expeditious reading than in the case of careful reading. Let us reflect on the *Global* and *selective local* classifications generated by the researcher's algorithm and consider how they align or differ from Khalifa and Weir's Expeditious global and Expeditious local classifications.

According to Khalifa and Weir, skimming is selective sampling of the text in order to establish the overarching theme of the text and whether it aligns with the reader's goals for reading. Skimming will necessarily include long jumps forward through the text with quick sampling of sections of the text. The reader will not attempt to 'carefully' parse individual sentence as Khalifa and Weir suggest that readers will use as few details as possible to arrive at their understanding. Skimming must necessarily operate across more than one sentence. The researcher's algorithm would classify this as selective global reading and Khalifa and Weir classify it as Expeditious global reading.

In search reading, Khalifa and Weir suggest the reader is searching for information on a pre-determined topic. It differs from scanning as when scanning the reader is looking for an exact lexical match. When searching the reader is considering any word which might provide a semantic link to the information they are searching for. Search reading will necessarily involve long jumps through the text and will be distinguished by an absence of sustained short forward moving fixations. Khalifa and Weir suggest that when the search is confined to a single sentence it is Expeditious local and when the search incorporates a wide area of text it is Expeditious global. These categories align with the researcher's definitions as seen in Table 2.

The difficulty arises when considering scanning. Khalifa and Weir suggest that because scanning involves hunting for a specific fact or piece of information that is likely to be *contained* within a single sentence, it should be considered as Expeditious local reading, even though the search may inevitably involve searching across larger portions of text (paragraphs or pages). The researcher's algorithm will necessarily classify this as selective global reading, because it extends beyond a single sentence.

*Table 2 Types of reading as identified by this study and Khalifa and Weir*

| Form of reading | More than 3 short forward to 1 regression? | Long jumps permitted? | Incorporates more than one sentence? | Algorithm classification | Khalifa and Weir classification |
|---|---|---|---|---|---|
| Careful | Yes | No | No | Careful Local | Careful Local |
| | | | Yes | Careful Global | Careful Global |
| Skimming | No | Yes | Yes | Selective Global | Expeditious Global |
| Searching | No | Yes | Yes | Selective Global | Expeditious Global |
| | | | No | Selective Local | Expeditious Local |
| Scanning | No | Yes | Yes | Selective Global | Expeditious Local |
| | | | No | Selective Global | Expeditious Local |

## 2.6 Conclusion and research questions

The early part (section 2.2) of this chapter concluded that reading into writing tasks are best placed to elicit knowledge transformation (Bereiter and Scardamalia 1987), which is central to academic literacy, from students. Section 2.3.1 went on to outline what is meant by a cognitive process as discussion of cognitive processes was central to section 2.3 which considered various models of writing and reading that provide a broader understanding of how reading into writing might operate. Section 2.3.4 went on to consider the extent to which reading and writing are reliant on common processes. This section concluded that whilst reading into writing may draw upon processes from both reading and writing, research suggests that it represents an independent skill rather than an amalgamation of reading and writing ability. Section 2.3.5 went on to consider the research which attempts to explain the process of reading into writing in greater depth. The account of cognitive processes suggested by Chan (2013) is comprehensive and, the researcher suggests, plausible. Khalifa and Weir's (2009) account of reading in an academic context could offer a detailed account of how the high-level reading proposed in Chan's 2013 model is realised.

The latter part of the chapter reviewed the literature in relation to eye tracking. Section 2.4 reviewed what the research into eye-movements in reading has revealed to date about reading and the researcher concluded by suggesting that eye tracking research may shed light on more specific reading processes. Therefore, in section 2.5 the researcher suggested how eye tracking data could be analysed using an algorithm developed by the researcher to quantify the reading activity of students undertaking an academic reading into writing task according to Khalifa and Weir's model of reading.

Therefore, the researcher suggests that analysis of eye tracking data from reading, could usefully shed light on the following questions:

1. What are the characteristics of reading during an academic reading into writing test task?
2. What are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?
3. What are the similarities and differences between the way high and low scoring participants tackle an academic reading into writing test task?

The answers to these questions, could help develop a deeper understanding of the role of reading in the reading into writing construct. However, because not all the processes can be evidenced through eye tracking data, other methodologies (retrospective think aloud and a cognitive processing questionnaire) will also be used to help contextualise reading within the broader reading into writing framework.

# 3 Methodology

## 3.1 Introduction

This study used a mixed-methods research design (Johnson, Onwuegbuzie and Turner, 2007) which incorporated eye tracking recordings, retrospective think aloud (RTA) recordings and a cognitive processing questionnaire to investigate the reading behaviour of 30 students completing an academic reading into writing task. The research questions were:

1. What are the characteristics of reading during an academic reading into writing test task?
2. What are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?
3. What are the similarities and differences between the way high and low scoring participants tackle an academic reading into writing test task?

Eye tracking data generated by the eye tracker was triangulated with the RTA and questionnaire data. The same data was used to answer all three research questions. However, for RQ1 the data for all participants was analysed. For RQ2 the data was split into two groups: first year undergraduate participants and third year undergraduate / postgraduate participants and the data for the two groups was compared. For RQ3, the data was limited to the five highest scoring and the five lowest scoring participants. The data for the high scoring group was compared with the data for the low scoring group.

The chapter starts by giving an overview of the data collection procedure 3.2. The participants are detailed in section 3.3 and section 3.4 outlines the research design. Section 3.5 reports the procedure for collecting the data. The development of the research instruments is described in section 3.6. Finally, section 3.7 reports how the data was analysed. Summary tables, detailing all the measures used for each research question are included in 0

## 3.2 Overview of data collection procedure

This section gives an overview of the data collection procedures. The first stage of the data collection procedure involved recording the eye-movements of participants whilst they completed an academic reading into writing task, followed by the retrospective think aloud session. Lastly participants completed a cognitive processing questionnaire.

The data was collected from the participants one at a time. The procedure required the use of an eye-tracker and only one was available. To check the feasibility of the study a pilot was conducted involving three participants before conducting the main data collection. Where information was gained from the pilot to inform methodological decisions, this is reported.

For the main study each data collection session lasted approximately one and a half to two hours. The sessions began with completion of the relevant paper work (see section 3.4 for a more detailed explanation of ethics and consent). Participants were asked to complete a one-hour academic reading into writing task (See section 3.6.1). The task was delivered on computer via an interactive webpage (referred to as a user interface, see section 3.6.5). The participants read the source texts on screen and typed their answers on the computer. The participants were given training on the interface before starting the reading into writing task. Whilst participants completed the reading into writing task an eye-tracker was used to collect information about their eye-movements. The eye tracking data from periods when participants fixated on the source texts was analysed to deduce reading patterns and behaviours. The participants were asked to complete the task within one hour. It was explained that after an hour, if the participant had not finished, the data collection process would stop. They could continue to complete the task, for their own satisfaction, but the data would be excluded from the study. In fact, only one of the participants felt she had not finished and continued working after the data collection had finished. Most participants finished before the hour deadline. On average the participants completed the task in 55 minutes.

For the pilot study, the researcher used replay of the eye tracking recordings to make judgements about the type of reading that participants were engaged in. This involved replaying the recording and stopping every five seconds to make observations about the type of reading behaviour. However, the judgements were very subjective, and the process proved extremely time consuming. When the researcher tried repeat-coding sections of the recording to test reliability, it was found that sometimes coding varied. This led the researcher to seek a more technical, data driven solution to coding the data which is described in section 3.7.11

After completing the reading into writing task, participants were asked to watch a recording of their reading and writing behaviour to help elicit a retrospective think aloud (RTA) account of the task. The session concluded with the participants completing a short cognitive processing questionnaire.

## 3.3 Participants

As one of the aims of the research is to compare the reading behaviours of more experienced and less experienced academic writers (RQ2), the researcher recruited 15 first year undergraduates and 15 third year undergraduates or postgraduate students from a range of universities. Recruitment was opportunistic, with the researcher initially seeking participants from groups of students at The University of Bedfordshire. The task used in this study was developed by a team at the University of Bedfordshire as part of a diagnostic assessment of academic skills for Y1 undergraduates. Therefore, the task was particularly suitable for Bedfordshire students as it had been developed based on the level of difficulty of texts found in the University of Bedfordshire's library.

The researcher sought permission from course leaders before approaching students enrolled on courses. The response rate from course leaders was low and only those course leaders that had worked with the researcher previously responded. Once permission was granted the researcher visited lectures to recruit participants. A payment of £20 was offered to participants in compensation for travel expenses. Technical difficulties with the collection of eye tracking data (see sections 3.6.4) and issues with the software led to the researcher having to conduct data collection sessions with around 45 participants before 30 recordings of sufficient quality were achieved.

Participants were allocated a participant number when the researcher first met with them, 55 participant numbers were allocated in total. However, some participants did the initial interview but then failed to turn up for the main data collection. In other cases, the participant started the task, but the eye tracking software crashed during recording or would not start to record properly. The researcher deemed that once participants had reviewed or started the task, it was not valid to try and 'restart' the task again as participants would have already read the question and perhaps some of the source materials. Therefore, these participants were eliminated from the study. In a few cases, participants completed the entire task, but the eye tracking data was unreliable, either during monitoring of the task the researcher could see they were reading but the eye tracker was not detecting their eyes or in one case, reading glasses seemed to be distorting the signal. A total of 55 participants were recruited but only 30 recordings of sufficient quality were generated.

The difficulties described above resulted in the researcher needing to recruit more participants than initially anticipated. Insufficient students from The University of Bedfordshire came forward, therefore the researcher also recruited students via word of mouth of friends and acquaintances. Students recruited in this way were also either Y1 undergraduates or year three / postgraduate students. Table 3 shows the numbers of participants recruited according to their place and level of study. Appendix 24 reports the score averages on the reading into writing task for the participants according to their level of study and place of study. The similarity of average scores suggests that, allowing for their level of study, the participants were at similar levels of academic ability whether recruited from University of Bedfordshire or elsewhere.

*Table 3 Participants according to level and place of study*

| | Y1 undergraduates | Year three undergrads / postgrads |
|---|---|---|
| **Place of study** | **No. of participants** | **No. of participants** |
| Uni of Bedfordshire | 7 | 2 |
| Other universities | 8 | 13 |

As a result, the data from 30 participants was used in the study. Ages ranged from 18 to 43 with an average age of 22. A summary of the participant characteristics is presented in Table 4 . All the non-native speakers of English, from all universities, rated themselves as C1 or above on the CEFR.

*Table 4 Participant characteristics*

| Level of study | No. | Gender | No. | Native / non-native speaker | No. |
|---|---|---|---|---|---|
| Y1 undergraduates | 15 | Male | 3 | Native Speaker | 3 |
| | | | | Non-native speaker | 0 |
| | | Female | 12 | Native Speaker | 9 |
| | | | | Non-native speaker | 3 |
| Year three undergrads / post-graduates | 15 | Male | 8 | Native Speaker | 8 |
| | | | | Non-native speaker | 0 |
| | | Female | 7 | Native Speaker | 5 |
| | | | | Non-native speaker | 2 |
| Total | 30 | | 30 | | 30 |

## 3.4 Research design

RQ1 investigates what the characteristics of reading during an academic reading into writing test are. To answer this question, descriptive statistics were used to report the eye tracking data in terms of the following measures

a) The distribution of fixations across the different parts of the screen (the participants own work, the buttons to move page, the screen instructions, the task instructions, the written source texts, the diagrams and areas of the screen that were outside all of the aforementioned). Section 3.7.5 provides a more detailed explanation.

b) The distribution of attention across the different sentences in the source texts. The results are reported for both individual sentences and for sentences when grouped according to their relevance to the writing task. (as per section 3.7.6 and 3.7.7)

c) The type of movement between fixations on the source texts; short forward movements, long forward movements, short regressions, long regressions. (as per section 3.7.8)

d) The patterns formed by fixations on the source texts that suggest the amount and type of reading undertaken. (as per section 3.7.11)

The RTA data (3.7.12) and questionnaire data (3.7.13) were used to triangulate and contextualise the eye tracking data and assist with interpretation. For RQ1 measures a-d were calculated and reported for all participants. For RQ2 the data was split into two groups (year one undergraduates and year three undergraduates / postgraduates). Measures a-d were calculated for each group and compared. For RQ3 the data was filtered to create two groups (the five highest scoring participants and the five lowest scoring participants). Measures a-d were once again calculated for each group and compared.

## 3.5 Procedures for collecting the data

This research was conducted in accordance with guidance set out by BAAL (British Association of Applied Linguistics) and permission to collect data was sought and granted by the University of Bedfordshire Research Ethics committee.

The researcher arranged to meet students that expressed an interest in taking part in the study individually. During the meeting students were given some background information about the researcher, given a research information sheet, told what would happen during the main data collection procedure and given the opportunity to ask questions before informed consent was sought. Appendix 1 includes a copy of the ethics committee letter giving approval for the study, Appendix 2 sets out the details of how this study complied with the BAAL guidance, Appendix 3 shows the research information sheet and Appendix 4 shows the research consent form (copies of completed forms have not been included to preserve the anonymity of the participants). Once informed consent had been given participants were interviewed informally to collect background information and a date for the main data collection was arranged.

During the data collection session, participants were seated at the monitor, the monitor height was adjusted, and the eye-tracker was calibrated (see Figure 7). Participants received a demonstration of the user interface used to present the task before starting.

Participants were asked to try to complete the task within 60 minutes. Participants were advised that if, after 60 minutes, they had not completed the task the researcher would stop recording their activity. If they wished to continue and complete the task they were free to do so, however their activity after the 60 minutes would not form part of the research. Only one participant elected to continue to work after the 60-minute deadline. Participants were advised that if they finished the task before the 60-minute time limit this was ok, and they could indicate that they had finished, and the recording would be stopped. The researcher monitored the participants' reading and writing activity during the task.

After the task was completed, a retrospective think aloud was conducted. Participants were played a screen recording which showed their emerging answer along with their eye-movements superimposed on the task rubric, the source texts and their own emerging answer. The researcher prompted the participants to recall their thinking during sections of recording.

Once the retrospective think aloud was completed the participants were asked to complete a short questionnaire.

## 3.6 Development of research instruments

To facilitate this research, the researcher needed to find a task that was representative of an academic reading into writing task, establish an environment and procedure that allowed good quality eye tracking data to be collected, design a retrospective think aloud protocol and finally design a cognitive processing questionnaire. Once all of the above were in place, the researcher needed to recruit participants. This section describes the development of these research instruments, explaining where information from a pilot study informed decision making.

### 3.6.1 The reading into writing task

As has been discussed in the literature review, reading into writing tasks closely reflect the demands placed upon students in the UK Higher Education context (Carson, 2001, Horowitz, 1986a, 1986b, Plakans, 2010). Therefore, the researcher sought a reading into writing task to elicit the source-based reading behaviour that students would ordinarily engage in whilst completing a written assignment. To be feasible the reading into writing task for this study also needed to meet the following criteria:

- Take no more than one hour – Eye tracking generates an enormous amount of data and tasks taking more than an hour were likely to prove difficult to analyse because of the volume of data. Much of the previous research using eye tracking had been limited to recording participants reading short sections of text or single target words or sentences in order to understand how our brains control our eyes during reading (for example the SWIFT model by Engbert et al., 2005; E-Z Reader by Reichle et al., 1998) or to understand how reading alters according to variations at word or sentence level (for example word frequency or syntactic complexity; see section 2.4 in the literature review for a more detailed discussion). The few studies which used eye tracking over longer periods (Bax, 2013, Rinck, Gamez, Diaz and De Vega, 2003, Hyönä, Lorch, Kaakinen, 2002) limited the metrics analysed or limited areas of interest.  As the researcher would be analysing a wide range of metrics, over the full period of the task, it was not practical to consider a task that would last more than one hour. The researcher also needed to consider the total length of time that the data collection session would last. Once time was added for calibrating the equipment, conducting the stimulated recall and completing the questionnaire the total amount of time required was close to two hours. The researcher found during the pilot study that asking participants to give up more than two hours of their time made recruiting participants very difficult indeed.

- Capable of being adapted for on screen delivery – In order to use eye tracking to monitor reading behaviour, the reading materials needed to be on-screen.
- As the researcher wished to compare the reading into writing behaviour of inexperienced academic readers, with the behaviour of more experienced academic writers the task needed to be accessible for first year undergraduates whilst still being relevant to year three undergraduates / postgraduate students.
- Incorporate multiple source texts – As the clear majority of academic writing involves synthesising information from multiple sources (Cooper and Bikowski, 2007), it was imperative that the task require participants to use at least two written source texts.
- Contain source texts of at least 250 words as texts shorter than this would not provide sufficient written text to potentially give rise to all the four types of reading proposed by Weir (see section 2.3 of the literature review).

The researcher reviewed current commercially available reading into writing test tasks. Presently four exams currently elect to include source-based writing as part of their exams. TOEFL iBT uses a single reading passage and an audible lecture excerpt as the prompts for their writing task and was therefore not suitable. The GEPT Advanced only uses a single paragraph as the source text prompt, which also ruled this out. Although at level IV, Trinity ISE requires candidates to read two sources, only one of them is a written text and the other is a non-verbal (diagram or graph) source. The IELTS academic paper requires participants to write about non-verbal sources, with the standard IELTS writing paper calling for a summary of a single source. For the reasons set out above, none of these test tasks was feasible. However, although not yet available to the public, the team at CRELLA have developed an academic reading into writing test, the Bedfordshire Academic Reading into writing Test (BART-W). A summary of test criteria is provided in Table 5.

*Table 5 Summary of reading into writing test criteria*

| | Tests | | | | |
|---|---|---|---|---|---|
| Selection criteria | TOEFL iBT | GEPT Advanced | Trinity ISE level IV | IELTS academic | Bedfordshire Academic RiW test |
| 1-hour time limit | Y | Y | N | Y | Y |
| Can be adapted to an on-screen task | Y | Y | Y | Y | Y |
| Accessible to all levels of HE | Y | Y | Y | Y | Y |
| Multiple written sources | N | N | N | N | Y |
| Text length greater than 250 words | N | N | N | N | Y |

The task from this test fits all the criteria detailed above. The wording of the task asks the students to "write a short essay drawing upon all the information (verbal and non-verbal) contained in the two articles". The source texts included two short articles. The first article included text (verbal information) and a line chart (non-verbal information), the second article included text (verbal information) and a pie chart (non-verbal information) . It goes on to set out four elements that the students must include, namely:

1. a title which summarises the content of the two articles.
2. a description of the problem and its causes as identified in the articles.
3. a summary of the different solutions suggested in the articles.
4. a conclusion which states which solution the student feels is most effective and explains why.

A full copy of the test has not been included in the appendix for reasons of test security, however the wording of the source texts can be seen in the left-hand column of Appendix 23 .

The BART-W task is designed to be given to first year undergraduates, upon arrival at university as part of a diagnostic screening process to identify students in need of academic support. The reading into writing task allows students 60 minutes to complete the task. Students need to write a short (200 word) essay based upon information drawn from two source articles. The articles have a shared theme and describe a problem, its causes and detail various potential solutions to the problem. Each source text is approximately 300 words long and includes a graph or diagram. Text 1 includes a line chart plotting the increasing number of working days lost due to work related stress. Text 2 incorporates a pie chart showing the causes of work-related stress. There is some repetition of information across the two source texts, and some information is highly relevant to the question, whilst other information is less relevant.

As described above, students are required to describe the problem, its cause(s) and identify solutions. Then students must conclude by stating which solution they favour and give their reasons. To score well students must include the most relevant information from both sources and organise their writing in response to the task. In addition, their answer needs to be cohesive and coherent; expressing ideas and information grammatically, accurately and clearly. Students are not penalised for poor referencing practice; however, candidates are warned against copying chunks of the text. The assessment criteria and rating scale for the test can be found at Appendix 5.

### 3.6.2 The specification of the eye-tracker

In reality, the researcher had no choice in selection of the eye-tracker. Eye-trackers are expensive and valuable pieces of equipment, even rental of them is prohibitively expensive. Therefore, the researcher used the only eye-tracker available to her; fortunately, it was well suited the researcher's purposes. Tobii reports that this model of eye-tracker 'is perfect for studies outside of the lab, and it is designed for fixation-based research' (Tobii, 2018).

The Tobii X2-60 used for this research enables the participant to freely move their head within a 500mm x 360mm 'window' at a distance of between 400mm – 900mm from the monitor. The system recommends a screen size up to 25'. The X2-60 samples at 60Hz (60 readings per second) with a latency of less than 35 milliseconds. Latency refers to the lapse in time between the eye-movement and the recording of the eye tracking data. Holmqvist et al. (2011) suggest that ideally latency should be no more than the time taken to collect three data samples. The X2-60 collects samples every 16.7 milliseconds (1000 millisecond/60 = 16.66ms) therefore, three samples are collected over 50.1 milliseconds.

The X2-60 uses both bright and dark pupil tracking and both are assessed during calibration to determine which yields the most accurate result depending on the light conditions and the physical features of the participants' eyes. The X2-60 is a binocular system, tracking both eyes. When only one eye can be detected, the position of the other eye is interpolated.

Whilst many articles report the use of eye tracking (for example: Hannus and Hyönä 1999), few set out a definitive list of the technical eye tracking information which researchers should present. Two articles which do focus on the reporting of technical information for eye tracking are Oaks (2010) and McConkie (1981). In addition to giving details of the eye-tracker used, Oakes suggests describing the geometry of the testing situation and stimulus, describing the calibration procedure, outlining procedures for dealing with missing data, describing the processing of the eye tracking data (including details of any data which was excluded) and lastly, how the areas of interest were defined. McConkie focuses on the reporting of factors relating to quality and propose that data relating to three key areas should be reported, namely the characteristics of the signal, the algorithm used to convert the raw data into fixations and saccades and finally the degree of accuracy. In the following sections the researcher has attempted to include all the data suggested by both articles.

### 3.6.3 Accuracy and precision of the eye-tracker

Tobii (2018) report accuracy of 0.4 – 0.6 of a visual degree and precision of 0.34 – 0.74 of a visual degree at a distance of 450mm – 800mm. The range of participants' movements, towards or away from the monitor fell within these distances. Using an average accuracy of 0.5 of a degree for accuracy the X2-60 was accurate to within 6mm on the screen at a range of 450mm – 800mm.

When the source texts were displayed on the screen characters were, on average, 7mm high with a clear 12mm between lines of text. With a screen height of 300mm or 1080 pixels (1080/300=3.6 pixels per mm of screen) and a screen width of 530mm or 1920 pixels

(1920/530=3.6) there are 3.6 pixels per millimetre. Therefore, the letters on screen were on average 25 pixels high with approximately 45 pixels between lines of text. The size of the characters in pixels is significant for the analysis of the eye tracking data which is detailed in sections 3.4

Tobii Studio reports the sampling rate. The sampling rate is the number of times Tobii Studio located the participants pupils as a percentage of the number of attempts. Appendix 25 reports the weighted gaze sampling rate which ranged from 37% to 80%. Although some of these rates are low by comparison to other eye-tracking studies, for this study the sampling rates were not reliable method of checking the quality of the data because there were significant periods when the participants were not looking at the screen. For example, some participants were not touch typists. This meant that whilst typing their answer they looked at the keyboard rather than at the screen. It was also common for participants to close their eyes or gaze around the room whilst thinking during writing activity. In addition, the recordings started with a short demonstration of the user interface. During this time participants frequently turned to look at the researcher whilst asking questions. The Tobii Studio Manual (2016:p40) warns that…

> *the eyes cannot be found when a person is looking away from the screen; this will result in a lower percentage. If the test is performed so that the participant is asked questions during the recording and occasionally looks at a test leader, an alternative method for determining the quality of the recording should be considered.*

Therefore, the researcher used a synchronised video recording to ensure that the data represented a full record of the participants reading activity. To do this a webcam was used to film the participants during data collection. Tobii Studio includes the option of recording the participant's face during the recording. The webcam footage of the participant can be replayed through Tobii Studio, precisely synchronised with the replay of the screen image overlaid with eye movements. By watching the synchronised replays, the researcher was able to verify that gaps in the eye movement data were due to participants closing their eyes or looking away from the screen (at the keyboard, the researcher or elsewhere in the room) rather than as a result of a failure to record reading activity. Any recordings with gaps in the eye movement data that could not be accounted for by participants looking away from the screen were excluded from the study. For this reason, and those reported in section 3.3, although 55 participants were recruited only the data for 30 participants was analysed.

Having excluded any recordings with unexplained gaps in the data, the remaining recordings were checked for quality. The software supplied with the eye-tracker, Tobii Pro Studio (2017), rates each fixation from 0, the highest level of confidence to 4, the data has been interpolated from surrounding data in the absence of a reading. The confidence for each eye is reported separately. A summary of the confidence scores (labelled validity score in the output tables from Tobii Studio) can be found in Appendix 6. For this study 94 per cent of fixations had the highest confidence rating for one or both eyes. Thus 6 per cent were interpolated.

### 3.6.4 Creating an environment to facilitate collection of accurate eye tracking data.

During the pilot phase the researcher established that a range of factors could negatively affect the reliability of the eye tracking data. These are now described and discussed in turn.

### 3.6.4.1 Calibration of the eye-tracker

Before starting to collect data the eye-tracker needs to be calibrated to each individual participant. Calibration involves asking participants to follow a dot which moves around the screen, whilst the eye tracker collects and analyses data about the participants eyes. The number of points that the dot moves to and the speed at which the dot moves can be selected by the researcher. The researcher found that using a slow nine-point calibration (the highest number of points permitted) gave the most accurate calibration. During calibration the moving dot pauses at nine different points: three rows (top, middle and bottom of the screen), each containing three calibration points (left, middle and right of the screen). During calibration the eye tracking software compares both light and dark pupil measurements to identify which readings are most accurate given the light conditions and the physical characteristics of the participant's eyes.

After the calibration the researcher performed a check to review the accuracy of the calibration. This involved re-displaying the calibration points and asking the participants to focus on the points one at a time. The researcher was able to see a screen with both the calibration points and a moving dot, representing the participant's gaze point as recorded by the eye tracker. This enabled the researcher to check the accuracy of the calibration. If the researcher was not satisfied with the accuracy, the calibration exercise was repeated.

### 3.6.4.2 Lighting

The overwhelming majority of eye tracking equipment (the Tobii X2-60 included) rely on tracking the reflection of light on the participants' cornea. If the lighting in the room alters during the session, the accuracy of the eye tracking data can be affected. For example, during the day, if the sun comes out and increases the light levels in the room this can affect the accuracy of the eye-tracker. When the researcher conducted the pilot a room with no windows was used. However, this room was not always available and / or participants were not necessarily able to travel to the campus where this room was located. Therefore, during the main study the researcher conducted several sessions where, as lighting conditions altered, the accuracy of the eye-tracker diminished and data from that session was unusable. Therefore, it is advisable to have a room where there is no natural daylight or where blinds or curtains can be used to minimise changes in levels of daylight entering the room. The room needs to be well lit; poor levels of lighting cause the participants' pupils to dilate and this can reduce the accuracy of the data gathered. The type of artificial lighting in the room also has an impact on the quality of the eye tracking data.

Although Holmqvist et al. (Holmqvist et al., 2011:17) suggest using fluorescent lighting, the researcher has found that if a fluorescent light is flickering this reduces the accuracy of the data, especially when the room is illuminated by a single fluorescent bulb. The researcher had greater success in a room lit with a series of LED lights which provided bright, evenly lit conditions. Data was gathered most successfully in a room with closed blinds to reduce the effects of natural daylight and multiple LED lights.

#### 3.6.4.3 Stability of the equipment

It may seem obvious but the surface on which the eye-tracker is placed needs to be extremely rigid. For this research the participants needed to be able to type on a keyboard without disturbing the eye-tracker mounted on the monitor. Once again during the pilot study and during early data collection sessions the researcher discovered that even small vibrations of the monitor, to which the eye-tracker is attached, proved sufficient to interfere with the accuracy of the data, once again leading to inadmissible data. The researcher suggests that as a rule of thumb if, when you place a glass of water a desk or table and proceed to gently thump the desk the surface of the water in the glass is disturbed, the desk is not a sufficiently sturdy surface on which to place the eye-tracker.

Holmqvist et al. (Holmqvist et al., 2011) suggest that the floor on which the desk or table stands should be solid concrete to prevent interference from the movement of people in the room. The researcher had the greatest success when the equipment was mounted on an extremely heavy oak table, standing on a concrete floor.

#### 3.6.4.4 Geometry of the arrangement of the equipment

Experience from the pilot study taught the researcher that participants seated in chairs fitted with castors can lead to participants moving around excessively during data collection sessions. Therefore, participants were seated in comfortable chairs with no casters to reduce the movement of participants during data collection. The monitor was positioned 650mm from the participant and adjusted so that the centre of the screen was level with the participants' eye line. The screen on the 25' monitor measured 532mm wide and 300mm high, with a resolution of 1920 x 1080 pixels. Figure 7 illustrates the set up.

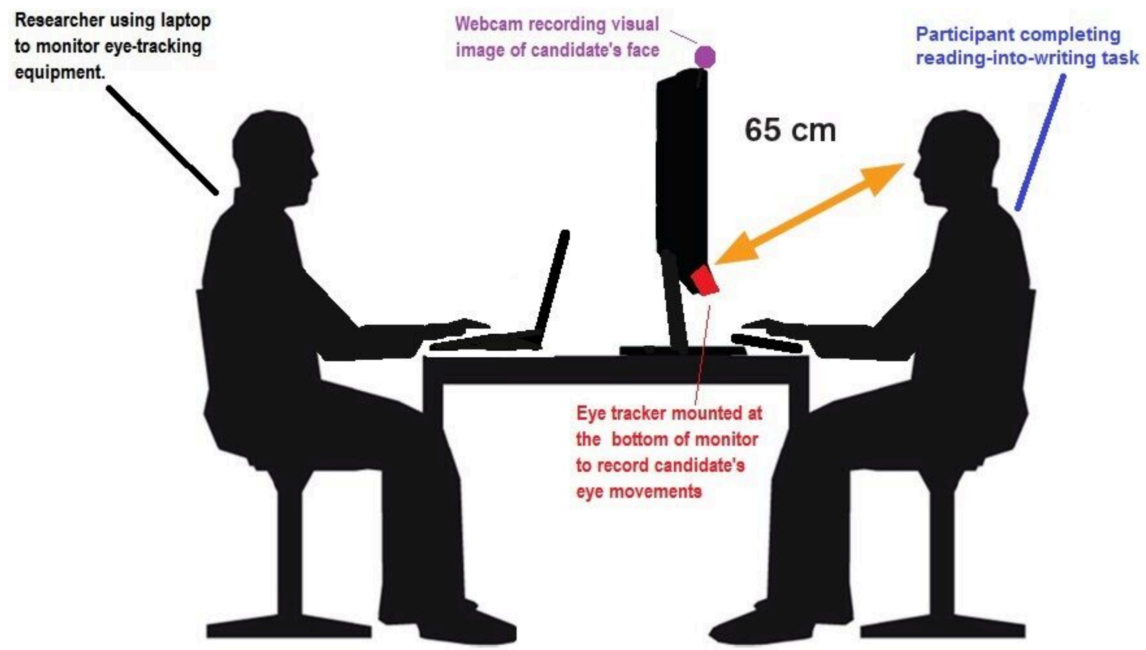



*Figure 7: Arrangement of eye tracking equipment for data collection sessions.*

## 3.6.5 Displaying the reading into writing task on screen

In order to eye-track the participants reading as they completed the reading into writing task the researcher had to develop a means of presenting the task on screen. There were several pages of source texts as well as the task instructions that needed to be displayed. They could not all be displayed at the same time because the text would have been too small, and the collection of accurate eye tracking data would not have been possible. Therefore, the researcher developed an interactive website to present the task which allowed the participants to 'turn the page' at will, allowing participants to select the reading material they wished to read alongside their emerging answer. This website will be referred to as the *user interface*. A screen shot of the user interface can be seen in Figure 8.

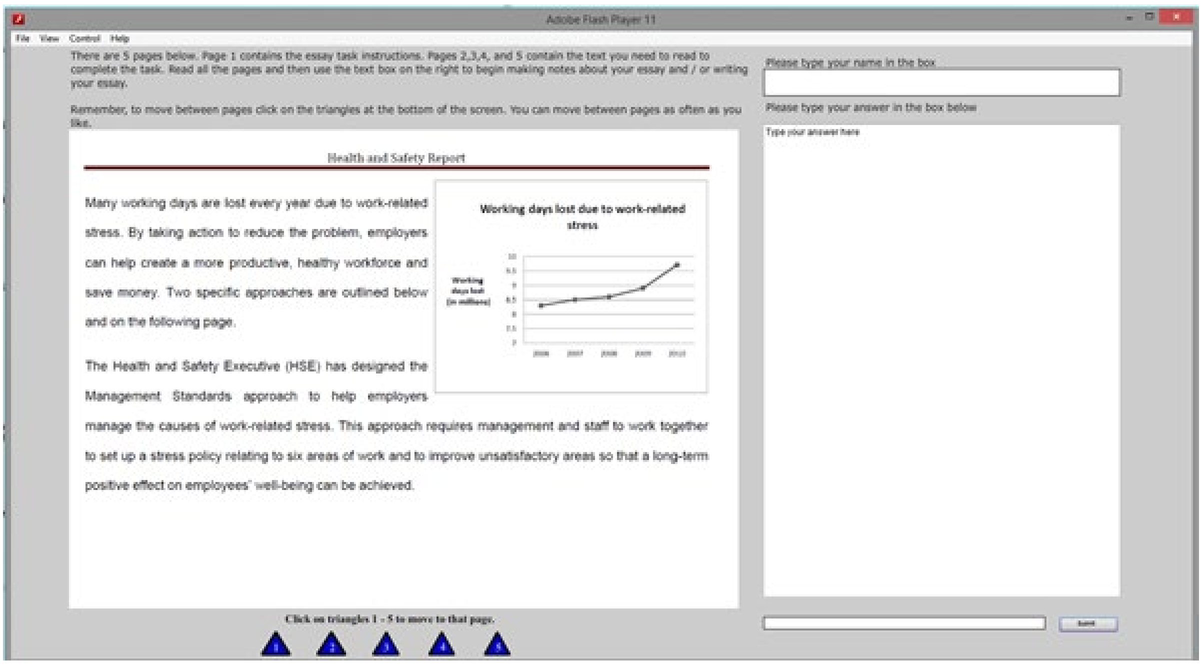



*Figure 8: Screen shot of the interface*

The five blue triangles along the bottom of the screen allow the participant to change the content displayed in the source content area directly above the triangles. The source content area is identified in Figure 9.

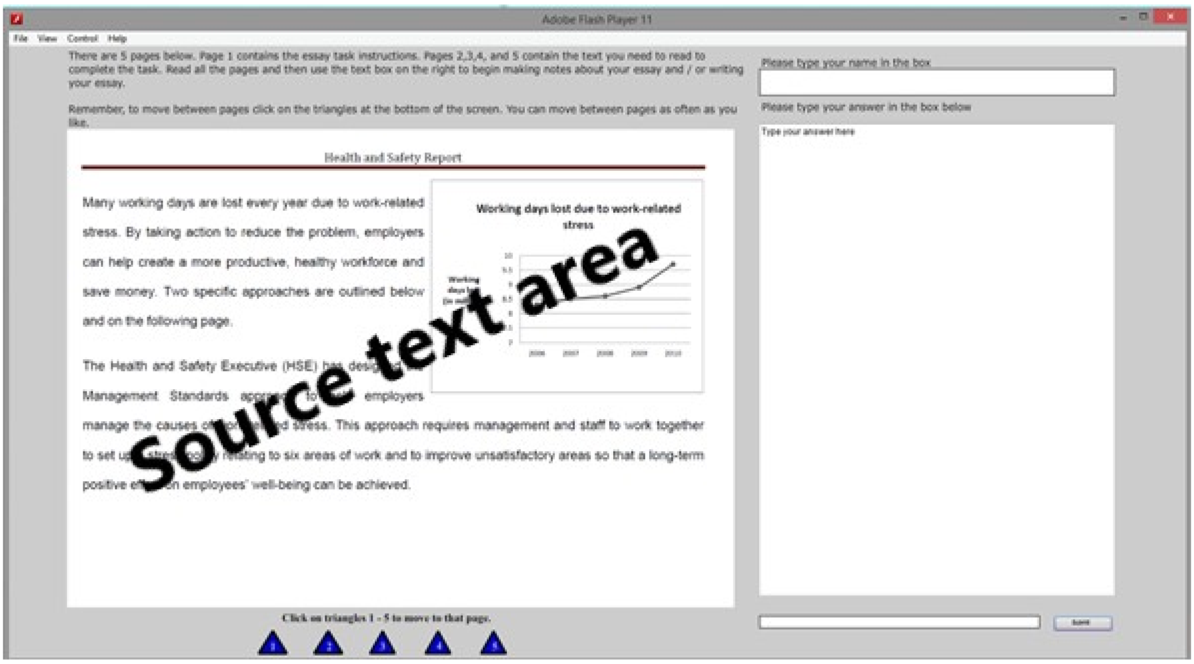



*Figure 9 Screen shot of interface with source text area identified*

As the participant clicked on the triangles numbered one to five the content displayed in the source text area changed. Page one contained the task rubric, page two contained the first part of source one, page three contained the second part of source one, page four contained the first part of source two and page five contained the second part of source two.

The participants were shown how to use the interface before starting the data collection. There was no limit to how often the participants changed the content displayed in the source text area. Text at the top of the screen acted as a reminder for participants and read as follows:

> *There are 5 pages below. Page 1 contains the essay task instructions. Pages 2,3,4 and 5 contain the texts you need to read to complete the task. Read all the pages and then use the text box on the right to begin making notes about your essay and / or writing your essay. Remember, to move between pages click on the triangles at the bottom of the screen. You can move between pages as often as you like.*

To compose their answer, participants clicked in the white rectangle on the right-hand side of the screen. The composition area is identified in Figure 10.

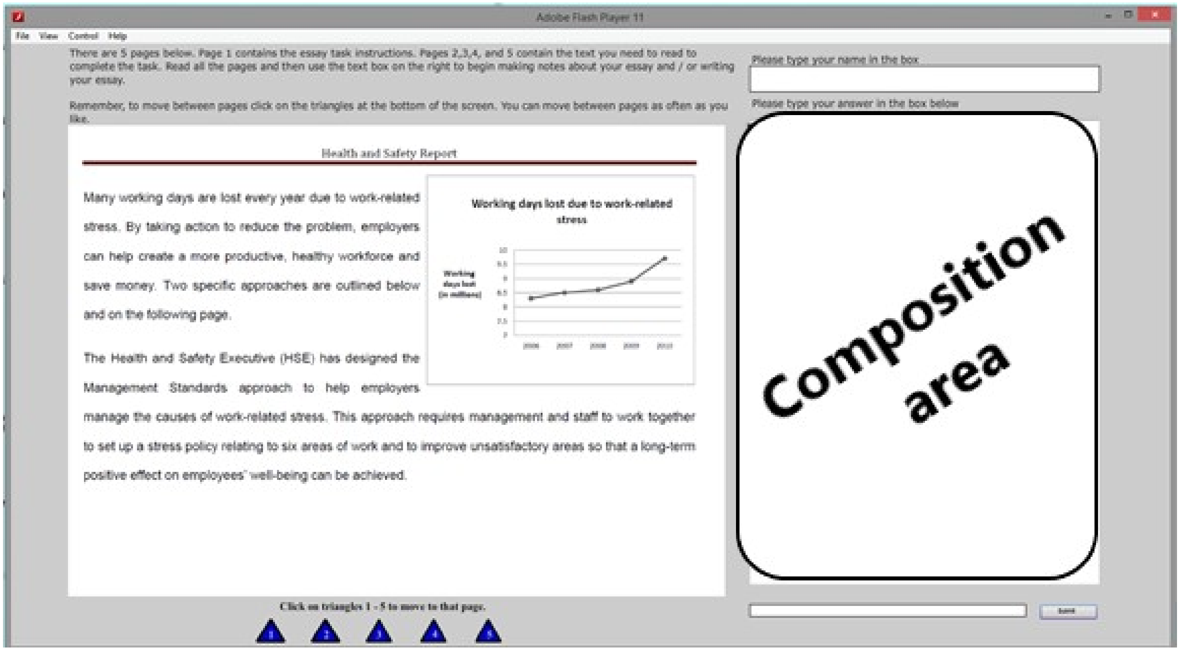



*Figure 10 Screen shot of the user interface with the composition area identified.*

The composition area acted as a 'page' on which the participants composed their answer. When the participant changed page in the source text area, the rest of the interface, including the composition area remained unchanged.

Whilst participants worked on the task, the researcher was able to observe both their eye-movements and their typing activity on another screen. Figure 11 shows a screen shot of the researcher's screen which displays a snapshot of a participant's eye-movements indicated by the red dots numbered one, two, three and four. These red dots indicate the participants fixations. Section 3.7 gives further information about what raw eye tracking data is generated and section 3.7.3 explains how fixations are identified from the raw eye tracking data.

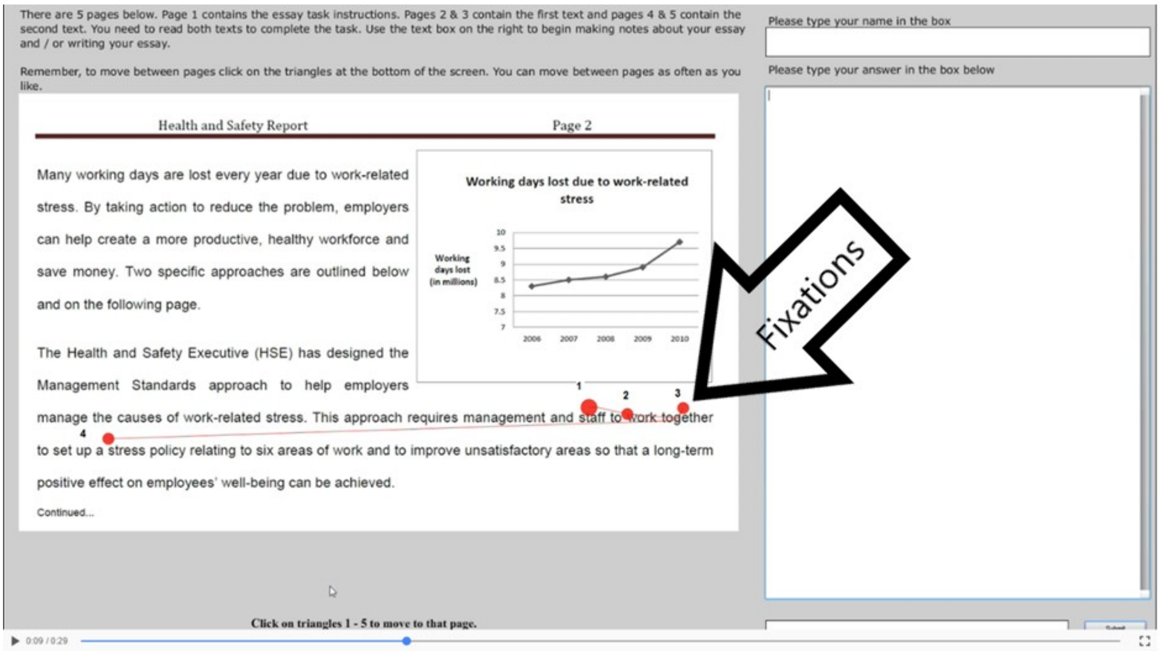



*Figure 11 Screen shot of researcher's screen*

Once participants completed the task they were asked to engage in a retrospective think aloud (RTA) session. The development of the script for the RTA session is explained next.

### 3.6.6 Development of the RTA protocol

The purpose of the RTA was to provide a broader context in which to situate the eye tracking data and to assist with interpretation of eye tracking data. In addition, the RTA data might indicate whether differences between groups of participants (in relation to RQ2 and RQ3) which might emerge from the eye tracking data were representative of differences across every stage of the reading into writing process.

Retrospective think aloud involves asking the participants to watch a replay of their reading and writing activity and asking them to comment on their thoughts and intentions as they progressed through the task. The researcher felt that retrospective stimulated recall would be preferable to concurrent verbalising of thoughts which might interfere with candidates reading, thinking and writing processes, especially for any candidates for whom English was not their first language.

During the pilot the researcher found it difficult to prompt participants to talk about their *thoughts* at the time of reading / writing, participants tended to simply describe what they could see themselves doing. Therefore, the researcher developed a protocol to help facilitate the process as per Gass and Mackey (2000). This helped the researcher to ensure that nothing was forgotten and to ensure that the wording used focused the participant on reporting their *thoughts at the time of the activity* rather than reporting their actions or their present reaction to the replay. The protocol was developed based on examples presented in Gass and Mackey (2000) and a copy can be found in Appendix 7.

After the participant was recorded completing the task, a replay of the participant's reading and writing activity was screened and the researcher used the protocol to prompt participants to verbalise their thinking and motivation at the time of reading / writing. This process was facilitated by the eye tracking software which generates a recording in which participants see a moving red dot superimposed on the screen recording, as illustrated in Figure 11. Effectively after the recording had finished the participants were able to see what the researcher had been observing during the task, a red dot moving to indicate where the participants gaze had been focused and their written answer gradually being typed. The replay of the recording could be fast-forwarded, rewound or paused to facilitate the retrospective think-aloud process. As the purpose of the study was to gain a deeper understanding of the cognitive processes of reading into writing the protocol was structured around Chan's (2013) cognitive processes, which have been discussed in depth in the literature review section 2.3.5. The full RTA protocol can be viewed in Appendix 7, but an example of how the prompts were linked to Chan's (2013) cognitive phases can be seen in Table 6.

*Table 6 Example of RTA prompts based on Chan (2013) cognitive processes of reading into writing*

| **Cognitive processes** | **Working definitions** | **Initial prompt** | **Further prompt** |
|---|---|---|---|
| Task representation | • Create an initial understanding of the task (e.g. the overall purpose of the test/assignment, structure of the test, time constraints, scoring criteria, word length, topic, genre and intended reader, rhetorical functions to perform) | It seems as though you had just read the question.<br><br>Can you tell me about that? / Can you recall what you were thinking at that time? / Do you remember thinking anything after you had read the question? / Can you tell me what you thought when you had read the question? | • At that point, what were you thinking about finding ideas to write about?<br>• At that point did you think about what style to use?<br>• At that point, did you think about how to organise your answer?<br>• At that point did you think about what to do next?<br>• At that point did you have any goals or targets? |

As suggested by Gass and Mackey (ibid.), during the RTA process the researcher tried to provide non-committal but encouraging responses such as ‘Oh / mmm / great / good / I see / uh-huh / ok’ in a bid to reduce researcher influence. Gass and Mackey (ibid.) suggest delaying any follow-up questions about specific comments made by the participant until the end of the stimulated recall session, once the replay has finished. The researcher tried to observe this advice also. On average the RTA sessions lasted about 15 minutes. RTA interviews were conducted with all the participants, however technical issues with the recording equipment resulted in three missing recordings and one very short recording. The coding and analysis of the RTA data is described in section 3.7.12

### 3.6.7 Design of the questionnaire

The purpose of the questionnaire was to provide a broader context in which to situate the eye tracking data but especially, to review whether differences between groups of participants (in relation to RQ2 and RQ3) which might emerge in the eye tracking data were representative of differences across every stage of the reading into writing process. Therefore, the questionnaire was also developed in response to Chan’s (2013) cognitive processes in reading into writing.

Although Chan had developed a questionnaire based on these cognitive processes, it tended to focus on establishing which activities writers engaged in during completion of a reading into writing task. In contrast, the current study wished to understand how the cognitive processes influenced participants’ reading and writing behaviour.

The format of the questionnaire consisted of a guidance statement asking participants to think about a particular aspect of the task, followed by a series of statements which participants had to rate on a 5-point Likert scale to indicate their level of agreement. A copy of the full questionnaire can be found at Appendix 8 however, Table 7 details the items related to each of the processes in Chan's (2013) framework.

*Table 7 Questionnaire items linked to Chan's cognitive processes*

| Items | Cognitive Process |
|---|---|
| **1-8** | Task representation |
| **9-18** | Macro-planning |
| **19-27** | High-level reading including connecting and generating |
| **28-32** | Organising |
| **33-44** | Monitoring and revising |

The rationale behind the formulation of the questions is now explained more fully, one cognitive process at a time.

#### 3.6.7.1 Task representation

Chan (2013) suggests that task representation determines the overall 'shape' and 'feel' of the participant's answer. Chan suggests that one of the key task representation decisions that writers need to make is what to include. Writers must decide whether to include only ideas from the source text or whether to add previous knowledge, comments or opinions. Therefore, questions one, two and three relate to participants perceptions of what they needed to include.

In addition, Chan suggests that during task representation writers may decide to use certain strategies to help them tackle the task. Question four checks whether participants are aware of using a conscious strategy to help them at this stage.

In 2.3.5.1review the researcher concludes that task representation may be a process which continues throughout completion of the task. Therefore, questions five and six assess participants inclination to return to the task instructions and whether their interpretation of the demands of the task developed over the course of the task.

Lastly, relating to task representation, question seven enquires whether participants set themselves any additional goals. Chan (2013) suggests that writers may give themselves targets over and above those outlined in the task.

#### 3.6.7.2 Macro planning

Table 6 appears to suggest some overlap between task representation and macro planning in that both refer to goal setting and the intended reader. However, task representation seems to focus on interpreting the question rubric whilst macro planning focuses on how the writer plans to adapt their writing to meet the perceived demands of the task. In summary macro-planning is concerned with writer's perception of how they can adapt their writing to incorporate relevant content, meet the expectations of the reader, and develop or adopt a suitable framework or over- arching theme which suits their purpose for writing.

Field (2004) suggests that skilled writers *continually* review and reorganise the ideas they select to include in their writing. In contrast, less skilled writers are unlikely to engage in reorganising main ideas, instead they focus on generating text, relying on the current idea to help suggest the next. Questions eight, nine and ten focus on understanding how far into the task macro-planning continued.

Field (2004) suggests that drawing upon world-knowledge the writer decides upon the goals of writing including consideration of the genre and level of formality. Questions 11 and 12 focus on genre and formality.

Kellogg (2006) emphasises that skilled writers shape their text to accommodate the way the reader is likely to interpret the text. Skilled writers do not produce a text to suit themselves, they produce a text that will suit the intended readership. This acute awareness of audience is reflected in the goals of the writer and affects the writer's decisions regarding genre and formality. Questions 13 and 14 ask about awareness of the reader and adaptation to suit the reader.

Shaw and Weir (2007) point to 'the relevance and adequacy of content to the task set, the appropriateness of the language used for the topic' as well as the effect on the reader as being indicators of the presence or absence of macro-planning. Accordingly, questions 15 and 16 concern relevance and adequacy.

#### 3.6.7.3 Reading

Chan argues that although it is easy to observe when and for how long someone engaged in a reading into writing task reads the source texts this information is far less important than which reading processes have been used. Khalifa and Weir's (2009) model of reading breaks reading into eight cognitive sub processes which range from low level processes such as word recognition through to high level processes such as creating an intertextual interpretation. See section 2.3 of the literature review for a full discussion of Khalifa and Weir's model. Khalifa and Weir suggest that readers deploy different kinds of reading depending on their purpose or goal. The reader's goal will also determine which cognitive processes are necessary to process the text under examination. Therefore, questions 19 – 23 focus on linking the different types of reading identified by Khalifa and Weir (2009) with the participants purpose or goal.

#### 3.6.7.4 High level reading including connecting and generating

Chan's (2013) model (Table 1) lists 'organising' before 'connecting and generating' which is distinct from reading. However, the researcher felt that overlap existed between some of the high-level reading processes and connecting ideas from the text with the reader's existing knowledge. Therefore, the researcher decided to group 'high level reading' and 'connecting and generating' together.

Kintsch and van Dijk (1978) and Grabe and Stoller (2011) propose broadly similar models of reading where the reader reviews the ideas from the text, creating a summary of the text in which ideas which are not repeated or are not key in linking to other ideas fade / are deleted from the reader's model of the text. Question 24 relates to evaluating the importance of ideas and question 26 relates to understanding how the main ideas within a text fit together.

Chan (2013:67) proposes that '(c)onnecting is a process in which writers generate links between ideas or new meaning by connecting ideas in the source texts with their own knowledge' based on the work of Spivey (1984,1990,1997). Question 25 relates to connecting new ideas to existing knowledge.

Rouet (2006) suggests that when having to integrate multiple sources of information the reader is confronted by three additional problems than when working from a single text. Firstly, the texts have not been written or designed to fit neatly together and so lack coherence. Secondly, there may be differences, discrepancies or even contradictions between the texts. Thirdly, when writing in response to a complex task, there may be no correspondence between the content of the documents and the task; unlike comprehension questions which are designed to 'fit' the text they accompany.

Rouet (ibid) suggests that when comprehending multiple documents readers construct a documents model. This model is made up of information about the sources (where the texts came from, who wrote them, when and implicit knowledge about the reliability of the source) and information about the contents of the various documents (details such as main ideas and propositions). The reader must form multiple links; who said what, whether the contents of one document support or contradict another, whether one source is more reliable than another. Goldman (2004) suggests that expert readers tend to 'make cross text comparisons to corroborate information, pay attention to the source of the research' and engage in strategic reading behaviour. Therefore, question 27 centres on linking main ideas across multiple texts.

#### 3.6.7.5 Organising writing

Chan bases her definition of organising on the work of Field (2004) and Shaw and Weir (2007). Field (2004:329) suggests that organising is a phase where the writer ''provisionally organises the ideas, still in abstract form, in relation to the text as a whole and in relation to each other '. Shaw and Weir (2007:38) suggest that organising is used to 'determine which are central to the goals of the text and which are of secondary importance'. Questions 28 and 29 relate to factors that influenced the ordering of ideas.

Spivey and King's (1989:11) definition of discourse synthesis includes three essential sub-processes, of which organising is the first. However, Spivey and King's organising (where writers think about the overall structure of their writing and the structure of the source texts) is perhaps closer to Chan's Macro-planning. Chan's definition of organising could perhaps be more closely related to Spivey's 'selecting' and 'connecting' processing in the sense that dominant ideas from the texts are identified and organised in relation to each other. In addition, Flower et al. (1990) suggest that writers will need to decide how to organise the ideas they have decided to include and reflect on issues such as whether an overarching idea / theme should be used to organise the ideas from the source texts. Accordingly, question 30 asks about participants use of a theme to help select ideas.

Several writers have suggested that good organising ability can be linked to the quality of the finished product. Kellogg (1994) concludes that when writers produce a more thorough plan, in the form of an outline, it enables writers to organise their ideas better prior to writing and results in a better final product. Chai (2006) also found that students whose writing plans showed a greater degree of organisation of idea units tended to produce better essays.

Parodi (2007) suggests that a lack of automaticity in lower level reading and writing skills leaves little working memory available for students to retain ideas across paragraphs and thus both perceive (when reading) and produce (when writing) a list of ideas rather than organising ideas into a more complex structure. Questions 31 and 32 ask about participants approach to combining ideas into more complex structures.

#### 3.6.7.6 Micro-planning

Micro-planning refers to planning the next individual idea. Field (2004:329) suggests that decisions about the content of the next sentence or paragraph are influenced by the macro-plan, the text so far (and to what extent it has met the goals outlined in the macro-plan) and 'whether an individual piece of information is or is not shared with the reader' based on what has been said in the text so far or based on shared common knowledge.

Shaw and Weir (2007) suggest that Field's model has advantages over the model proposed by Grabe and Kaplan (1996) in which the text so far has little impact.

It should be stressed that in Field's Micro-planning stage the propositions are still in the form of abstract ideas and have not yet been converted into linguistic form. The conversion to linguistic form takes place during the Translation phase (see section 3.6.7.7). As the processes involved in micro-planning would be repeated again and again in the course of completing the task the researcher felt that it would be better to use the RTA process to enquire about specific micro-planning incidents rather than asking participants to generalise.

#### 3.6.7.7 Translating

Field (2004) suggests that translating is when the proposed 'content that has been assembled...undergoes a process of conversion from abstract to linguistic form'. Because this process is brief and fleeting, and in many ways subconscious, it does not lend itself to retrospective recall. Even verbal recording at the time of translating is of limited value because of the automated subconscious nature of this process. Therefore, the researcher decided to exclude Translating from this study.

#### 3.6.7.8 Monitoring and revising

Chan (2013) proposes that monitoring is when the writer checks the quality of the text produced and, if dissatisfied with the text, embarks on an episode of revision. Chan sub-divides monitoring and revising into low level aspects such as accuracy or high-level aspects such as argument and coherence.

The Chenoweth and Hayes model of writing (2001) breaks writing into four components one of which is termed the reviser. The reviser is responsible for monitoring both emerging language which has not yet been transcribed and modifying existing text. When problems are identified, changes are initiated. Whilst according to Chenoweth and Hayes (2001) revision can be seen as a process which can occur during the translation process, before the emerging language has been typed or transcribed, such behaviours cannot be observed in behaviour and their fleeting, temporary nature (much like the translation process) make them difficult for participants to recall. For these reasons, revision prior to transcription will be excluded from this study.

Hayes, Flower, Schriver et al. (1987) offer a thorough break down of the revision process and explain the difference in the revision processes of expert and novice writers. They suggest that expert writers perceive revision as a whole-text task and consider their plans and / or notes as suitable targets for revision. Expert writers are consistently better at detecting problems, particularly at global level. Expert writers are also more likely have effective strategies for resolving problems once they are detected. In contrast, novice writers perceive revision as a process of looking for errors rather than as a means of improving the quality of the text. This leads to a narrower range of revisions. Questions 33 to 41 ask about the type of revisions participants made, some local, some global.

Kellogg (1996), Field (2004) and Shaw and Weir (2007) suggest that monitoring and revising are extremely cognitively demanding. This idea was singled out for attention by Quinlan et al. (2012) who concluded increased load on working memory reduced the ability of participants to successfully monitor and revise. Therefore, the researcher would expect to see less experienced writers making the majority of revisions after they had completed the bulk of their writing. Questions 42 and 43 reflect this.

After drafting the questionnaire, it was piloted with six university students, who were asked to complete the questionnaire after they had finished one of their course work tasks which required them to complete some source-based writing. Following their feedback changes were made to the layout, a model question was introduced and the ‘Think about’ statements where printed in blue to draw the participants’ attention to them.

As the questionnaire was intended to find if there were any broad differences between groups of participants, the researcher did not analyse the data for correlation within groups of questions (factors). The analysis of the questionnaire was limited to comparison of the results between groups in RQ2 and RQ3. (See section 3.7.12 for further details).

## 3.7 Procedure for analysing the data

In this section the way that the data was processed is reported. The procedure for the collection of eye tracking data is reported first (sections 3.7.1to 3.7.11.4) followed by the procedure for the RTA data (3.7.12) and lastly the questionnaire data (3.7.13).

### 3.7.1 Analysis of the eye tracking data

The raw eye tracking data is processed by the eye tracking software supplied with the eye tracker hardware and converted into fixations using an algorithm. There are a number of algorithms available for conversation of raw data into fixations and selection of the algorithm will inevitably have an effect on the fixation data that is generated. Therefore, the researcher will begin with a brief explanation of how raw eye tracking is converted into fixations because selection of the appropriate algorithm relies upon an understanding of this process.

### 3.7.2 The raw eye tracking data

As indicated in section 3.6.3 the eye tracker used takes 60 readings per second. Each reading records the time from the start of the recording (in milliseconds), the x (horizontal) coordinate of the participants gaze point, the y (vertical) coordinate of the participants gaze point for both the left and right eye. The coordinate locations are represented as pixels on the computer screen. In some cases, there will be an absence of data, such as when a participant has looked away from the screen or blinked. It is possible that readings for only one eye have been detected. Although the eyes appear to work in tandem, when recorded in minute detail it is evident that there are fractional differences between the left and right eye; for example, in the speed of the eye lids during a blink. Therefore, eye-tracker also generates a validity score for each eye for every reading. As discussed in 3.6.3the validity code indicates the confidence level that the left or right eye has been correctly identified. A sample of raw gaze data can be seen in Figure 12.

| ParticipantName | RecordingTimestamp | KeyPressEvent | GazePointLeftX (ADCSpx) | GazePointLeftY (ADCSpx) | GazePointRightX (ADCSpx) | GazePointRightY (ADCSpx) | ValidityLeft | ValidityRight |
|---|---|---|---|---|---|---|---|---|
| P49 | 1041716 | | 817 | 596 | 849 | 606 | 0 | 0 |
| P49 | 1041733 | | 833 | 598 | 861 | 633 | 0 | 0 |
| P49 | 1041750 | | 837 | 601 | 840 | 627 | 0 | 0 |
| P49 | 1041766 | | 828 | 613 | 844 | 624 | 0 | 0 |
| P49 | 1041783 | | 834 | 590 | 854 | 631 | 0 | 0 |
| P49 | 1041799 | | 829 | 592 | 848 | 603 | 0 | 0 |
| P49 | 1041816 | | 842 | 605 | 850 | 616 | 0 | 0 |
| P49 | 1041833 | | 838 | 612 | 847 | 615 | 0 | 0 |
| P49 | 1041850 | | 840 | 619 | 843 | 612 | 0 | 0 |
| P49 | 1041866 | | 831 | 588 | 856 | 618 | 0 | 0 |
| P49 | 1041883 | | 820 | 587 | 858 | 625 | 0 | 0 |
| P49 | 1041900 | | 840 | 603 | 850 | 646 | 0 | 0 |
| P49 | 1041916 | | 839 | 615 | 844 | 611 | 0 | 0 |
| P49 | 1041933 | | 836 | 607 | 864 | 611 | 0 | 0 |
| P49 | 1041950 | | 843 | 607 | 854 | 597 | 0 | 0 |
| P49 | 1041966 | | 837 | 589 | 860 | 613 | 0 | 0 |
| P49 | 1041983 | | 793 | 488 | 841 | 512 | 0 | 0 |
| P49 | 1042000 | | 780 | 338 | 798 | 305 | 0 | 0 |
| P49 | 1042016 | | 797 | 385 | | | 0 | 4 |
| P49 | 1042033 | | 790 | 345 | 800 | 313 | 0 | 0 |
| P49 | 1042050 | | 796 | 346 | 799 | 327 | 0 | 0 |
| P49 | 1042066 | | 783 | 353 | 803 | 334 | 0 | 0 |
| P49 | 1042083 | | 790 | 353 | 799 | 338 | 0 | 0 |
| P49 | 1042100 | | 786 | 360 | 806 | 365 | 0 | 0 |

*Figure 12 A sample of raw eye tracking gaze data*

As reported in the literature review, eyes are deemed to focus on a single point (a fixation) before jumping very quickly (a saccade) to the next point of focus. The reality is that eyes tremble, making tiny adjustments, during a fixation rather than remaining absolutely static (Holmqvist et al. 2011:22). Whether each reading from the eye tracker forms part of a fixation or a saccade is determined by the fixation filter applied to the raw data.

### 3.7.3 Fixation filter

The raw data for this research was processed using the Tobii I-VT fixation filter. For full details of the algorithm please refer to Olsen (2012) however, a brief explanation is included here. The I-VT (Velocity-Threshold Identification) Fixation filter, as its name suggest uses the speed of the eyes' movement from one reading to the next to determine whether each reading forms part of a fixation or not. The speed of eye-movement is measured in visual degrees per second. After combining the readings for the left and right eye, the speed at which the eyes are moving is calculated. When the speed of the eyes falls below a threshold, the eye is deemed to be fixating. Figure 13 illustrates the I-VT being applied to data. The dark red line indicates the location of the eyes on the vertical plane, the blue line indicates the location of the eyes on the horizonal plane. The green line indicates the speed at which the eyes are moving. When the speed of the eyes drops below the threshold of 30 visual degrees per second (indicated by a straight, bright red horizonal line) a fixation is deemed to be taking place. Tiny, short-lived increases in speed are permitted (as seen near the beginning of fixation two and the beginning of fixation 3) to allow for trembling movements without causing an end to the fixation. The fixation filter is a complex algorithm which considers not only speed but also other factors such as the time delay from the end of one fixation and the start of the next of the fixation and the difference in visual angle between fixations before deciding whether adjacent fixations should be merged or remain distinct from one another.

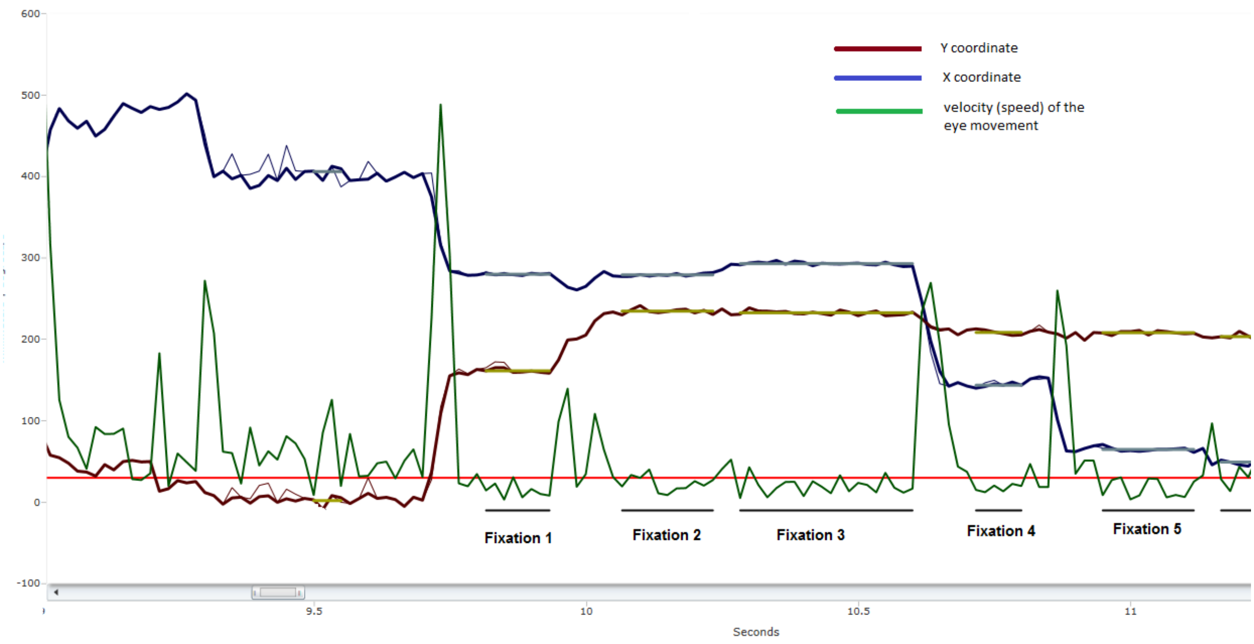



*Figure 13 Velocity chart illustrating raw eye tracking data*

After the raw data has been processed by the algorithm, the data is summarised into a single row of data per fixation, each with a start time, a duration and a coordinate expressed in pixels. The data is now 'fixation' data rather than raw eye tracking data and additional information about the location or area of interest (AOI) of the fixation can be applied and the recording can be broken into segments.

### 3.7.4 Segmenting the recordings

As detailed in section 2.9, the user interface had multiple 'pages'. To establish which text participants were reading it was necessary to divide each recording into segments according to the page being displayed on screen. As well as recording the eye-movements of participants, Tobii Studio also records every key stroke and mouse click during the recording. By checking every mouse click the researcher was able to identify mouse clicks which instigated a change of page and create a new segment. Segments of the recording were then allocated to a scene (Scene one = the interface displaying page one in the source text area, scene two = the interface displaying page two in the source text area and so on.)

### 3.7.5 The distribution of fixations across different areas of interest

The areas of the screen can be sub divided into smaller sections for analysis and coding. These sections are called areas of interest (AOI). The researcher divided the user interface into broad AOIs as illustrated in Figure 8: Screen instructions, Source text area, Diagram (please note that only pages two and four included diagrams), Move page controls, and finally Area for typing answer (this AOI is labelled 'Own work').

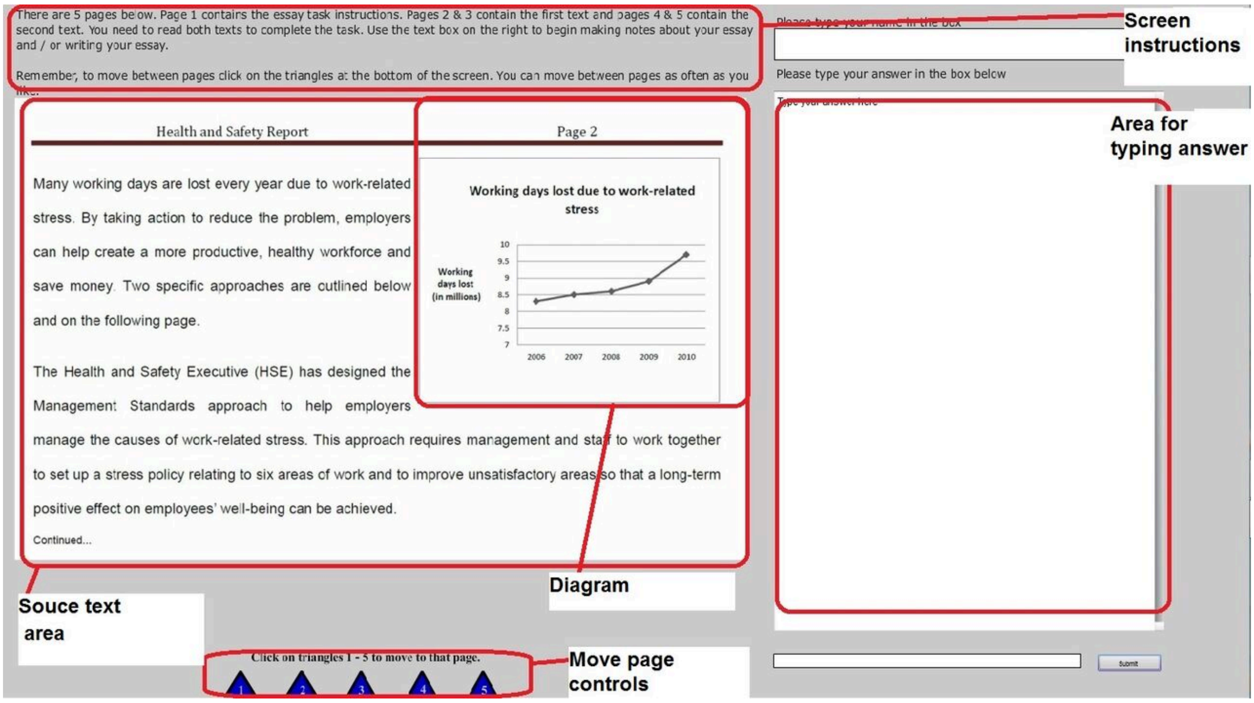



*Figure 14 User interface showing broad AOIs*

The eye tracking data was analysed to report the number of fixations and the total duration of fixations on each broad area of interest for each participant. However, as this research aimed to understand how participants read the source texts, the fixations on the source texts area were subjected to analysis in greater depth. The further levels of analysis involved breaking down the fixations into sentence level areas of interest (3.7.6) and considering the direction and distance between fixations (Section 3.7.8) as well as the patterns of fixation that might suggest the type of reading (3.7.11) taking place.

### 3.7.6 The distribution of fixations across sentences in the written source texts

The broad AOI 'source text area' contained the verbal (written) content. In the same way that the broad areas of interest were defined, it was possible to subdivide the source text area into smaller areas of interest which correlated with individual sentences. The researcher refers to these sentence level areas of interests as 'Narrow AOI'. Fixations on the narrow AOIs were analysed for evidence of reading activity.

Narrow AOI were only applied to source text pages 2,3,4 and 5. All other areas of the screen were excluded from analysis for reading activity. This is because fixation analysis to suggest reading is only applicable to fixations on text. In addition, the analysis for reading relied on calculations relating to the font size and line spacing of text. The text on the diagrams and the task rubric on page one did not conform to the same font size and line spacing as the verbal sources on pages 2-5. The same factors applied to the participants' own work; the font was much smaller with smaller distances between lines. Additionally, the area for typing participants responses scrolled to allow participants to use as much space as they wished for note taking and setting out their answer. The technique used to suggest reading from fixation data only works on static, not scrolling, text.

Therefore, only the fixations on verbal source texts on pages 2-5 were further subdivided into sentence level areas of interest. The number and duration of fixations on each sentence was considered as a percentage of fixations on the source texts. The total length, in words, of the two source texts was calculated and represented as 100 per cent of the source text length. The length of each sentence in words was also measured. The researcher calculated what percentage of the total source text length each sentence accounted for. The researcher then compared the percentage of fixations (both number of fixations and total fixation duration) on each sentence to the percentage of the whole source text that each sentence accounted for. For example, if a sentence contained 5 per cent of the total number of words, you could expect it to receive 5 per cent of the fixations if the reader were reading all sentences in the same way. This exercise allowed the researcher to judge whether some sentences attracted more of the participants attention than others when adjusted for length.

### 3.7.7 The distribution of fixations across sentences in the written sources according to sentence relevance

The analysis of individual sentences was supplemented with analysis of the sentences when grouped according to their relevance to the writing task. The researcher and two further expert raters rated each sentence according to the relevance of the idea unit as follows: 4-highly relevant (core ideas); 3-relevant (additional details); 2-less relevant for the task; 1-irrelevant. The total relevance score for all three raters was averaged to determine the relevance score for each sentence. (Kendall's W was run to determine if there was agreement between the three raters' judgements on the relevance of the sentences. The three raters statistically significantly agreed in their assessments, $W = .732$, $p < .0005$.) Raters were not asked to rate the page titles and these were excluded from this analysis.

The researcher then compared the percentage of fixations (both number of fixations and total fixation duration) on each group of sentences to the percentage of the whole source text that each group accounted for. Once again this enabled the researcher to consider whether some groups of sentences attracted more of the participants attention than others when adjusted for length.

### 3.7.8 Direction of movement and distance between fixations

In order to analyse the type of movement between fixations or to consider whether eye-movements formed a pattern which was indicative of reading several measures had first to be calculated.

The first factors to be calculated are the position of a fixation in relation to the previous fixation (whether it moved forward through the text or regressed backward) and the distance between those two fixations. Once the fixation filter (see section 3.7.3) has been applied to the raw data the data is condensed into an Excel spreadsheet with one line per fixation (Table 8).

*Table 8 An example of the Excel spreadsheet showing fixation data*

| Participant Name | Fixation Number | Scene Name | Time in ms | Fixation Duration | X coordinate in pixels | Y coordinate in pixels | Area of Interest |
|---|---|---|---|---|---|---|---|
| P35 | 1 | page 1 | 421615 | 67 | 218 | 243 | P1S1 |
| P35 | 2 | page 1 | 421681 | 83 | 273 | 236 | P1S1 |
| P35 | 3 | page 1 | 421798 | 83 | 306 | 249 | P1S1 |
| P35 | 4 | page 1 | 422165 | 117 | 350 | 246 | P1S1 |
| P35 | 5 | page 1 | 422331 | 133 | 404 | 256 | P1S1 |
| P35 | 6 | page 1 | 422831 | 83 | 470 | 231 | P1S1 |
| P35 | 7 | page 1 | 422948 | 100 | 525 | 240 | P1S1 |
| P35 | 8 | page 1 | 423615 | 83 | 569 | 245 | P1S1 |
| P35 | 9 | page 1 | 423766 | 85 | 625 | 235 | P1S2 |
| P35 | 10 | page 1 | 424537 | 135 | 279 | 254 | P1S1 |
| P35 | 11 | page 1 | 424823 | 250 | 691 | 253 | P1S1 |
| P35 | 12 | page 1 | 425090 | 183 | 746 | 247 | P1S1 |

From this data, using formulas in Excel, it is a simple matter to calculate whether a fixation is to the left or right and above or below the previous fixation. Referring to section 3.6.3, the distance in pixels between the centre of one line and the centre of the line above or below is 70 pixels. For example, if we consider Figure 15 , the numbers in blue indicate the fixation number from Table 8. Fixations 1-9 all progress forward through the text and that is reflected in the increasing value of their x coordinates and relative stability of their y coordinates. Although the y coordinates fluctuate a little between fixations 1-9, they do not decrease or increase sufficiently (plus or minus 35 pixels) to indicate a change to the line above or line below.

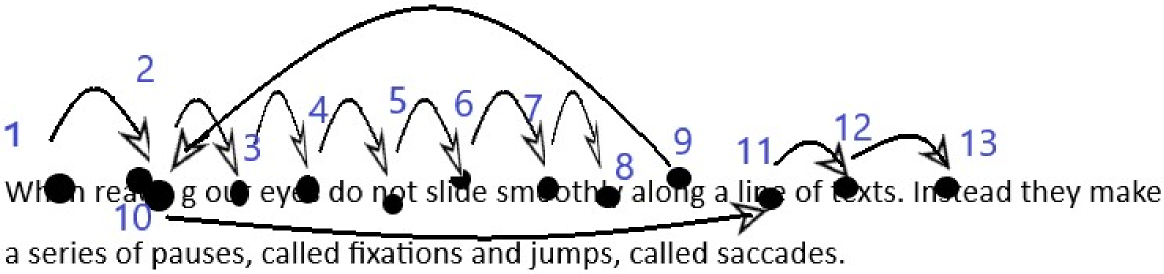



*Figure 15 Illustration of fixations*

Fixation 10, however, regresses back to an area of the text already read. This is evident from the 246-pixel reduction in the x coordinate, whilst maintaining relative stability of the y coordinate.

Therefore, after loading the data into an Excel spreadsheet, the first stage in determining the direction and distance of saccades is the addition of columns which calculate the difference between the current and previous fixation. A negative value in the ‘x difference’ column indicates a move to the left, a positive value a move to the right. The size of the ‘x difference’ value indicates how far left or right the eyes have moved. For example, on average characters occupied a space on screen equivalent to 11 pixels. Therefore, an ‘x difference’ of between 1 and 180 indicates a move to the right of 16 characters or less. A value over 180 indicates a move to the right of more than 16 characters.

Similarly, a negative value in the ‘y difference’ column indicates a move up the page, a positive value a move down. The size of the ‘y difference’ value indicates how far up or down the eyes have moved. A ‘y difference’ figure of less than minus 35 pixels indicates a move to a line above, a figure of greater than 35 pixels indicates a move to a line below. A ‘y difference’ value between -35 and +35 indicates that the fixation remained on the same line as the previous fixation. Therefore, after the fixation data, as per the example shown in Table 8, was loaded into an Excel spreadsheet the first stage involved adding columns which calculated the difference between the x coordinate of the previous fixation and the current fixation and the difference between the y coordinate of the previous fixation and the current fixation.

*Table 9 Fixation data with x and y coordinate movement calculations.*

| Participant Name | Fixation Number | Scene Name | Time in ms | Fixation Duration | X coordinate in pixels | Y coordinate in pixels | Area of Interest | X movement | Y movement |
|---|---|---|---|---|---|---|---|---|---|
| P35 | 1 | page 1 | 421615 | 67 | 218 | 243 | P1S1 | | |
| P35 | 2 | page 1 | 421681 | 83 | 273 | 236 | P1S1 | 55 | -7 |
| P35 | 3 | page 1 | 421798 | 83 | 306 | 249 | P1S1 | 33 | 13 |
| P35 | 4 | page 1 | 422165 | 117 | 350 | 246 | P1S1 | 44 | -3 |
| P35 | 5 | page 1 | 422331 | 133 | 404 | 256 | P1S1 | 54 | 10 |
| P35 | 6 | page 1 | 422831 | 83 | 470 | 231 | P1S1 | 66 | -25 |
| P35 | 7 | page 1 | 422948 | 100 | 525 | 240 | P1S1 | 55 | 9 |
| P35 | 8 | page 1 | 423615 | 83 | 569 | 245 | P1S1 | 44 | 5 |
| P35 | 9 | page 1 | 423766 | 85 | 625 | 235 | P1S2 | 56 | -10 |
| P35 | 10 | page 1 | 424537 | 135 | 279 | 254 | P1S1 | -346 | 19 |
| P35 | 11 | page 1 | 424823 | 250 | 691 | 253 | P1S1 | 412 | -1 |

### 3.7.9 Algorithm to detect reading

Once the columns which calculate the difference between the current and previous fixation have been added it is possible to begin to use Excel formulas to test each fixation to see whether it meets a range of criteria.

### 3.7.10 Breaks in reading

Having explained the basic principle of how distance and direction is calculated, the researcher stresses that it is only relevant to consider the direction and distance between fixations when reading is proceeding actively. Therefore, consideration also needs to be given to the time that has elapsed between fixations to identify times when participants may have closed their eyes to think or gazed around the room, indicating that reading has paused. On average saccades from one reading fixation to the next last 40-60 milliseconds (Rayner, Foorman, Perfetti, Pesetsky and Seidenberget (2001), however the overwhelming majority of blinks last between 100ms and 500ms (Holmqvist et al. 2011:325). Therefore, a pause of less than 500ms may well be caused by a blink, however pauses longer than this are likely to indicate a pause of some kind. Therefore, in addition to calculating the differences between fixation coordinates, it was also necessary to calculate the time lapse from the end of a fixation to the start of the next fixation. This was achieved by subtracting the end time of the previous fixation (time in ms column plus fixation duration in Table 9) from the start of the current fixation (time in ms). Therefore, on the Excel spread detailed in Table 9 another column was added which contained the time lapse in milliseconds from the end of one fixation to the start of the next. Another additional column used a formula to test the time lapse from the end of a fixation to the start of the next fixation. If the formula detected a delay of more than 500ms the word “RESET” was entered in the cell.

### 3.7.11 Analysis of the type of movement between fixations

A pause between fixations was not the only reason to exclude a fixation from analysis of the type of movement between fixations. It was only relevant to consider the type of movement between fixations once three criteria had been met. Therefore, as well as columns which calculated the movement between fixations (as detailed in Table 9 ) additional columns and formulas were used to test whether each fixation

1. was on the source text (not elsewhere on the page)
2. occurred less than 500ms after the end of the previous fixation
3. occurred on the same page as the previous fixation.

Fixations which did not meet these criteria were marked. In other words, the formulas resulted in the phrase “OFF TEXT” or “RESET” appearing in a cell on the relevant row in the spreadsheet.

Having marked fixations which were not eligible to be considered as potential reading fixations more formulas were used to give each fixation a score or numeric value according to whether they had moved a short distance forward through the text (less than 16 characters), a longer distance forward through the text (more than 16 characters), regressed a short distance (on the same line or to the line above) or finally, regressed to more than one line above. The reasons for selecting a 16-character demarcation are discussed in section 2.4.1 of the literature review. If the word “RESET” or “OFF text” appeared in the row, no value was assigned.

Thus far, using an Excel spreadsheet each individual fixation has been marked with either a word or value. The next section explains how, also using Excel, formulas were used to label fixations according to the patterns formed by series of fixations.

### 3.7.12 Method for detecting careful reading

The system for detecting whether reading was taking place, and if so what type of reading, is also achieved by using a series of formulas in an Excel spread sheet. These formulas categorise each eligible fixation according to the change in X and Y coordinates between this fixation and the last fixation and allocate it a score accordingly. By accumulating the score of eligible fixations until the next "reset" allows broader patterns to be established.

For ease of understanding, the series of tests and categories applied to the data by the Excel formulas has been divided into four phases which are explained below and illustrated with a series of flow charts (Figure 16 - Figure 19). The combination of Excel formulas used will, henceforth, be referred to as an algorithm.

#### 3.7.12.1 Stage 1- checking for potential reading

Stage one, illustrated in Figure 16, excludes certain types of fixations that should not be analysed as potential reading fixations (as described in section 3.7.9 and 3.7.10). Firstly, it screens out fixations that are not focusing on text; they cannot be reading. Secondly it identifies whether a fixation is the first fixation on text after a change of page. Thirdly, it checks if a pause has occurred between this fixation and the last fixation on this page. It is not possible, in the case of a change of page, to analyse a fixation in relation to the previous fixation. In the case of a return to text after a pause, it is not logical to consider a fixation in relation to the previous fixation as that 'phase of reading' can be considered to have come to an end. In all three cases, the fixation is marked as a 'reset' and stage two is bypassed, jumping to stage 3. If none of the conditions described above have arisen, the fixation is left unmarked and passes to stage 2.

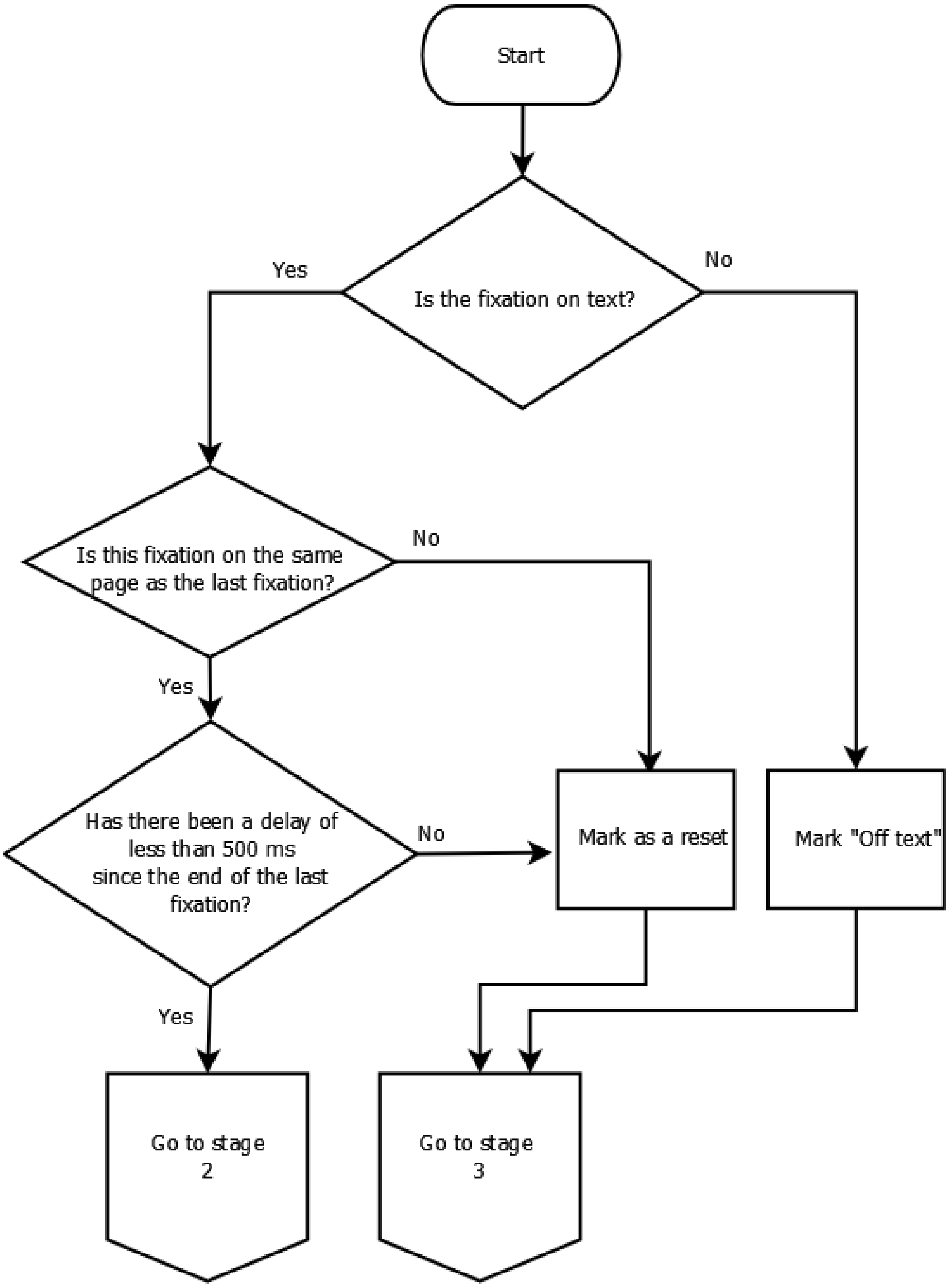



*Figure 16 Stage one of the fixation analysis process*

#### 3.7.12.2 Stage 2 – Distance and direction of movement

Stage two, illustrated in Figure 17, analyses the distance and direction moved between the last fixation and this fixation.

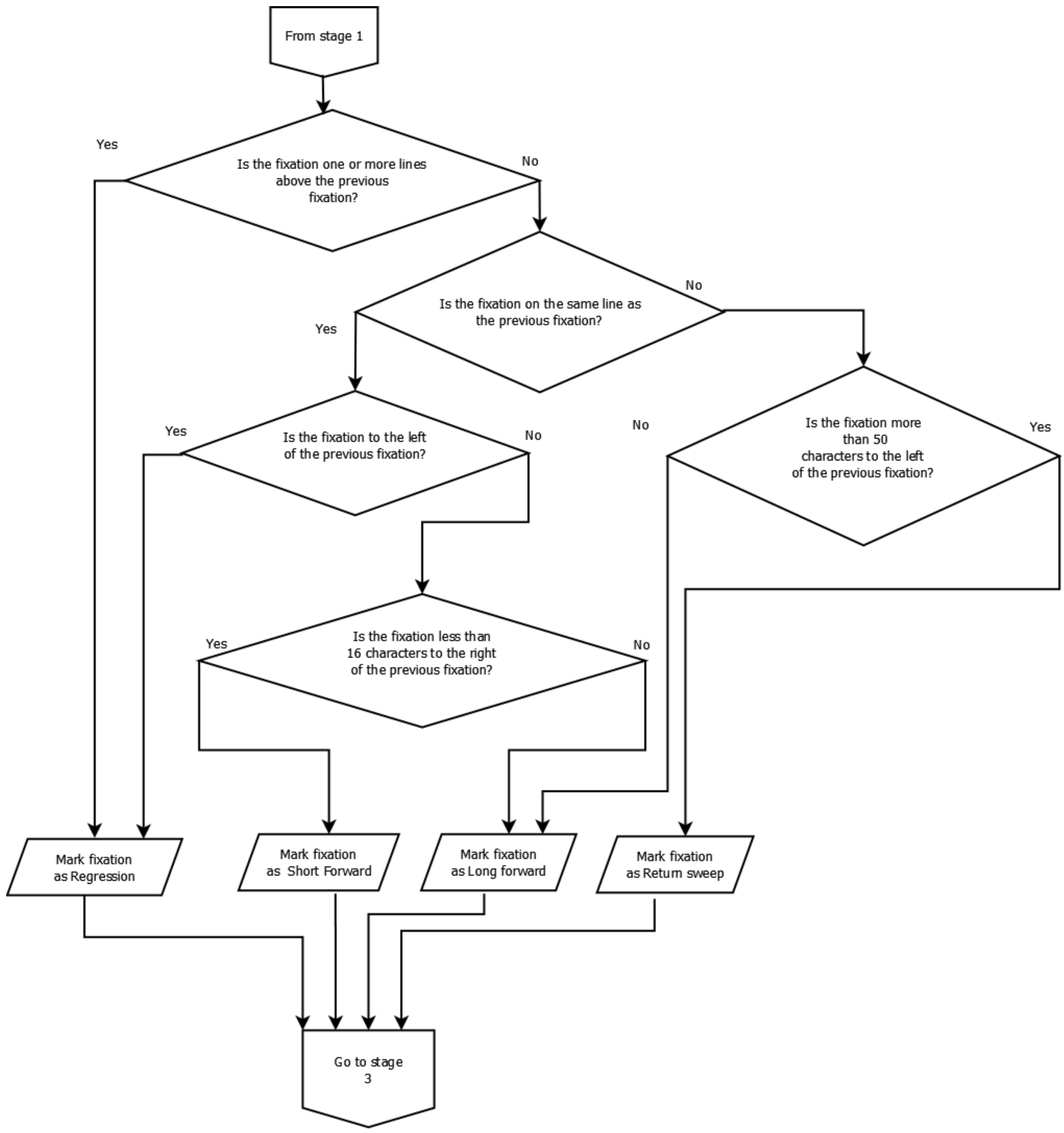



*Figure 17 Stage two of the fixation analysis process*

As explained in section 3.7.8, the coordinate (in pixels) of each fixation can be used to ascertain whether the reader has progressed further through the text or regressed. If the fixation has not been marked as “ineligible” in stage 1, stage 2 classifies the fixation according to whether it has progressed or regressed through the text. The classifications applied to fixations during stage two are used in stage three to determine whether the scan path pattern suggests that careful reading is taking place. Stage 2 labels fixations as one of four different types of movement: regression, short forward, Return sweep or long forward. Each type of movement is now discussed in turn.

Regressions are any movement in which the current fixation focuses on an earlier part of the text than the previous fixation. Therefore, any fixation which moves up a line must be a regression, likewise any fixation which is on the same line and left of the previous fixation must be a regression.

Short forward moving fixations must occur on the same line and slightly to the right (within 16 characters, see discussion in section 2.4.1) of the previous fixation. These fixations are seen as representative of careful reading.

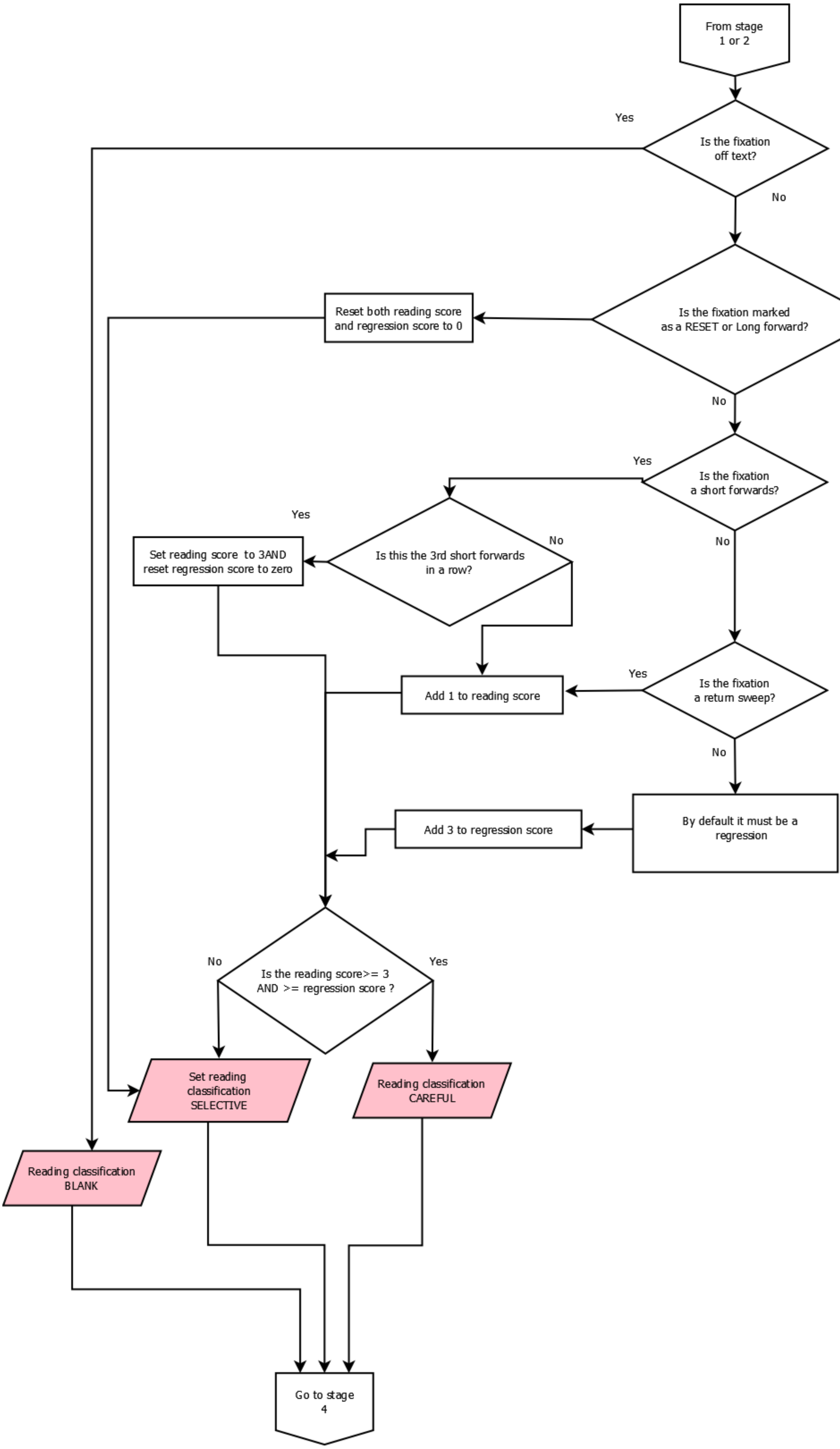
Return sweeps, which move the reader from the end of one line and to the beginning of the next are another form of short forward. Therefore, any fixation which moves to the line below and more than 50 characters to the left is labelled as a return sweep and scored in the same way as a short forward.

Several different types of movements can be labelled as long forward. Firstly, any fixation which more than more than 16 characters to the right on the same line is labelled as a long forward. Also, any fixation which is on any line below the previous fixation and is not labelled as a return sweep is classified as a long forward.

Long forward movements are considered to signal the end of careful reading. Long forward movements indicate that the participant is moving their focus to another area of the text and this type of movement is much more likely to represent scanning or skimming activities. After the fixations have been labelled the process moves on to stage three.

#### 3.7.12.3 Stage 3 – Does the pattern suggest careful reading?

Stage 3, illustrated in Figure 18, uses a pooled scoring method, an adaptation of the method used by Campbell and Maglio (2001) to look for patterns over a series of fixations.



From stage
1 or 2
Is the fixation
off text?
Yes
No
Is the fixation marked
as a RESET or Long forward?
Reset both reading score
and regression score to 0
No
Is the fixation
a short forwards?
Yes
No
Is this the 3rd short forwards
in a row?
Yes
No
Set reading score to 3AND
reset regression score to zero
Is the fixation
a return sweep?
Yes
No
Add 1 to reading score
By default it must be a
regression
Add 3 to regression score
Is the reading score>= 3
AND >= regression score ?
No
Yes
Set reading
classification
SELECTIVE
Reading classification
CAREFUL
Reading classification
BLANK
Go to stage
4

*Figure 18 Stage three of the fixation analysis process*

Stage three scores each fixation according to the properties of the movement between the last fixation (assessed and categorised in stage 2) and uses a simple arithmetic calculation to decide whether a sufficient number of the 'right' type of fixations have occurred in conjunction with a limited proportion of regressions.

Holmqvist et al. (2003) suggested that for reading to be taking place, a minimum of three fixations which fit the physical properties of reading fixations need to have taken place. In addition, those fixations need to be moving a short distance to the right, or a return sweep for careful reading to be taking place. Finally, the number of regressions taking place should account for no more than a limited proportion of fixations.

The researcher found that the regression rates in this study were much higher than those reported in the literature. As discussed in section 2.4.1 much of the literature suggests a regression rate of 10-15 per cent is typical when reading. However, when the researcher set the algorithm in stage three to allow no more than one regressive fixation per six short forward moving fixations (15 per cent) very little reading activity was classified as 'careful reading'. During the retrospective think aloud participants were asked to describe their reading goals. Participants frequently reported periods when they were reading carefully to gain a complete understanding of the text. The researcher noted that during these periods the participants made high numbers of regressions. Carver (1992) suggests that reading to remember information for later use in answering questions is slower with a higher incidence of rereading than 'rauding' or reading to understand. For these reasons the researcher experimented with adjusting the permissible regression ratio in the algorithm and found that a ratio of one regression to three careful reading fixations resulted in high levels of agreement between the algorithm and the RTA reporting.

Stage three assigns fixations a numeric value according to the category marked in stages one and two. Values were assigning as detailed in Table 10.

Table 10 Values assigned to fixation types

| Fixation label | Value assigned | Rationale |
|---|---|---|
| Regression | Adds 3 to regression score | Regressions can only occur at a ratio of 1 regression for every 3 short forward moving fixations during careful reading |
| Short forward | Adds 1 to the reading score | Short forward movements are likely to form part of careful reading |
| Long forward | RESET | Long jumps forward through the test signal an end to that episode of careful reading |
| Return sweep | Adds 1 to the reading score | Return sweeps are effectively a short forward movement and are likely to form part of careful reading |

The classification of careful reading can only occur after a minimum of three fixations on text. This is because a minimum of three short forward moving fixations must occur before careful reading can be said to be occurring (Holmqvist et al., 2003). Therefore, the first and second fixation on text will always return a classification of selective reading because it is impossible for a reading score of three to have been accumulated yet. Only once the reading score equals three is it possible for a classification of careful reading to occur. Even when the reading score equals three or more the regression score must also not exceed the reading score. Increasing the regression score by three every time a regression occurs, whilst only increasing the reading score by one every time a short forward moving fixation occurs ensures the ratio of three short forward to one regression is maintained during periods classified as careful reading. Three short forward moving fixations in a row are considered to signal a new episode of careful reading therefore, if there are three successive short forward fixations the regression total is reset to zero, and the reading score is set to three ensuring that a new episode of careful reading is established.

#### 3.7.12.4 Stage 4 – Does this phase of reading extend beyond a single sentence?

Stage four considers whether each fixation within a phase of reading has remained within a single sentence or crossed sentence boundaries into a new sentence and is illustrated in Figure 19.

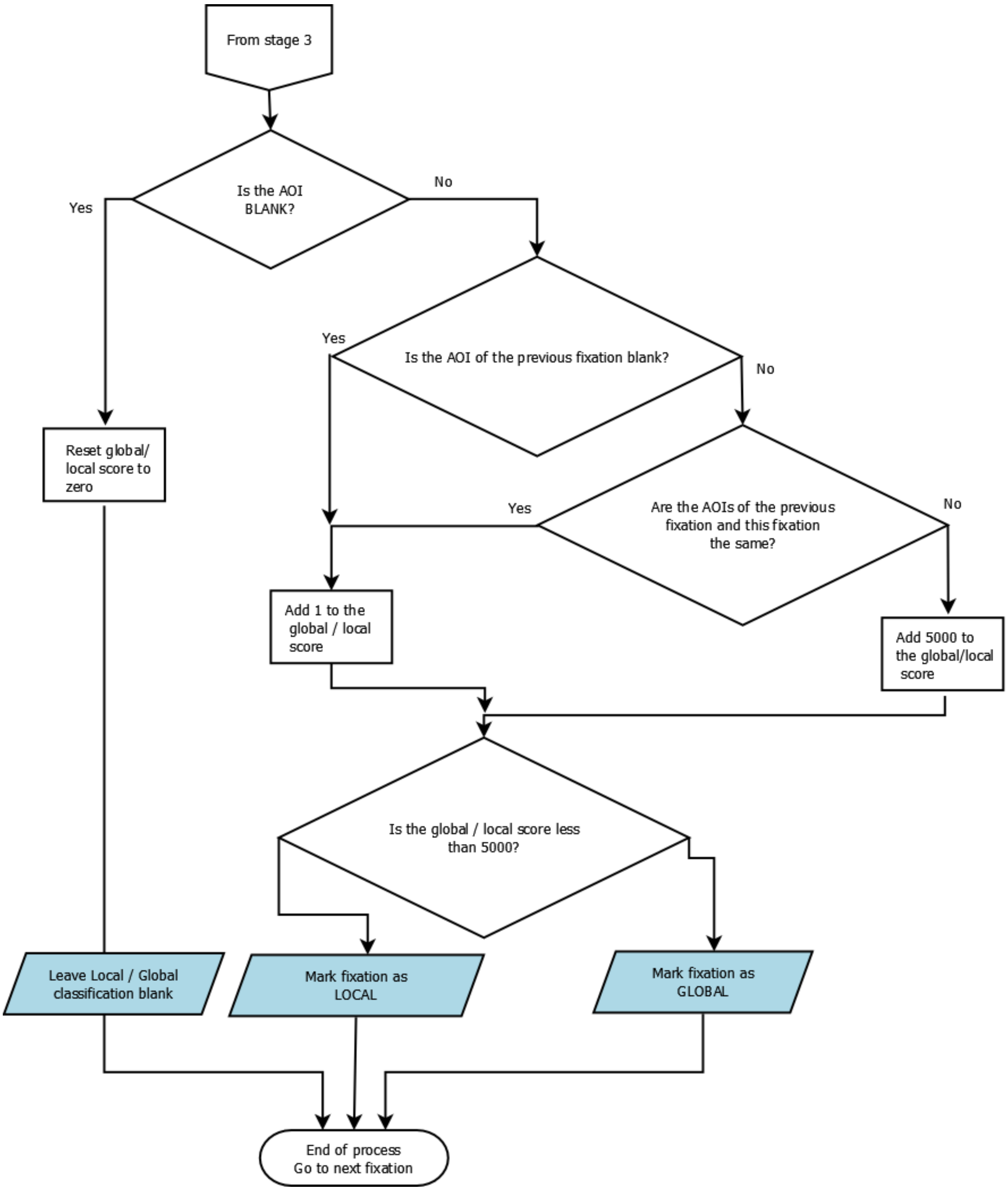



*Figure 19 Stage four of the fixation analysis process*

Khalifa and Weir (2009:59) distinguish between careful reading at sentence level (local reading) which focuses on understanding the sentence proposition in isolation from the rest of the ideas in the text. Khalifa and Weir suggest that once reading extends beyond sentence boundaries (global reading) readers become engaged in 'integrating new information into a mental model and perhaps finally creating a discourse level structure for the text' (Khalifa and Weir, 2009:59).

The stage four process for assigning a local or global considers the AOI of each fixation and checks whether this fixation and the preceding fixation occurred in the same sentence. The first stage is to check for an AOI (sentence indicator). If the AOI is blank, the fixation was not on the written source text and the global /local score is reset to zero and the global / local classification is left blank.

If there is an AOI for this fixation, the next stage is to check if the previous AOI was blank. If the previous fixation's AOI was blank, then this fixation must necessarily be the first fixation in an episode and one is added to the global/local score (The global /local score is used later to determine whether to apply a global / local label).

If the last fixation AOI was not blank, the AOI of this fixation and the previous fixation are compared. If the AOIs are the same, then one is added to the global / local score. If there has been a change of AOI, 5000 is added to the global / local score.

Now that all the possible combinations have been scored, the global / local score is tested. If the global / local score is less than 5000, a label of local is applied. If the global / local score is 5000 or more a global classification is applied.

It is necessary to use a scoring system rather than simply comparing the current fixation to the previous fixation because otherwise, as soon as two successive fixations occurred in the same sentence, the global/local label would return to local. By using this system of scoring each fixation, every fixation in the current episode of reading is taken into consideration and once the episode leaves the sentence in which the episode of reading started, the global label will be applied and maintained. The only disadvantage of this system is that if 5000 successive fixations in an episode stayed within the same sentence the algorithm would falsely apply a global label. However, this is extremely unlikely, given that most participants only generated about 7000 fixations in total.

The global and local classifications, when applied to fixations that are marked as Careful reading, correspond to Khalifa and Weir's 2009 classification for Careful global reading and Careful local reading. However, the global and local classifications when applied to reading classified as 'selective' do not correspond directly to Khalifa and Weir's Expeditious global reading and Expeditious local reading. The reasons for this, and the disparities between the classifications, are discussed further in section 5.2.2.2 of the discussion chapter.

After processing through stages one to 4, each fixation has been given a reading classification (indicated in the red output parallelograms in Figure 18) and every instance of reading has been marked as local or global (indicated by the blue output parallelograms in Figure 19).

This results in every fixation on the written source texts being allocated to one of four types of reading as follows:

1. **Careful local** reading which has a ratio of at least three short forward moving fixations for every regressive fixation. This category will not tolerate any long forward jumps or any long regressions to the previous sentence.
2. **Careful global** reading therefore, still requires the ratio of three short forward to every regression but will allow the reader to progress to the next sentence but only by means of a short forward or a return sweep. Long regressions which jump back to an earlier part

of the text would be permitted as long as the regression rate remains within permitted limits. A long glance back to an earlier part of the text to resolve confusion could be part of careful global reading. However, if the reader then used a long jump forward through the text to return to the previous section of text this would prompt a reset, therefore a minimum of three short forward moving fixations would need to occur to return the reading classification to careful again.

3. **Selective local** reading occurs when an episode of reading does not meet the ratio of at least three short moving fixations to every regression but stays within a single sentence. For example, there may be a very high amount of regressions or there may be long forward saccades but within a single sentence.
4. **Selective Global** reading occurs when an episode does not meet the requirements for careful reading and extends beyond a single sentence. For example, the forward / regression ratio for careful reading is not met, or the episode includes long forward moving fixations and the episode crossed from one sentence into another.

### 3.7.13 Analysis of the RTA data

Following the data collection, the RTA recordings were transcribed. Three of the RTA had failed to record for technical reasons so there were 27 recordings. Of the remaining 27 recordings, one recording had experienced technical difficulties and so was very short. In total the RTA recordings generated 46,068 words of transcription. Unfortunately, all of the recording failures occurred with the Y1 undergraduate participants and therefore there was a considerable imbalance between the amount of data generated for the Y1 undergraduate group and the year three/ postgraduate group (see Appendix 9). For this reason, results of the RTA analysis for RQ2 were also reported as percentages for each group to facilitate comparison.

Table 11 Examples of RTA transcripts coded according to cognitive phases of reading into writing as defined by Chan (2013)

| **Cognitive Process Coding Category** | **Definition** | **Example of dialogue from transcripts** |
|---|---|---|
| Task Representation **TRP** | The writer's thoughts about the demands of the task. | **'So, I defined in my head, okay problem I need to find what the problem is but also I saw there was data. I would look at that to find what's the problem in the data. In the discussion I did the same with the causes and solutions '.** (P12, line 17) |
| Macro-planning **MAP** | The writer's thoughts about how they might respond to the task. | **'I've taken the question and translated it on to the page as a guide, so I know what I'm looking for.'** (P12, line 12) |
| High level reading processes **HLR** | The writer's thoughts about reading the source texts. | **'I think because I was trying to read the causes first, and I knew the back end of it was the solution, so I thought I would read the second article again, the start of it…'** (P37, line 66) |
| Connecting and generating **C&G** | The writer's thoughts about how to make connections between the source texts and / or their own ideas. | **'Do I really need to focus on what has been written here or could I add some of my general knowledge in this kind of problem?'** (P45, line 92) |
| Organising **ORG** | The writer's thoughts about how to order their ideas. | **'It just didn't sit well with me because I realised I had said there with two different approaches but actually I'd only written one paragraph, so I needed to split it up... .... I split it up and changed it around again'** (P41, line 173) |
| Micro planning **MIP** | The writer's thoughts about composing at local level. | **'I think it was because I was trying to work out how to word it. How I could start…'** (P37, line 74) |
| Low level monitoring & revising **LMR** | The writer's thoughts about making changes at local level. | Making changes to vocabulary **'If there were words that are too similar to each other in the same sentence. I think I have workers and the employers work together. Then I changed workers…to staff. '** (P42, line 166) |
| High level monitoring and revising **HMR** | The writer's thoughts about making changes at a global level. | Actively considers the reader when editing **'What would I want it to read like? Have I proved my point? Has it come across well?'** (P50, line 142) |

The transcriptions were coded using the cognitive phases defined by Chan (2013). Examples of the application of coding can be seen in Table 11 . Therefore, the researcher studied the transcripts of the RTA, whilst watching the accompanying replay on screen. Each comment from the participants was recorded on a spread sheet, with the participant number, transcript line number, time stamp from the replay. As the comment was logged on the spread sheet, the researcher categorised the comment according to Chan's (2013) cognitive processes as can be seen in Table 11.

After this all the comments were reviewed by the researcher and an academic colleague to identify emerging sub-themes. The sub-theme categories can be seen in Table 12, along with examples taken from the spread sheet. The researcher coded all the data according to the subthemes and her colleague checked a selection (approximately 10 per cent) to review categorisation. (The rate of agreement using Cohen's kappa *k*=.942)

*Table 12 Themes emerging from RTA categories*

| **Cognitive processes** | **Example** |
|---|---|
| **Task Representation** | |
| Genre | 'I was wondering if I have to use academic or formal, or I can use contractions some things I came back to first pages, and I saw that your lecturer asked you to write and I thought that has to be academic or formal.' |
| Goal setting | P26 reports wanting to 'test' herself to 'see if [her] language skills have improved'. |
| Recall of task | Uses note taking to reinforce memory 'It's easier for me to remember things when I have them written down instead of just from memory' |
| Rhetorical function | 'So, I defined in my head, okay problem I need to find what the problem is but also I saw there was data. I would look at that to find what's the problem in the data.' |
| Time constraints | 'I have this mental budget in terms of time in my mind. I have this minutes to do reading…' |
| **Macro-planning** | |
| Finding content | 'normally if I was reading something that's on paper I'd highlight it. So, it's what I would be highlighting up' |

| | |
|---|---|
| Structure | 'Basically, it [the task instructions] gave me the organisation of the text. I was reading the first three because I was thinking that's basically how we organise the text within.' |
| **High level reading processes** | |
| Deeper understanding | .' I was like, 'Okay. I will read again'. Just to understand more.' |
| Extracting info | 'I just think there I was using figures from the graph…so I wanted to make sure I was writing those down correctly ' |
| Failure to understand | Repetitious reading of sentence because 'it didn't go in straightaway'. |
| Gist | 'I was just trying get the main message from the article. I don' t feel like I needed to read that at the time.' |
| Lack of engagement | 'Do you know when you're reading something, and you're not thinking about it, you're just reading it and that's what happens to me.' |
| Search reading | 'I scanned until I found it, then I make quick notes on it. I think I read through it all again, until I found all of it.' |
| Targeted | 'I read with a purpose to find this particular information for this specific task that has been given to me.' |
| **Connecting and generating** | |
| Familiarisation | 'I probably have to understand it for a little longer to write about it because, I don't know, I'm not totally familiar with it.' |
| Forming opinions | 'So, I have a vague idea. I can start forming my own opinions if I needed to'. |
| Generating new ideas | 'I just need to use my-- paraphrasing thoughts and just creating new ones. Mine, a little bit, not copying anyone else's' |
| Makes links between text ideas | 'I was just checking, reading it. Then, thinking of a way how I was going to bring both of them into one sentence. Introduce them both at the same time.' |
| Own experience | 'Then that kind of led up into having work-related stress, like they might not-- you know, kind of go to them and say 'oh I'm feeling a bit stressed out. Is there anything you could help—' they might not have done that. They might have just taken a day off work…' |
| **Organising** | |
| Categorise | 'I was thinking for quite a while if any of these had fallen into any category, if there's anything due to management, poor management for example, or in one of the articles, it mentions lack of communication. I was trying to see if there's anything in common in these causes of stress. ' |

| | |
|---|---|
| **Micro planning & translating** | |
| Prompting | *Intense period of reading a paragraph, a lot of reading and regressing back.* 'I think it was because I was trying to work out how to word it. How I could start…' |
| Reflecting | *She suggests that she needs to 'gaze' elsewhere when she is trying to compose* 'But I need to close my eyes or just look somewhere else just to focus.'. |
| Rewording | 'I think I wrote the words and then I realised I'd already said it, I didn't want to repeat it, so I was thinking about what I could put instead of it.' |
| **High-level monitoring and revising** | |
| Amount of info | So, it's -- yes, just like re wording it and then maybe adding more information… that is not enough information.' |
| Coherence / cohesion | *Moves sentence and restructures paragraph because* 'I thought the information was more relevant to the previous paragraph.' |
| Improving expression | 'No, there's definitely a much better way to write that'. |
| **Low level monitoring & revising** | |
| General | **Interviewer:** What were you're looking for?<br><br>**Interviewee:** Mistakes mainly |
| Grammar | 'Once it's down then, go back. Then, make it fit into the sentence.' |
| Plagiarism | 'Also, you can see at the start of it…in the part 'not more than five words' as well, so then I go back and double check and make sure it wasn't over the word count.' |
| Spelling | 'This spelling 'through', so that's why I kept…' |
| Vocabulary | 'I'm trying to use words that are more professional.' |
| | |

Where differences of opinion emerged, (two cases) the section of the RTA was replayed, reviewed and discussed so that agreement was reached. Other comments in these categories were subsequently reviewed to check for other differences of opinion, but none emerged.

The number of comments relating to each category and sub-category was then calculated. To facilitate easy comparison between the two groups the number of comments recorded per group in each category and subcategory was presented as a percentage of the group total alongside the actual numbers.

### 3.7.14 Analysis of the questionnaire data

The questionnaire was exploratory in nature with the aim of providing a broader context that would assist with the interpretation of the eye tracking data. With only 30 participants it was not relevant to conduct inferential statistics. Instead descriptive statistics (means and standard deviations) were calculated for each item. The results were then compared between the groups for RQ2 and RQ3.

## 3.8 Summary of the data analysis

RQ1 analysed the data from all the participants using eye tracking data on all parts of the screen, RTA data and questionnaire data to investigate the full range of processes suggested by Chan (2013). In addition, eye tracking data on the written source texts was analysed in terms of the fixations on individual sentences, sentences grouped by relevance, direction and distance moved between fixations and the patterns formed by fixations that suggested reading. Table 13 below summarises the data analysed for RQ1.

*Table 13 Summary of data analysis for RQ1*

| RQ | Research Method | Cognitive Process investigated | Measure | Data set |
|---|---|---|---|---|
| RQ1 : All participants | Eye tracking data | All | Number of Fixations<br>Fixation duration<br>Mean duration per fixation | Fixations on all parts of the screen |
| | | | | Fixations on all parts of the screen. Data subdivided into fixations per AOIs |
| | | Reading | Number of Fixations<br><br>Fixation duration | Data set limited to fixations on written source text. Data subdivided into fixation per sentence. |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided into relevance groups. |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided by direction and distance moved between fixations |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided by reading type |
| | RTA data | All | Number of comments<br>Qualitative analysis of comments | Whole task |
| | Question-naire data | All | Mean score on Likert scale | Whole task |

Therefore, RQ2 compared the data from the Y1 and Y3+ participants using eye tracking data on all parts of the screen, RTA data and questionnaire data to investigate the full range of processes suggested by Chan (2013). In addition, eye tracking data on the written source texts was compared in terms of the fixations on individual sentences, sentences grouped by relevance, direction and distance moved between fixations and the patterns formed by fixations that suggested reading. Table 14 summarises the data analysed for RQ2.

*Table 14 Summary of the data analysis for RQ2*

<table>
<tr><th>RQ</th><th>Research Method</th><th>Cognitive Process investigated</th><th>Measure</th><th>Data set</th></tr>
<tr><td rowspan="8">RQ2 :Y1 compared to Y3+</td><td rowspan="6">Eye tracking data</td><td>All</td><td rowspan="2">Number of Fixations<br>Fixation duration<br>Mean duration per fixation</td><td>Fixations on all parts of the screen</td></tr>
<tr><td>All</td><td>Fixations on all parts of the screen. Data subdivided into fixations per AOIs</td></tr>
<tr><td>Reading</td><td rowspan="4">Number of Fixations<br>Fixation duration</td><td>Data set limited to fixations on written source text. Data subdivided into fixation per sentence.</td></tr>
<tr><td>Reading</td><td>Data set limited to fixations on written source text. Data subdivided into relevance groups.</td></tr>
<tr><td>Reading</td><td>Data set limited to fixations on written source text. Data subdivided by direction and distance moved between fixations</td></tr>
<tr><td>Reading</td><td>Data set limited to fixations on written source text. Data subdivided by reading type</td></tr>
<tr><td>RTA data</td><td>All</td><td>Number of comments<br>Qualitative analysis of comments</td><td>Whole task</td></tr>
<tr><td>Question-naire data</td><td>All</td><td>Mean score on Likert scale</td><td>Whole task</td></tr>
</table>

Therefore, RQ3 compared the data from the five lowest scoring and five highest scoring participants using eye tracking data on all parts of the screen, RTA data and questionnaire data to investigate the full range of processes suggested by Chan (2013). In addition, eye tracking data on the written source texts was compared in terms of the fixations on individual sentences, sentences grouped by relevance,  direction and distance moved between fixations and the patterns formed by fixations that suggested reading.

*Table 15 Summary of the data analysis for RQ3*

| RQ | Research Method | Cognitive Process investigated | Measure | Data set |
|---|---|---|---|---|
| RQ3: 5 highest scoring ***compared to*** 5 lowest scoring | Eye tracking data | All | Number of Fixations Fixation duration Mean duration per fixation | Fixations on all parts of the screen |
| | | All | | Fixations on all parts of the screen. Data subdivided into fixations per AOIs |
| | | Reading | Number of Fixations Fixation duration | Data set limited to fixations on written source text. Data subdivided into fixation per sentence. |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided into relevance groups. |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided by direction and distance moved between fixations |
| | | Reading | | Data set limited to fixations on written source text. Data subdivided by reading type |
| | RTA data | All | Number of comments Qualitative analysis of comments | Whole task |
| | Questionnaire data | All | Mean score on Likert scale | Whole task |

This concludes the reporting of the methodology for the study. Next the findings of the study will be presented.

# 4 Findings

## 4.1 Introduction

This chapter will report the data that was generated using the methodology described in the methodology chapter in response to each of the three research questions and explain what the analysis of the data might imply. As an aide memoire the research questions are:

1. What are the characteristics of reading during an academic reading into writing test task?
2. What are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?
3. What are the similarities and differences between the way high and low scoring participants tackle an academic reading into writing test task?

The purpose of RQ1 is to consider whether there are any characteristics of reading behaviour which are common to all readers as they progress through the task. The purpose of RQ2 and RQ3 is to consider whether differences emerge between groups of participants (first year and third year / postgraduate participants for RQ1 and low and high scoring participants for RQ3). The data included three types of data: eye tracking data, retrospective think aloud (RTA) data and questionnaire data. The eye tracking data was used to answer RQ1, drawing on the results of the retrospective think aloud data where appropriate, to assist with interpretation of the results. Data from all three research methods was compared between groups to answer RQ2 and RQ3. Section 4.2reports the data relating to RQ1, section 4.3reports that data relating to RQ2 and section 4.4 reports the data relating to RQ3.

## 4.2 RQ1 What are the characteristics of reading when reading to write?

To characterise the reading behaviour of 30 participants on the reading into writing test task eye tracking equipment was used to record their eye-movements during the test and the resulting data was analysed. Eye-movements during reading have been described in depth in the literature review but it is perhaps useful to include a brief reminder of the key points here. When we read, our eyes make a series of jumps along the line interspersed with pauses. The jumps are called *saccades* and the pauses are called *fixations*. Not every word is fixated, with high frequency words and predictable words more likely to be skipped.

According to Rayner et al. (2012) during fixations, which last typically last about 250 milliseconds (a quarter of a second), although the range can vary from just 50 milliseconds to 550 milliseconds, the eye remains relatively still. The eye then moves extremely quickly during the saccade to the next fixation. Typically, saccades last just 40 milliseconds and during saccades no visual information is registered; the eye is moving too quickly. On average, reading saccades move forward 8-character spaces although both saccade distance and fixation duration are known to be influenced by both text difficulty and reading ability. More challenging texts elicit more densely packed and longer fixations. Skilled readers will make longer saccades between fixations and their average fixation duration will be shorter than less skilled readers.

In a bid to characterise reading during the hour long reading into writing task the eye tracking data was analysed in terms of:

a) The distribution of fixations across the six different areas of interest (AOI). The areas were referred to as broad areas of interest and included the question rubric, the verbal source texts, the diagrams, own written work, move page controls and an ‘outside any’ category for any AOIs not included in one of the aforementioned AOIs. The methodology for this was described in 3.7.5 and the findings are reported here in section 4.2.2.

b) The distribution of attention across each of the different sentences in the source texts. The results are reported for both individual sentences and for sentences when grouped according to their relevance to the writing task. The methodology for this was described in section 3.7.6 of the methodology chapter and the findings are reported here in section 4.2.3.

c) The type of movement between fixations on the source texts; short forward movements, long forward movements, short regressions, long regressions. The methodology for this is described in section 3.7.11.2 of the methodology chapter and the findings are reported here in section 4.2.5

d) The patterns formed by fixations on the source texts that suggest the amount and type of reading undertaken. The methodology for this is described in section 3.7.11.3and the findings are reported here in section 4.2.7.

Before reporting the results of each of the sub questions a-d above, section 4.2.1 reviews the broad characteristics of the eye tracking data across the whole task.

### 4.2.1 Fixation data for students’ eye-movements during the test

In this study, students’ reading behaviour during a reading into writing task was measured using eye tracking. Reading behaviour was recorded in terms of participants’ fixations (the number, duration and location of fixations) and the patterns formed by these fixations were analysed to deduce reading behaviour. Participants were asked to complete the task within one hour. The researcher was unsure how much of that hour participants would spend fixating on screen and how many fixations that would generate as to the best knowledge of the researcher as no other research of this nature has been published. In fact, the 30 participants generated a total of 215,052 fixations, about 7000 per participant, although there was a great deal of variation between participants (ranging from less than 4,000 to over 10,000). These fixations added up to over 13 hours of data (about 27 minutes per participant), although once again there was extensive variation (ranging from 11 mins to 40 mins) between participants.

When analysing the data, the researcher considered the number of fixations (both total and averages) per participant or group of participants, the average fixation duration per participant or group of participants, as well as the total duration of all the fixations when added together. Therefore, the data can be summarised as the number of fixations, the average duration per fixation and the total duration of fixations. Fixations are measured in milliseconds (ms) because fixations generally last only fractions of a second, however the total duration time quickly becomes difficult to interpret when reported in millions of milliseconds. Therefore, where appropriate, total duration is reported in minutes and seconds to aid comprehensibility.

*Table 16 Summary of fixation data*

| | Mean | SD | Min | Max | Range |
|---|---|---|---|---|---|
| **Number of fixations per participant** | 7168 | 1839 | 3,903 | 10,363 | 6,460 |
| **Total Fixation duration** (minutes: seconds) **per participant** | 27:28 | 07:47 | 11:02 | 40:18 | 29:16 |
| **Fixation duration** (milliseconds) **per fixation** | 230 | 219 | 60 | 6,267 | 6,207 |

As can be seen in Table 16 there is a wide range of number of fixations and total duration per participant. A summary of the totals for each participant can be found in Appendix 10. There is also a wide range in how long individual fixations lasted with the shortest fixation lasting 60 milliseconds and the longest fixation lasting 6,267 milliseconds (over six seconds) however, 90 per cent of fixations lasted less than 400 milliseconds.

Figure 20 illustrates the distribution of fixations by duration. The fixations were categorised per 50 milliseconds according to their duration. By plotting the number of fixations in each duration category, Figure 20 provides a visual representation of the distribution of fixation duration.

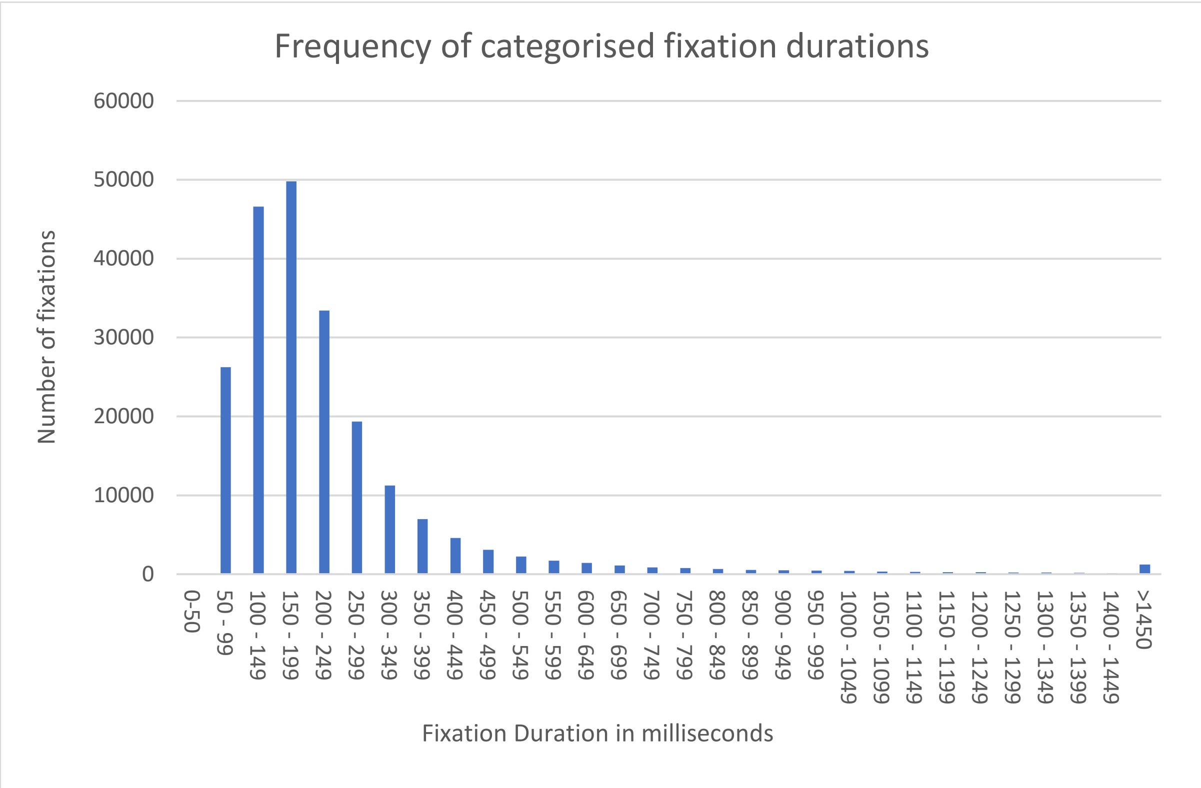



*Figure 20 Frequency of fixations grouped by fixation duration*

This distribution of fixation durations and the statistics reported in Table 16 are similar to those reported in the literature for careful reading at undergraduate level (Rayner et al., 2012) with the average fixation figure reported by Rayner as 218 milliseconds and a range of 66 – 416 milliseconds (Rayner et al, 2012:93). However, there were a number of long fixations which exceeded 1000 milliseconds. These long fixations account for 0.7 per cent of the total number of fixations and 3.8 per cent of the total duration. At more than double the maximum average duration, this raises the question of whether any of the fixation data should be excluded as outliers. However, in this case the researcher suggests that the small number of fixations with very long duration values, in excess of 1000 milliseconds, should not be eliminated from the data analysis for two main reasons.

Firstly, these long fixations were not evenly distributed across the broad areas of interest described in the next section. Instead 94 per cent of fixations over 1000 milliseconds in duration (1354 of 1440 fixations, and 1,781,820 of 1,894,700 milliseconds of total duration) occurred when participants were looking at their own work. The primary goal of this research is to investigate the reading of source texts for reading into writing tasks, rather than the reading of the writers' own emerging work. For this reason, fixations on this area of interest will not be subject to in depth analysis and therefore are not in danger of unduly influencing statistical analyses.

Secondly, very long fixations on the participants' written output (Own Work), whilst participants think about what to write next or reflecting on their writing so far, are perhaps not unexpected when writers reread what they have written so far. Indeed, Caporossi, Alamargot and Chesnet (2004) described long fixations occurring at various stages of the writing process whilst their participants were engaged in rereading their own work during writing, although, Caporossi et al. (2004) do not quantify how long the fixations lasted. During the retrospective think aloud, participants sometimes commented when seeing their gaze fixed for more than a second in one place. One participant commented 'that was the spelling', meaning that he was gazing at a word, thinking about whether it was spelt correctly (P42, line 69 of transcription). This suggests that periods of reflection, an inevitable part of the writing process, are likely to generate long fixations and therefore it seems logical to leave the data set intact.

Much of the literature on reading fixations reports the average saccade length of the eye tracking data. But, much of this literature is based on participants reading single sentences or small sections of text, rather than multiple pages on text, as in this research. However, when participants read across multiple pages of text, it is not possible to calculate an accurate average saccade length for all fixations. For short saccades between words on the same line, saccade distance can be accurately calculated. However, when saccades jump multiple lines of text, or jump from one area of the screen to another it is not logical or accurate to convert the distance travelled between fixations to character spaces. Therefore, any average which included significant numbers of long saccades would be rendered inaccurate and thus potentially misleading. For this reason, saccade lengths are only reported and compared to the literature when this can be done meaningfully.

The large differences between participants in terms of number of fixations and total durations reported in Table 16 are perhaps to be expected as there are several factors which can interact to influence fixations patterns. As long ago as 1979 Rothkopf and Billington concluded that differences in individual eye-movement patterns were marked and it is widely accepted that reading skill influences patterns (Rayner et al., 2012) with more able readers making shorter fixations and longer saccades. It has also been demonstrated that readers vary their rate of reading and the information they pay attention to according to their purpose for reading (Pichert and Anderson, 1977, Anderson, Pichert and Larry, 1983). However, in a large-scale study Fisher (1983) concluded that individual differences remain consistent across tasks, for example when changing from careful reading to search reading.

Given the factors discussed above, large variations between participants in the number and duration of fixations are inevitable. However, it is important to explain the implications of this individual variation this study. Firstly, this study is interested in establishing if there are metrics or patterns of behaviour that apply broadly across the group, irrespective of individual differences. Secondly, this study is interested in comparing groups of participants (for example first year and third year / postgraduate participants). By converting the group data to percentages and comparing percentage differences between groups, individual differences are ameliorated. In conclusion, although there were large differences between participants in terms of both the number of fixations generated and total duration, this is to be expected and should not obscure any findings.

To summarise the broad characteristics of the eye tracking data, on average participants spent 27 minutes and 28 seconds fixating on the screen. The average fixation duration of 230 milliseconds was in line with averages reported in the literature (Rayner et al., 2012:96 report mean of 231ms over a range of reading materials). Although there were some very long fixations these were not excluded as outliers because the overwhelming majority of fixations over one second in length occurred whilst participants were looking at their own written work. The average length of saccades cannot be accurately report calculated based on the full data set. This is because some long saccades cross lines of text diagonally (making it impossible to accurately calculate the distance in characters) whilst other long saccades cross from one area of the screen to another (making it illogical to calculate their distance in characters).

Having reviewed the number of fixations generated and their duration information, the researcher will now describe the data relating to how the participants divided their attention between the different AOIs.

### 4.2.2 The distribution of fixations across different AOIs

As discussed in 3.7.5, the screen was divided into six broad areas of interest (AOI). The AOIs are: the area where participants typed their answer (1:Own Work), the arrows allowing participants to click to move to a different page (2:Move Page), the instructions reminding participants how to move page (3:Screen Instructions), the task instructions (4:Task Instructions), the written content from the source texts (5:Written Sources), the graphic content from the source texts (6:Diagrams) and finally, any area of the screen that does not fall into one of the categories listed above (7:Outside any AOI).

*Table 17 Summary of fixation data across broad AOIs*

| | Number of fixations | | | | Duration of fixations | | | |
|---|---|---|---|---|---|---|---|---|
| AOIs | No. of fixations | Mean / participant | SD / participant | % of all fixations | Total fixation duration (hh:mm:ss) | Mean dur / fixation (ms) | SD of duration / fixation (ms) | % of total duration |
| 1) Own Work | 101,081 | 3369 | 1070 | 47% | 07:41:39 | 274 | 285 | 56% |
| 2) Move Page | 3,366 | 112 | 66 | 2% | 00:16:31 | 294 | 269 | 2% |
| 3)Screen Instructions | 545 | 18 | 23 | 0% | 00:01:37 | 177 | 87 | 0% |
| 4) Task instructions | 12,355 | 412 | 197 | 6% | 00:40:22 | 196 | 116 | 5% |
| 5) Written text | 79,383 | 2646 | 872 | 37% | 04:03:34 | 184 | 100 | 30% |
| 6) Diagrams | 7,199 | 240 | 155 | 3% | 00:25:19 | 211 | 150 | 3% |
| 7) Outside any AOIs | 11,123 | 371 | 190 | 5% | 00:34:50 | 188 | 171 | 4% |
| Total | 215,052 | 7168.4 | | 100% | 13:43:50 | | | 100% |

Table 17 shows the results of the analysis of eye tracking data across the broad areas of interest. The area which received the most attention was ‘own work’ receiving 47 per cent of fixations and 56 per cent of the total duration. The written source text was the second most attended area of interest receiving 37 per cent of fixations. This was followed by Task Instructions (6 per cent), Outside any AOI (5 per cent), Diagrams (3 per cent), Move Page (2 per cent) and lastly Screen Instructions (a fraction of one per cent) of fixations. The researcher will review the findings for each AOI in turn.

#### 4.2.2.1 AOI 1: Own work

Table 17 shows that 47 per cent of fixations and 56 per cent of fixation duration was focused on the participants own work. Therefore, participants devoted almost half of their fixations on screen to composing, editing and re-reading their written work. This is a key finding and suggests that reading the emerging text plays an important role.

The differences between the percentage of fixations and the percentage of fixation duration occur because not all fixations last the same amount of time. Some fixations are very short, and others can be very long. When the long or short fixations are clustered in one AOI this results in differences between the percentage of numbers of fixations and percentage of fixation duration. Table 17 broadly suggests that decreases in average fixation duration result in a drop in the percentage of total duration compared to percentage of fixations. Conversely, increases in mean fixation duration result in an increase in percentage of total duration compared to percentage of fixations. The greatest disparity between the number of fixations and the total duration percentages occur on Own Work (47 per cent/56 per cent) and Written Sources (37 per cent/30 per cent). This variation between the percentages of number of fixations and fixation duration suggest that these two areas of interest received high numbers of fixations with durations at the extremes of the range of fixation durations. To confirm this the researcher investigated the data further in two ways. Firstly, the data for long fixations (over 1000 milliseconds or one Second) which were considered as potential outliers in

4.2.1was analysed and is presented in Table 18 . Table 18 shows that 94 per cent (1354 of 1440) and 94 per cent of the total duration (1,781,820 of 1,894,700 milliseconds) of long fixations were focused on the Own Work area of interest.

*Table 18 Number of fixations with a duration in excess of 1000 milliseconds*

| | Broad Areas of Interest | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Diagram | Screen instructions | Move page | Outside any AOI | Own work | Task Instruction Page 1 | Pages 2-5 | Grand Total |
| Number of fixations over 1000ms | 9 | 0 | 25 | 28 | 1,354 | 10 | 14 | 1,440 |
| Cum dur of fixations over 1000ms | 11,020 | - | 32,440 | 37,620 | 1,781,820 | 12,620 | 19,180 | 1,894,700 |

Therefore, the disparity between the percentage of fixations and percentage of total duration on participants' own work is accounted for by this area of interest receiving the overwhelming majority (94 per cent) of the fixations which exceeded one second in duration. The researcher went on to examine the data to assess what accounted for the difference between percentage of fixations and duration for the written source texts.

#### 4.2.2.2 AOI 2: Move page

Attracting two percent of both the number of fixations and cumulative fixation duration the move page buttons attracted 3,366 fixations and attracted the fifth highest number of fixations out of the six areas of interest. The move page buttons had the highest average duration (294 milliseconds) and the second highest standard deviation (269 milliseconds); the highest standard deviation (285 milliseconds) was on participants' own work. Therefore, the move page buttons attracted a high proportion of long fixations as well as some very short ones. This was a little surprising as one might think that using the mouse to click an onscreen button required only brief fixations. However, Caporossi, Alamargot and Chesnet (2004) analysed eye tracking data from participants hand writing the end of a story (participants had been given the start of the story). Caporossi et al. describe long fixations whilst participants moved the pen to the point in the text where their eyes were fixating. This might help explain the long fixations on the move page buttons as participants guided the mouse cursor towards the appropriate button to click and thus change the page. On other occasions the mouse cursor may have been in close proximity to the button, thus requiring much briefer fixations.

#### 4.2.2.3 AOI 3: Screen instructions

The least fixated area of interest was the screen instructions. The screen instructions had both the lowest mean duration (177 milliseconds) and the narrowest range of durations (indicated by a standard deviation of 87 milliseconds) as wells as accounting for a fraction of one percent (0.25 per cent) of the number of fixations and total duration. This suggests that the participants paid little attention to the screen instructions only glancing at them very briefly. The researcher suggests this is because the screen instructions reinforced information that participants were given during the demonstration prior to starting the test; only serving as a reminder of information that participants were already familiar with.

#### 4.2.2.4 AOI 4: Task instructions

After the participants' own work and the written source texts, the task instructions received the next highest number of fixations (12,355). The average fixation duration (196 milliseconds) is slightly higher than that of the written sources (184 milliseconds) and the standard deviation is also slightly higher (116 milliseconds) than that the written sources (100 milliseconds). This is an important finding as it suggests that the task instructions were read slightly more slowly and carefully than the written sources. However, it is also notable that per word, the task instructions were much less well attended to than the written source texts.

#### 4.2.2.5 AOI 5: Written source texts

The fixations on the written source texts were more homogenous than the fixations on the other AOIs such as AOI 1: own work. This was demonstrated by the low standard deviation (100 milliseconds) of duration per fixation in relation to the mean (184 milliseconds). This mean duration is low by comparison to the mean fixation durations reported in the literature for reading as 230 milliseconds (Rayner et al., 2012). To get a more thorough insight into the factors that resulted in a low mean the data was investigated further.

The written source texts received a greater percentage of fixations (37 per cent) compared to the percentage of total fixation duration (30 per cent). This suggests that the written source texts received high numbers of shorter than average fixations. To confirm whether this was the case the distribution of fixation durations was examined. This involved grouping fixations according to their duration. So, all the fixations with a duration of less than 70 milliseconds were put into the first group, then fixations with a duration of 71-90 milliseconds were put into the next group and so on, up to 1000 milliseconds. The small number of fixations over 1000 milliseconds are not plotted on the graph because extending the horizontal axis to accommodate these long fixations (some exceeding 6,000 milliseconds / six seconds) detracts from the readability of the graph. Also, these long fixations have been investigated in Table 18 and very few were on the written source texts. Figure 21 shows the resulting distribution of fixations categorised by fixation duration for each broad area of interest.

Figure 21 enables the researcher to visualise the distribution of fixations across the range of durations. The numbers on the x axis represent the midpoint of the 20 milliseconds groupings, so the first point plotted at 60 milliseconds represents the number of fixations in the group 50-70 milliseconds, the data plotted at 80 milliseconds represents the number of fixations in the 90-110 milliseconds group and so on.

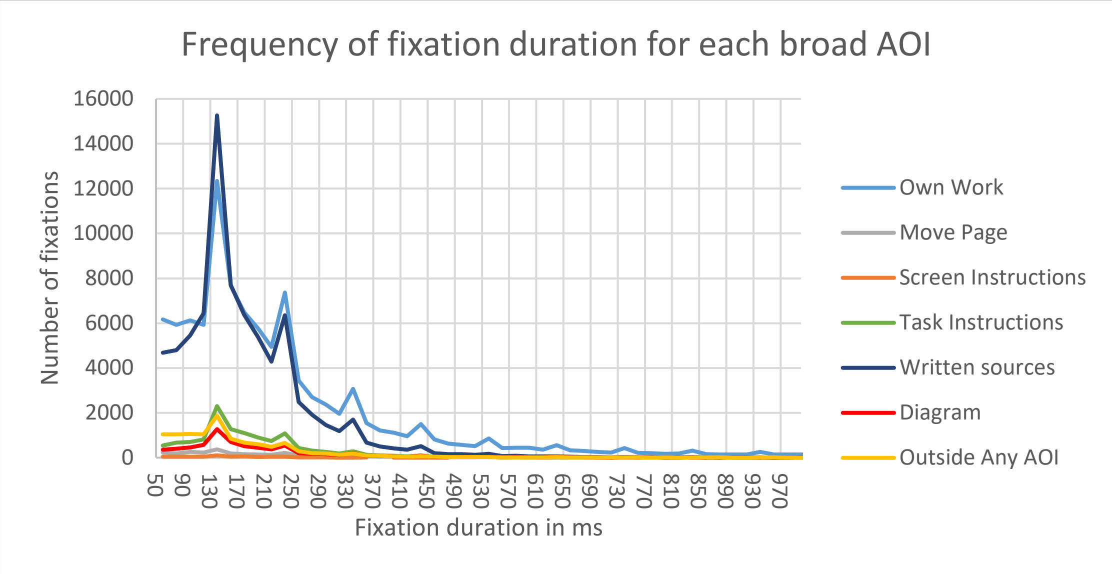



*Figure 21 Distribution of fixation durations for each broad AOI*

Figure 21 serves to illustrate the differences in the distribution of fixation duration between the participants own work and the written sources. The written source texts attracted the highest number of short fixations whilst the participants' own work attracted a higher number of fixations over 200 milliseconds. The high number of short fixations on the written sources combined with low numbers (14) of fixations with durations in excess of 1000 milliseconds account for the variance between the percentage of fixations (37 per cent) and the percentage of cumulative duration (30 per cent) on the written sources. The graph also illustrates the relatively small but consistently higher numbers of longer fixations on participants own work. In fact, 39 per cent (39,786 of 101,081) of fixations on participants own work were longer than the mean duration of 230 milliseconds compared to 24 per cent (19,024 of 79,383) of fixations on the written source texts. However, what is also noticeable is that the numbers of fixations over the mean duration decline at a similar rate for both areas. Therefore, the differences between these two groups can be accounted for by three main characteristics; the over whelming majority (94 per cent) of very long durations were on participants' own work, in addition participants' own work attracted a consistently higher number of fixations with an above average duration whilst the written sources received a greater number of short fixations (around 130 milliseconds duration).

The written sources contained 599 words and the task instructions page contained 209 words, therefore the source texts contained roughly three times as many words as the task instruction. We might therefore, expect to see the written source text receiving about three times as many fixations as the task instructions. However, the written sources received 79,383 fixations compared to 12,355 on the task instructions, almost six and a half times as many fixations. This suggests that the participants fixated more than twice as many times on each word of the written sources than on each word of the task instructions.

In conclusion the fixations on the written source texts were well attended to (accounting for approximately 30 per cent of all fixations) and were more homogenous than fixations on the other AOIs. Surprisingly, fixations on the written source texts had a lower mean fixation duration (184 millisecond) than might be expected when compared to the means reported in the literature for careful reading (230ms).

#### 4.2.2.6 AOI 6: Diagrams

The diagrams received the fifth largest number of fixations (7,199) and accounted for three per cent of the number of fixations and cumulative fixation duration. The mean fixation duration (211 milliseconds) was higher than for the written sources and the task instructions. The standard deviation (150 milliseconds) also exceeded that of the written sources and the task instructions. This suggests that whilst the participants only looked at the diagrams infrequently, or for short periods of time but when they did look at them, they looked at them quite carefully, fixating for slightly longer than when reading textual information.

#### 4.2.2.7 AOI 7: Outside any area of interest

Fixations outside any area of interest had the fourth highest number of fixations (11,123) and a low average fixation duration of 188 milliseconds and a relatively narrow range of fixation durations (standard deviation of 171 milliseconds); accounting for four per cent of the total duration of participants. When the researcher conducted a pilot for this study, participants were questioned about periods when their eyes 'roamed' across the page, before settling on particular areas. Participants reported 'looking' for different parts of the page. Although the researcher notes that rates of fixations outside any area of interest occurred consistently throughout the hour-long sessions and do not lessen as the task progresses. Therefore, the research suggests that participants eye-movements may be roaming across the page whilst they are thinking about what to do next, only settling on an area once the decision has been made.

The fact that participants spent 5 per cent of fixations outside any area of interest only serves to underline that 'looking' is not 'reading'. In other words, whilst participants gazed outside any area of interest, they were in fact fixating on blank AOIs.

In conclusion, the analysis of the data across the broad areas of interest shows that participants focused for longest (total duration) on their own work (56 per cent of all total duration). Participants' attention on their own work was typified by a majority of short fixations (68 per cent of fixations on own work were less than 250 milliseconds in duration) combined with a small proportion of very long fixations (3.5 per cent of fixations on own work had a duration in excess of 1,000 milliseconds). The next highest fixated area were the written sources. Participants' attention on this area was typified by a high number of short fixations (84 per cent of fixations lasted less than 250 milliseconds) and only a very small number of fixations exceeded 1000 milliseconds (14 fixations out of 79,383). The clustering of the very long fixations on participants own work is an indication that participants read their own emerging work in a different way to the way they read the written sources. The increasing mean fixation duration through the task instructions and diagrams suggest that these areas were read or considered slightly more slowly than the written sources.

This concludes the findings in relation to how participants divided their time between the broad AOIs. The most notable findings were that the participants devoted approximately half of their fixations to their own work and that fixations on their written source texts were more homogenous than fixations on the other AOIs and had a low mean fixation duration compared to reading reported in the literature. The next stage of investigation focused exclusively on the written source texts AOI 5) and explored how attention was divided across the individual sentences of the written source texts.

### 4.2.3 The distribution of fixations across sentences in the written sources.

The previous section presented the findings on the way participants fixated on the six AOIs and commented on how these findings might be interpreted. However, as this research also wishes to examine the ways in which participants read during the reading into writing task, this section pursues a more in-depth analysis of participants' fixations just on the written source texts. To remind the reader, of the 215,052 fixations made on screen, 79,383 were on the written source texts. The total fixation duration on screen was 13 hours, 43 minutes and 50 seconds of this four hours three minutes and 34 seconds were on the written source texts.

The techniques which allow analysis on the written source texts could not be applied to either the participants’ own work, the task instructions or the diagrams because the text these areas did not conform to the same font size and line spacing as the written source texts (for a more in-depth explanation refer to 3.7.11 of the methodology chapter).

As we saw from the previous section, participants fixations on the written sources were typified by a high number of short fixations with only a handful of fixations with a duration of more than one second and a low mean fixation duration compared to the data in the literature. We now consider how those fixations were distributed across the individual sentences of the written sources.

By breaking the fixation data down into sentence level areas of interest it is possible to consider which parts of the text attracted more attention and to consider the relevance of each sentence to the reading into writing task that the participants were completing. Each area of interest was named according to the page number and serial position, therefore the title on page one is labelled P1 Title, the first sentence on page one is labelled P1S1, the second sentence on page one is labelled P1S2 and so on.

*Table 19 Numbers of fixations and cumulative fixation duration shown as percentages of total*

| Areas of interest (sentence level) | No. of fixations | Total duration (hh:mm:ss) | No. of words per AOI | Average fixation per word | Average duration per word (seconds) |
|---|---|---|---|---|---|
| P2 TITLE | 210 | 00:00:36 | 4 | 53 | 9 |
| P2S1 | 1,897 | 00:06:01 | 12 | 158 | 30 |
| P2S2 | 3,515 | 00:11:10 | 20 | 176 | 33 |
| P2S3 | 1,224 | 00:03:56 | 11 | 111 | 21 |
| P2S4 | 4,915 | 00:14:19 | 22 | 223 | 39 |
| P2S5 | 6,253 | 00:18:34 | 39 | 160 | 29 |
| P3 TITLE | 109 | 00:00:17 | 5 | 22 | 3 |
| P3S1 | 2,686 | 00:08:35 | 15 | 179 | 34 |
| P3S2 | 2,214 | 00:07:00 | 12 | 185 | 35 |
| P3S3 | 3,724 | 00:12:10 | 29 | 128 | 25 |
| P3S4 | 1,046 | 00:03:06 | 8 | 131 | 23 |
| P3S5 | 3,730 | 00:11:38 | 20 | 187 | 35 |
| P3S6 | 3,010 | 00:09:17 | 16 | 188 | 35 |
| P3S7 | 2,598 | 00:08:01 | 19 | 137 | 25 |
| P3S8 | 3,329 | 00:09:44 | 31 | 107 | 19 |
| P3S9 | 2,130 | 00:06:13 | 31 | 69 | 12 |
| P4 TITLE | 339 | 00:00:57 | 3 | 113 | 19 |
| P4S1 | 1,627 | 00:05:03 | 14 | 116 | 22 |
| P4S2 | 1,909 | 00:05:46 | 15 | 127 | 23 |
| P4S3 | 2,449 | 00:07:15 | 14 | 175 | 31 |
| P4S4 | 4,813 | 00:14:22 | 34 | 142 | 25 |
| P4S5 | 4,305 | 00:12:43 | 29 | 148 | 26 |
| P4S6 | 3,374 | 00:10:02 | 29 | 116 | 21 |
| P5 TITLE | 89 | 00:00:15 | 4 | 22 | 4 |
| P5S1 | 3,837 | 00:12:30 | 30 | 128 | 25 |
| P5S2 | 4,844 | 00:16:07 | 28 | 173 | 35 |
| P5S3 | 1,398 | 00:04:20 | 14 | 100 | 19 |
| P5S4 | 1,799 | 00:05:29 | 17 | 106 | 19 |
| P5S5 | 637 | 00:01:56 | 5 | 127 | 23 |
| P5S6 | 1,794 | 00:05:31 | 21 | 85 | 16 |
| P5S7 | 1,957 | 00:05:47 | 26 | 75 | 13 |
| P5S8 | 1,623 | 00:04:54 | 22 | 74 | 13 |
| **Total** | **79,384** | **04:03:34** | **599** | **Mean: 133** | **Mean: 24** |

As reported in section 3.7.6 of the methodology chapter, the length of each sentence in words was also measured. The researcher opted to use the number of words per sentence rather than the number of characters per sentence because for university readers the reading process normally begins with decoding at word level; recognising words being the lowest level in Khalifa and Weir's model of reading (2009:43). Only children, beginning to learn to read decode at character level, which Rayner et al. (2012:280) describe as a the 'laborious' process of 'reading words sound by sound'. Therefore, it is reasonable to assume that fixations are more likely to be influenced by word boundaries rather than word length. Table 19 reports the fixation data for each sentence, the number of words in each sentence and the rate of fixations per word.

Figure 22 illustrates the data from Table 19. This allows the reader to more easily compare the number fixations and fixation duration per word. The graph plots the sentences in order of attention per word, with those least attended to on the left through to those most attended to on the right.

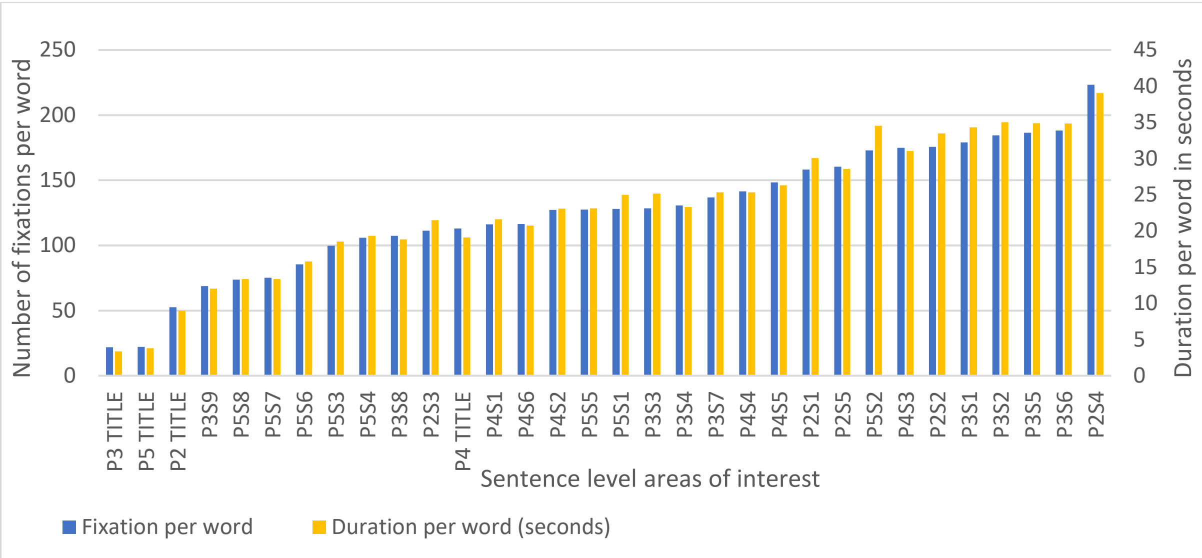



*Figure 22 Bar chart illustrating attention per across all the sentences of the source texts.*

The graph illustrates two things very clearly. Firstly, differences between the number of fixations and the fixation duration per word are relatively small. When converted to percentages, all the sentences have differences of a fraction of one per cent between the number of fixations and duration per word. This suggests that the range of fixation durations was consistent across all sentences, in other words long or short fixations were not focused in particular sentences. This is confirmed by analysis of the mean fixation duration for each sentence (Appendix 11) which range from a mean of 154 milliseconds for the title on page three (P3 Title) through to a mean of 200 milliseconds for the second sentence on page five (P5S2). The standard deviations range from 75 milliseconds to 112 milliseconds. Compared to the range of means and standard deviations seen across the different AOIs discussed in 4.2.2 (means 177 milliseconds – 294 milliseconds, SD 87 milliseconds – 285 milliseconds) fixations on the written source texts are much more homogenous. This is an important finding.

The second thing that the graph illustrates clearly is that there were marked differences between sentences in terms of the amount of attention per word. The range of attention per word varies greatly between those fixated least (P3 Title, P5 Title, P2 Title, P3S9) and those fixated most (P2S4, P3S6, P3S5 and P3S2). To establish whether these differences can be accounted for by the relevance of the content of each sentence, the sentences were rated for their relevance and analysed again. The results of attention according to sentence relevance are reported the in next section.

### 4.2.4 The distribution of fixations across sentences in the written sources according to sentence relevance.

The aim of this analysis is to determine whether the differences in the attention paid to individual sentences (after adjusting for length) were related to sentence relevance (and therefore reflects to some extent the participants' purpose for reading) or whether the differences related to some other factor.

Pichert and Anderson (1977) demonstrated that readers vary their rate of reading and the information they pay attention to according to their purpose for reading (Pichert and

Anderson, 1977, Anderson, Pichert and Larry, 1983). However, as discussed in the literature review, fixation patterns on text can be influenced by a wide range of other factors such as text difficulty (Kliegl, Nuthmann and Engbert, 2006), semantic ambiguity (Frenck-Mestre 2005) and word frequency (Inhoff and Rayner 1986); all of which may account for some of the differences in attention devoted per word of each sentence.

To identify which sentences could expect to receive the most attention, the sentences were rated by three expert raters according to their relevance to the writing task. The levels of inter-rater agreement are reported in section 3.7.7. Analysing the variance between the relevance score and attention allocated per word of each sentence, gives an indication of the extent to which the differences in attention per word can be accounted for by sentence relevance.

The sentences were rated using the following categories: 4-highly relevant (core ideas); 3-relevant (additional details); 2-less relevant for the task; 1-irrelevant. Raters were not asked to rate the page titles because the titles were repeated at the top of the page when the articles continued onto a second page. For example, the title on page two read 'Health and Safety Report' and the title on page three read 'Health and Safety Report continued'. This repetition of information made the titles difficult to rate in comparison to other sentences. Experts did not categorise any sentences as 1- irrelevant. The total relevance score for each sentence are reported in Table 20.

*Table 20 Results of sentence rating exercise*

| Sentence | Relevance score | Relevance category | Fixation per word | Duration per word (seconds) |
|---|---|---|---|---|
| P5S7 | 6 | Less relevant | 75 | 13 |
| P5S6 | 6 | Less relevant | 85 | 16 |
| P5S5 | 6 | Less relevant | 127 | 23 |
| P5S8 | 7 | Less relevant | 74 | 13 |
| P2S3 | 7 | Less relevant | 111 | 21 |
| P4S1 | 7 | Less relevant | 116 | 22 |
| P5S1 | 7 | Less relevant | 128 | 25 |
| P3S3 | 7 | Less relevant | 128 | 25 |
| P3S7 | 7 | Less relevant | 137 | 25 |
| P4S3 | 7 | Less relevant | 175 | 31 |
| P4S6 | 8 | Relevant | 116 | 21 |
| P4S5 | 8 | Relevant | 148 | 26 |
| P2S2 | 8 | Relevant | 176 | 33 |
| P3S1 | 8 | Relevant | 179 | 34 |
| P3S2 | 8 | Relevant | 185 | 35 |
| P3S9 | 9 | Relevant | 69 | 12 |
| P5S3 | 9 | Relevant | 100 | 19 |
| P5S4 | 9 | Relevant | 106 | 19 |
| P3S8 | 9 | Relevant | 107 | 19 |
| P4S2 | 9 | Relevant | 127 | 23 |
| P2S5 | 9 | Relevant | 160 | 29 |
| P3S4 | 11 | Highly relevant | 131 | 23 |
| P2S1 | 11 | Highly relevant | 158 | 30 |
| P4S4 | 12 | Highly relevant | 142 | 25 |
| P5S2 | 12 | Highly relevant | 173 | 35 |
| P3S5 | 12 | Highly relevant | 187 | 35 |
| P3S6 | 12 | Highly relevant | 188 | 35 |
| P2S4 | 12 | Highly relevant | 223 | 39 |

Once the sentences had been classified according to their relevance, they were grouped and the fixation per word and duration per word calculated. The results of these calculations are illustrated in Figure 23.

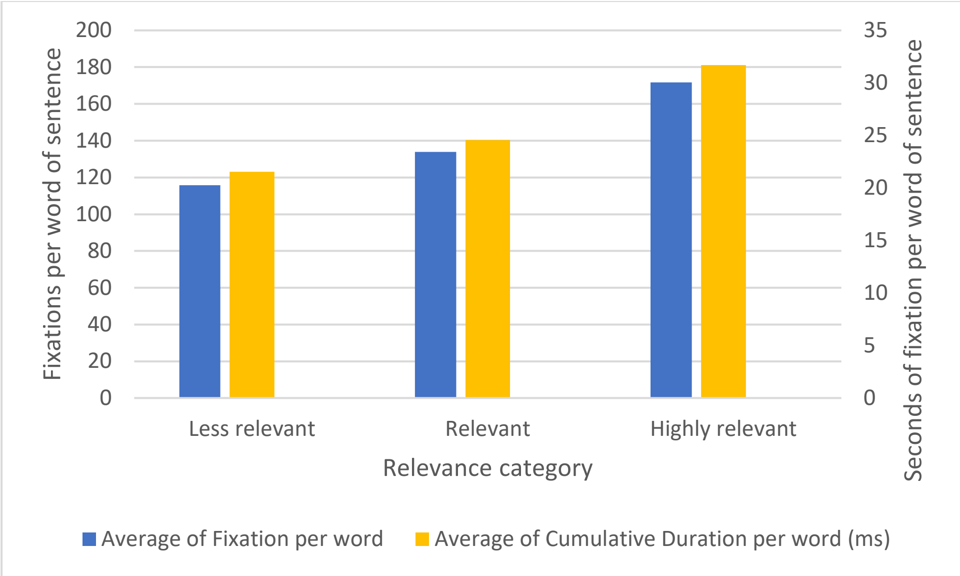



*Figure 23 Average number of fixations and fixation duration per word for sentences when grouped by relevance*

From Figure 23 we can see that, after adjusting for sentence length, participants did fixate more on the sentences judged by the raters to be more relevant. In order to assess the significance of the correlation between the total relevance score awarded by the raters and both the number of fixations and the fixation duration Pearson's correlation was run.

There was a moderate correlation between relevance score and number of fixations per word (r=.531) and there was also a moderate correlation between relevance score and fixation duration per word (r=.510). Therefore, coefficient of determination (calculated as the square of the correlation coefficient) results in sentence relevance statistically explaining 28 per cent of the variability in number of fixations and 26 per cent of the variability in fixation duration. This suggests that participants did alter their fixation patterns in relation to their purpose for reading however this does not account for the majority of difference between attention on sentences. During the retrospective think aloud, when seeing participants repeatedly reading the same sentence, the researcher asked participants to comment. Four themes emerged as participants accounted for repeated reading of sections.

Obtaining a better understanding of the text was cited as the most common reason for rereading a section (18 times). Comments such as ‘The more I did my reading, the more I could take from the article, the deeper my understanding was.’ (Participant 45 line 69 of transcription) were typical of participants. This suggests that participants were deliberately spending time to gain a full understanding of the most important aspects of the source texts.

The second most common reason cited (17 comments) was needing to reread due to a failure to understand; ‘The information hasn't gone in, so I'm reading it over and over again until it does just sink in.’ (P42 line 75). It is not evident from these comments whether the lack of understanding it due to factors such as sentence complexity or the abstract natures of the ideas contained in some of the sentences.

The third most common reason (12 comments) given for rereading areas of text was that participants were extracting information, usually in the form of note taking or paraphrasing. This is clearly linked to the participants’ purposes for reading.

Lastly the participants also reported having to reread sections of the text because although their eyes had ‘read’ the sentence, they had no recollection of what the sentence said. Nine comments such as ‘I have no clue of what I was actually reading, so I started again to make sure’ (P42 line 33), were made.

In conclusion, participants did focus more of their attention on the more relevant sentences. This is an import finding. That is not to say that sentence relevance was the only factor that influenced attention, the statistical analysis suggested that relevance rating explained 28 per cent of the variability in number of fixations and 26 per cent of the variability in fixation duration. The literature on eye-movements in reading (as discussed above) suggests that the predictability of words, the structural complexity of sentences and word frequency all have a role to play determining the amount of visual attention sentences receive. However, it would be extremely complex to try and unpick the combined effects of all these factors on the individual sentences. Therefore, the researcher notes the moderate correlation and suggests that participants did alter their fixation patterns in relation to their purpose for reading. The evidence from the RTA suggests that participants' most common motives for allocating more attention to the most relevant sentences were to gain a deeper understanding and to overcome a failure to understand.

Having considered the way attention was distributed across the written source text sentences, we turn to considering the characteristics of the fixations.

### 4.2.5 The type of movement between fixations

This section reviews the data for the classification of fixations according to whether they move the reader forward through the text, or whether they regressed back to an earlier point in the text, and how far the fixation was from the last fixation.

Analysis of the characteristics of the fixations on the written sources can offer a broad indication of reading behaviour. To remind the reader, when the reader engages in careful reading, they will typically advance through the text with fixations occurring at a distance of 8-character spaces (Rayner et al. 2012). There will not be a relentless progress forward through the text. The reader will, often unconsciously, back track; fixating on words or sentences already read. These back tracks are called regressions. Regressions are thought to occur due to difficulty parsing the current sentence which result in the reader returning to an earlier part of the current sentence, or due to a more general difficulty understanding the text which will result in a regression to an earlier sentence or section of the text (Rayner 2012). Therefore, this analysis sought to identify which fixations progressed forward through the text in relation to the last fixation and which fixations represented a regression in relation to the previous fixation.

Having distinguished forward moving fixations from regressive fixations, the researcher sought to subdivide both categories again. In the case of forward moving fixations the researcher sought to discriminate between forward moving fixations which only moved a short distance (less than 16 character spaces) forward through the text and were likely to represent careful reading behaviour and forward moving fixations which jumped a longer distance (more than 16 character spaces) through the text and so were more likely to represent other types of selective reading such as skimming, search reading and scanning (for a more detailed discussion of these types of reading see Khalifa and Weir 2009:46). In the case of regressive fixations, this stage sought to distinguish between regressions which remained close to the point of reading (on the same line or the line above) likely to represent an attempt to resolve sentence level difficulties and regressions back to an earlier part of the text, likely to represent a more general difficulty in understanding how the main ideas of the text relate to one another.

The mechanism used to categorise fixations is described in 3.7.8 This categorisation generated four groups: fixations following a jump of 16 character spaces or less forward through the text (short forward s), a fixation following a jump of more than 16 character spaces through the text (long forward s), a fixation which moved backward through the text on the same line or the line above the previous fixation (short regression) and finally, a fixation which moved back up two or more lines of text (long regression). As reading skill can influence saccade length, with more skilled readers making longer saccades, the distinction between a short and a long saccade was set at double the average saccade length reported in the literature. This ensured that the even the reading saccades of skilled readers would be recorded as short saccades.

To remind the reader, of the 215,052 fixations made on screen, 79,383 were on the written source texts. The total fixation duration on screen was 13 hours, 43 minutes and 50 seconds of this four hours three minutes and 34 seconds were on the written source texts. Table 21 reports distribution of the fixations across the four categories.

*Table 21 Fixations on source texts according to direction and distance moved between fixations for all participants*

| Fixation type | No. of fixations | Total duration (hh:mm:ss) | Mean duration / fixation | SD of duration / fixation |
|---|---|---|---|---|
| Short forward | 39,061 | 02:04:03 | 190ms | 104ms |
| Short regression | 24,159 | 01:13:22 | 182ms | 100ms |
| Long forward | 12,945 | 00:36:48 | 171ms | 86ms |
| Long regression | 3,219 | 00:09:21 | 174ms | 93ms |
| Total | 79,384 | 04:03:34 | | |

First of all, we will review the numbers of fixations of each type and what insight the retrospective think aloud and questionnaire data can offer.

#### 4.2.5.1 Short forward moving fixations

From Table 21 we can see that almost half (49 per cent) of the fixations on the written sources moved forward a short distance suggesting that the participants spent a considerable amount of the time on texts engaged in careful reading. Indeed, in the questionnaire all 30 participants agreed or strongly agreed that they ‘read the whole of that source text slowly and carefully’ to get a ‘deep understanding of the main ideas in a source text’. This would seem logical and is to be expected.

#### 4.2.5.2 Short regressions

Regressions on the same line or to the line above (short regressions) accounted for 30 per cent of fixations on the written sources. This is surprisingly high. The rates of regression reported in the literature for careful reading suggest that 10-15 per cent rates are common for careful reading (Rayner 2012). However, during the retrospective think aloud, out of 101 comments which referred directly to reading behaviour, 34 referred to regressing during reading. Comments in response to seeing their eye-movements repeatedly covering the same sentence or lines of text included 'My brain can't process it first time, that's why. I was just trying to understand –' (P47, line 24), 'it didn't go in straightaway' (P48, line 19); these comments seem to refer to an inability to understand a specific idea in the text. Other comments included, ' I was like, 'Okay. I will read again'. Just to understand more.' (P26 line 40) and another participant reported reading 'over and over again' until he understands it 'completely' (P52, line 5). These comments seem to suggest that participants kept going back over sections of the text to gain a deeper understanding of the text, which is slightly different to failing to understand. This also suggests that many of these regressions were part of a deliberate strategy rather than a sub-conscious process.

#### 4.2.5.3 Long forward moving fixations

Table 21 reports that fixations which moved more than 16-character spaces forward through the text (long forward) accounted for 16 per cent of the fixations. This type of fixation is most likely to be associated with the expeditious reading outlined by Khalifa and Weir (2009:46) as readers make long jumps through the text, sampling sections in the case of scanning or searching for specific words. In the retrospective think aloud participants made 13 comments after seeing their eye-movements making long, rapid saccades. In nine of these cases participants reported searching for a word or piece of information. However, participants also reported reading for gist 'I'm not reading, like, properly. I'm just looking for [the] main thing, just words which can be used in my text. It was just words, like ideas.' (P26, line 113) and being able to jump directly to a point in the text to locate information 'I find that I quite easily understand the structure of the text, and if I ever have to come back, I know where. Sometimes I even remember, if I'm looking at a book, I remember where on the page. Is it bottom or middle-' (P38, line 140).

#### 4.2.5.4 Long regressions

In terms of regressions which jumped back up two or more lines just four per cent of fixations followed long regressive saccades. There were a couple of comments in the retrospective think aloud alluding to long regressions such as; P34, line 52: ''I was trying put in both pictures [pointing to the two sections of text which contain separate solutions in article 1] in my brain [unintelligible 00:04:39] so I could go back and forth because they're both the same but they're not the same at the same time. It's hard to remember them both at the same time.'. Also, P52 transcription line 19, when the participant explained a return to an earlier paragraph is motivated by a lack of understanding. However, there were also examples of long regressions as part of deliberate strategies such as 'I was going to go through all the solutions separately...I found out where each one stopped, and then I look back to the first solution' (P37, line 100). Also, several of the episodes when participants commented on behaviour in which their eyes made large saccades, both forward and backward through the text, the participants reported being engaged in searching for specific pieces of information. Therefore, the researcher suggests that long regressive saccades could be indicative of attempts to resolve text or paragraph level misunderstandings / confusion or could equally form part of search reading behaviour.

In summary, the most notable finding from this analysis is that regressions were present at a much higher rate than rates reported in previous eye tracking studies of reading. Short regressive fixations seem to result from frequent bouts of rereading sentences or sections of text. Longer regressions may be due to search reading or could relate to attempts to resolve text or paragraph level misunderstandings.

### 4.2.6 Differences in mean fixation durations between types of fixation

Having compared the number of short forwards, short regressions, long forward and long regressions, we consider what the fixation duration information for each type of fixation can add to the picture of participant behaviour.

The first question to consider is whether the fixation durations are distributed evenly across all the four categories of fixation, in other words, did one type of fixation have high number of fixations with long durations whilst another type of fixation had high numbers of fixations with short durations. To answer this question, as with sections 4.2.2 and 4.2.3 the fixations were grouped into 20 milliseconds duration groups according to their duration and plotted on Figure 24 to illustrate the frequency of fixations in each duration group for each fixation type.

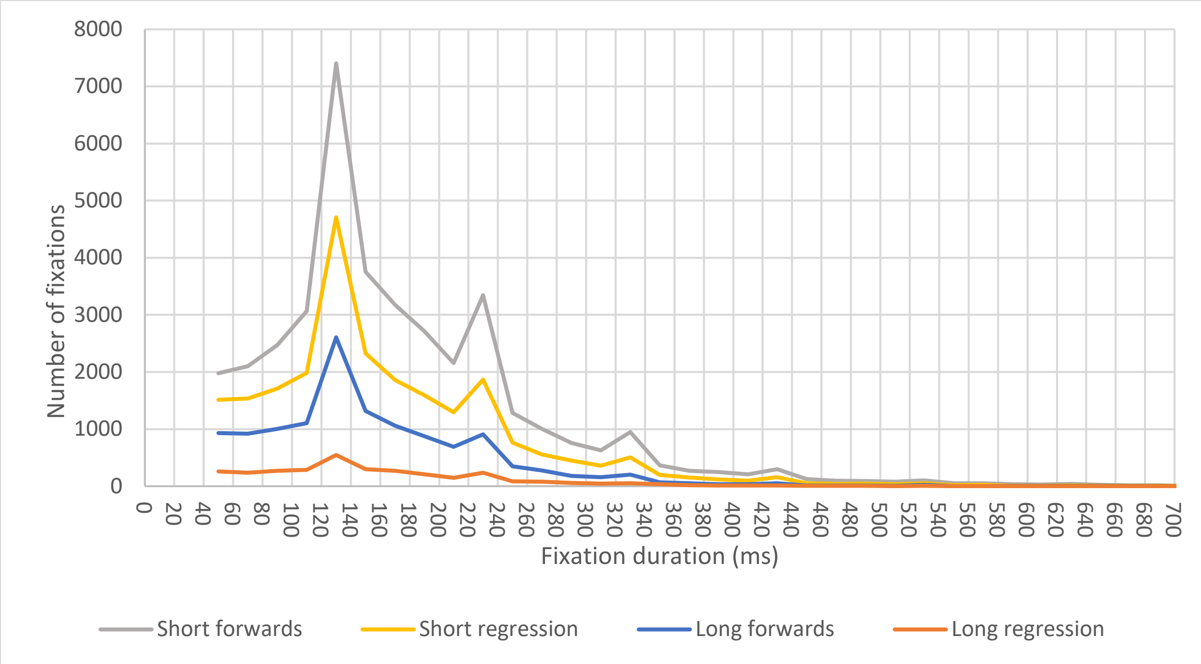



*Figure 24 The frequency of fixations, grouped by direction and distance, when categorised by fixation duration.*

From Figure 24 we can see that fixations following a short jump forward (short forward s) consistently had the highest number of fixations at all durations, fixations after a short jump backward (short regressions) had the second highest number of fixations at all durations and fixations following a long forward movement (long forward s) had the third highest number. For these three types of fixation we can see a very similar distribution of fixation durations with the number of fixations peaking at 130 milliseconds, followed by a decline at a similar rate until a second peak at around 230 milliseconds and so on. The pattern of distribution is much ‘flatter’ for fixations following a long jump back through the text (long regression) with much less variety in the number of fixations at each duration. Whilst not totally consistent, the frequency of fixations when categorised by fixation duration does not suggest that any particular type of fixation was disproportionately responsible for influencing fixation durations.

The mean fixation duration times reported in Table 21 reflect the differences in the distribution of fixation durations. As expected the fixations following a short jump forward have the longest mean fixation duration (190 milliseconds). This type of fixation is associated with careful reading and therefore we would expect to see slightly longer average duration times. Fixations after a short regression (short regression) had a slightly shorter average fixation duration of (182 milliseconds). The research was unable to find any reference in the literature to average fixation durations for fixations after a short regression making it difficult to interpret this data.

Fixations after a long jump forward through the text and fixations after a long regression had average durations of 171 milliseconds and 174 milliseconds respectively. As discussed earlier in this section, the retrospective think aloud data seems to suggest that both long forward s fixation and long regressions may be examples of the expeditious reading such as skimming, search reading and scanning. This would naturally result in shorter fixation durations.

To understand whether the different types of fixation occurred consistently over the course of the task the researcher calculated the frequency of each type of fixation during each minute of the hour-long task. The results are displayed in Figure 25

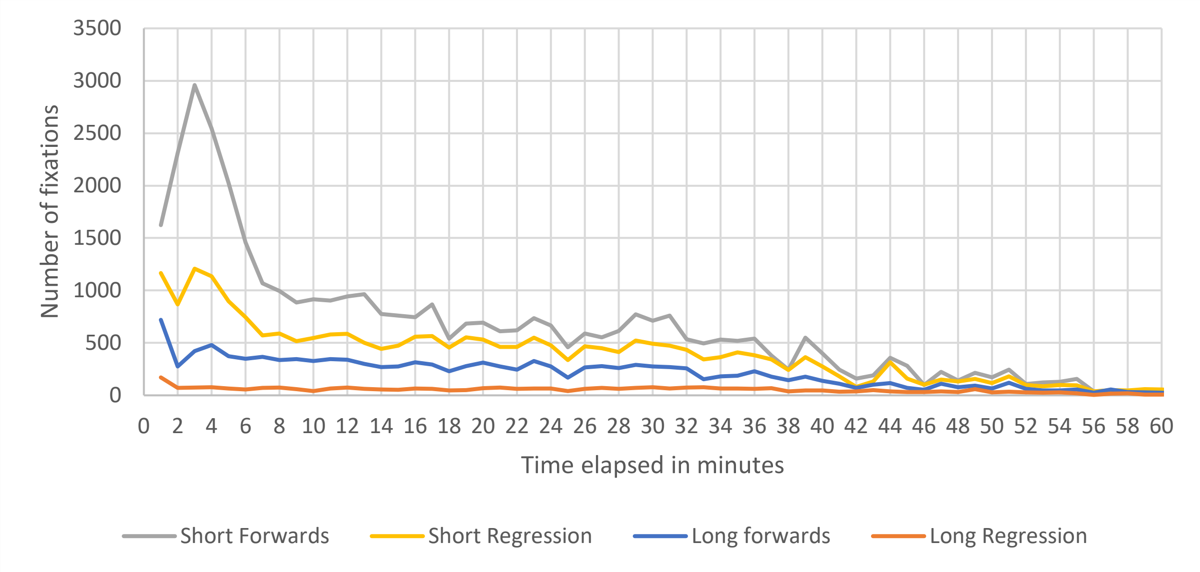



*Figure 25 Distribution of fixations during the hour-long task*

From Figure 25 we can see that during the first seven minutes of the task short forward moving fixations (grey line) dominate, outnumbering short regressions (yellow line) by approximately three to one in minute three. The data for the first ten minutes of the graph is presented as percentages in Table 22 to facilitate easy comparison of rates of regression to rates of forward movements during the first ten minutes.

*Table 22 Percentage of fixations types for the first ten minutes of the task*

| Minutes elapsed | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Short forward | 44% | 66% | 63% | 60% | 60% | 56% | 51% | 50% | 49% | 50% |
| Short regression | 32% | 25% | 26% | 27% | 27% | 29% | 28% | 30% | 29% | 30% |
| Long forward | 20% | 8% | 9% | 11% | 11% | 13% | 18% | 17% | 19% | 18% |
| Long regression | 5% | 2% | 2% | 2% | 2% | 2% | 3% | 4% | 3% | 2% |
| Grand Total | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

From Table 22 we can see that after the first minute of the task the proportion of short forward moving fixations increases rapidly from 44 per cent to 66 per cent. To review whether the mean fixation duration and mean saccade length for short forward moving fixations remained constant throughout the duration of the task, the researcher examined the average fixation duration and saccade length for each minute of the task.

For accuracy, it was necessary to eliminate all fixations in this group that moved from the end of one line to the beginning of the next. This is because although they represent a small jump forward through the text in terms of characters, the distance in pixels between the end of one line and the beginning of the next was disproportionately long.

The results of the mean fixation duration, calculated minute by minute, for short forward moving fixations is illustrated in Figure 27.

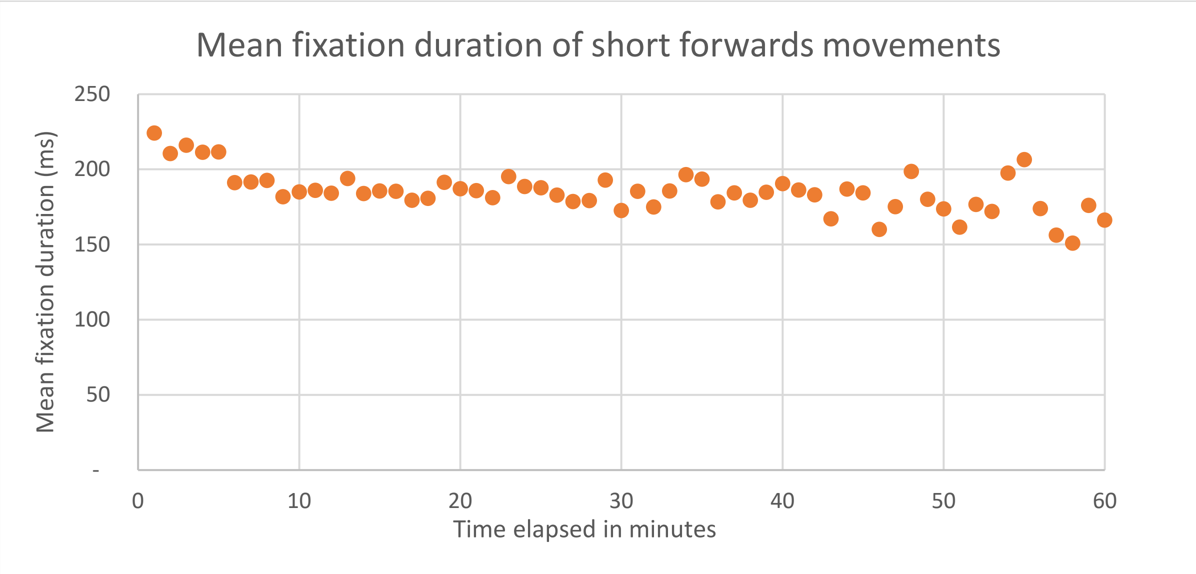



*Figure 26 The average, per minute of the task, fixation duration for short forward moving fixations*

Figure 26 illustrates that fixations after a short jump forward were slightly longer during the first ten minutes of the task. As the task progresses the mean fixation duration for short forward moving fixations fluctuates more. This may be because after 35 minutes some participants finished the task. As more participants finished the task and numbers of participants reduced, individual participant's statistics will have had a greater impact on the mean duration, resulting in greater volatility of the mean.

Figure 27 illustrates the results of the mean saccade length, calculated minute by minute, for short forward moving fixations.

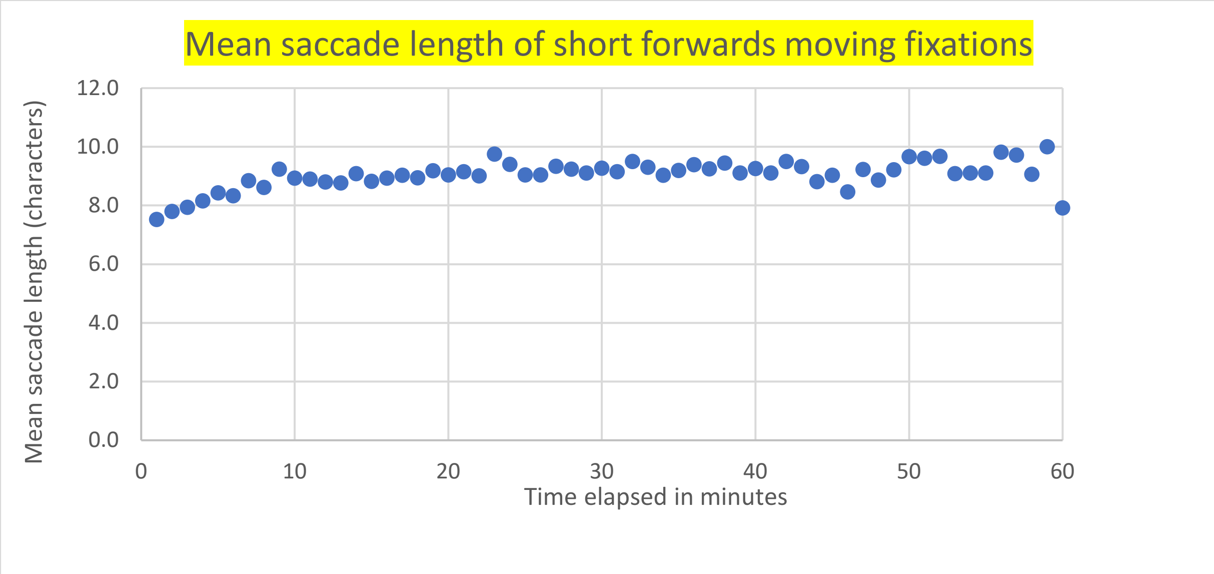



*Figure 27 The average, per minute of the task, that short forward moving fixations moved (in characters)*

From Figure 27 we can see that the average saccade distance for short forward moving fixations increased during the first ten minutes of the task. From an average of seven and a half characters during the first minute the saccade distance average gradually increases to 9 characters by the tenth minute of the task. This suggests that participants read more carefully at first, gradually making longer saccades (jumps) between fixations until minute ten when the saccade length seems to level off. As with the mean fixation duration, the mean saccade length becomes more volatile towards the end of the hour-long task. Perhaps once again due to reducing numbers of participants.

Both of these facts suggest that participants started reading rather slowly and carefully, gradually increasing their reading speed until minute ten when it seems to plateau. This early peak in reading activity is a key finding which is explored further in section 4.2.7.3

Whilst helpful, the analysis of the characteristics of fixations still left many questions unanswered. Therefore, the researcher went on to look for patterns of fixations that might more clearly indicated the participants' reading behaviour as they progressed through the task.

### 4.2.7 Fixation patterns suggesting the amount and type of reading

This section reviews the data for the classification of fixation patterns suggesting the amount and type of reading. Once again, only the fixations on the written source materials are suitable for analysis in this way. Arguably, the data from the last section analysing the characteristics of fixation can be used to suggest the type of reading that is taking place however, conclusions can only be drawn in rather broad terms. Therefore, researcher sought to find a method which allowed a minute by minute classification of the participants' eye-movements. The reader may find it useful to refer to section 2.4.2 of the literature for a review of methods for detecting reading from eye tracking data, and section 3.7.11 of the methodology chapter for a full explanation of the researcher's method of detecting careful reading.

The research sought primarily to distinguish careful reading from other types of selective reading. To remind the reader, when the reader engages in careful reading, they will typically advance through the text with fixations occurring at a distance of 8-character spaces (Rayner et al. 2012). There will not be a relentless progress forward through the text. The reader will, often unconsciously, back track; fixating on words or sentences already read. These back tracks are called regressions. Regressions are thought to occur due to difficulty parsing the current sentence which result in the reader returning to an earlier part of the current sentence, or due to a more general difficulty understanding the text which will result in a regression to an earlier sentence or section of the text (Rayner 2012).

In the literature the rate of regression recorded for careful reading is typically 10-15 per cent. 15 per cent regression rate equates to six short forward moving fixations to every regression. However, as the researcher detailed in the last section, the rates of regression whilst participants completed this task were much higher, with the retrospective think aloud data suggesting that participants regularly made regressions to reread sections or sentences to further their understanding of the text. Therefore, the researcher concluded that to allow the higher rates of regression that typified reading in this study the algorithm would be set to report careful reading when participants made at least three short forward moving fixations to every regressive fixation. Although the regression rate during the first ten minutes of the task was slightly higher than this, the researcher suggests that not all the short regressions will have been part of the careful reading seen on the replays of eye tracking behaviour and described in the retrospective think aloud recordings.

The researcher also decided to distinguish between episodes of reading that remained within a single sentence and episodes that extended beyond a single sentence. Please refer to section 3.7.11 for a complete description of the categorisation of reading type by the algorithm.

The number of fixations classified as each type of reading are reported in Table 23 below.

*Table 23 Fixations categorised by reading type*

| Reading type | No. of fixations | Total duration (hh:mm:ss) | Mean duration /fixation (ms) | SD of Duration/ fixation |
|---|---|---|---|---|
| Careful Local | 7,459 | 00:24:12 | 195 | 102 |
| Careful Global | 15,515 | 00:51:58 | 201 | 105 |
| Selective Local | 28,421 | 01:23:26 | 176 | 98 |
| Selective Global | 27,989 | 01:23:58 | 180 | 96 |
| Total | 79,384 | 04:03:34 | | |

First, the numbers of fixations and total fixation duration for each reading category will be reviewed. Table 23 shows careful global fixations accounted for 20 per cent of number and 21 per cent of total duration with careful local fixations accounting for approximately 9 per cent of number and 10 per cent of total duration. Therefore, careful reading accounted for less than 30 per cent of the reading activity.

Selective local fixations accounted for over a third (36 per cent of number and 34 per cent of total duration) of the fixations, selective global fixations were very similar, accounting for 35 per cent of the number of fixations and 34 per cent of total fixation duration. Together, therefore, selective types of reading accounted for over 70 per cent of the reading activity.

The low rates of careful reading are perhaps surprising, when we reflect on the previous section, which showed that almost half of all fixations were short forward moving fixations. However, as careful reading limits the number of regressive fixations that can occur to a ratio of three short forward moving fixations to every regressive fixation, high numbers of regressions are likely to limit when careful reading is deemed to have occurred. To clarify the proportion of short forward, long forward moving fixations and long and short regressive fixations which the algorithm had coded as belonging to each reading type the researcher conducted further analysis. The results are reported in Table 24.

*Table 24 The direction and distance properties of fixations in each reading category reported as a percentage*

| | Short forward moving fixations | Short Regressive fixations | Long forward moving fixations | Long Regressive fixations | Total |
|---|---|---|---|---|---|
| Careful Local | 81.69% | 18.30% | 0.00% | 0.01% | 100% |
| Careful global | 80.72% | 18.91% | 0.00% | 0.37% | 100% |
| Selective Local | 36.14% | 35.61% | 20.83% | 7.42% | 100% |
| Selective Global | 36.35% | 34.80% | 25.10% | 3.76% | 100% |

#### 4.2.7.1 Careful reading

Table 24 seems to confirm that algorithm works well to classify careful global reading behaviour because fixations identified as belonging to careful global reading consisted of an overwhelming majority of short forward moving fixations (81 per cent) and a small percentage of short regressions (19 per cent). Interestingly, this level of regressions accords with the highest rates of regression reported by Rayner (2012:96) for college age readers reading different types of texts. From Table 24 we can see that careful global reading has the longest average duration per fixation (201 milliseconds) although it should be noted that when the readers in Rayner's study, had a regression rate of 18 per cent they had an average fixation duration of 264 milliseconds. careful global reading reports 0.37 per cent of long regressions, showing that readers did occasionally regress back to an earlier point in the text to resolve confusion. This is perhaps an underestimate of the times that this may have happened, as the algorithm was unable to report regressions back to previous pages due to technical limitations.

Careful local reading reported a very similar combination of fixation types with 82 per cent of short forward moving fixations compared to 18 per cent regressions. This appears to confirm that the algorithm is consistently identifying patterns likely to represent careful reading. The average fixation duration is slightly shorter for careful local reading (195 milliseconds) than for careful global reading (201 milliseconds). This suggests that as participants progressed beyond a single sentence their rate of reading reduced slightly, perhaps because they were forced to consider additional propositions, introduced by subsequent sentences.

#### 4.2.7.2 Selective reading

'Selective' reading consisted of a more diverse range of fixations types. Selective local consisted of 36 per cent short forward moving fixations, 36 per cent of short regressive fixations, 21 per cent of long forward and 7 per cent of long regressions. To remind the reader, long forward represented fixations after a saccade which moved more than 50-character spaces forward through the text. Long regressions moved the reader up two or more lines. The researcher examined the data to understand how long forward or long regressions formed part of local reading. The long forward and long regressions were always the first fixation in an episode of reading. The reader made a long jump forward or a long regression to a sentence and then proceeded to move back and forth within the sentence. The high rate of regressions suggests that the reader was either having difficulty parsing the sentence or searching for particular words or information within the sentence. The lower average fixation duration (98 milliseconds) compared to careful reading (local: 103 milliseconds, global: 107 milliseconds) perhaps suggests that participants were more likely to be engaged in search reading because difficultly parsing a sentence might be expected to produce longer fixation durations as the reader slowed down whilst they attempted to resolve.

Selective global reading consisted of very similar levels of short forward and short regressions but differed in the proportions of long regressions. long forward accounted for 25 per cent and long regressions accounted for four per cent of fixations classified as selective global reading. This suggests that during periods identified as selective global reading participants were predominantly moving forward through the text, signified by the low percentage of long regressions compared to long forward. The levels of short forward and short regressions suggest that there were periods of intensive backward and forward saccades, perhaps before jumping on to another sentence. When considering all these factors together, it seems reasonable to suggest that during periods when participants' fixations are classified as selective global they may have been engaged in reading behaviour that resembled gist or search reading behaviour.

The most important finding to emerge from the analysis of the types of reading is that selective reading accounted for 70 per cent of the reading of the written source texts. Only 30 per cent of reading appeared to reflect the type of careful reading described in the literature. To help with interpretation of this finding the researcher went on to investigate how the different types of reading were deployed over the course of the task.

#### 4.2.7.3 When the different types of reading occurred

The researcher then examined when the behaviour classified as each type of reading occurred. Therefore, number of fixations assigned to each type of reading were plotted according to the minute of the task when they occurred. The resulting graph is Figure 28.

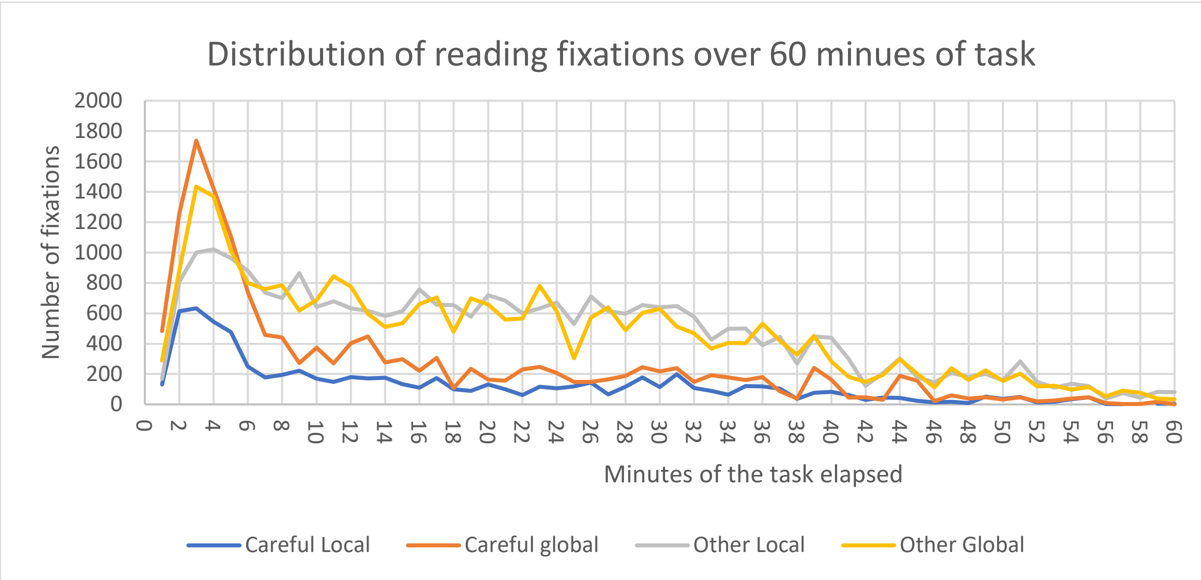



*Figure 28 Graph illustrating changes in the numbers of fixations in each reading category over the 60 minutes of the task*

Figure 28 illustrates the changing numbers of fixations for each reading type over the course of the hour-long task. The implications of these findings will now be considered for each type of reading.

Fixations classified as careful global reading dominated for the first six minutes, peaking at 1700 fixations in minute three before declining to about half that rate by minute six. This suggests that participants engaged in careful global reading of the source texts in the opening six minutes. After six minutes the amount fixations classified as careful global reading drops away considerably to around 400 fixations per minute at eight minutes and continues to decline steadily to virtually nothing by the end of the task. This accords closely with the researcher's observations during the data collection when participants seemed to devote much of their attention to a systematic reading of the texts in the opening minutes of the task. As they began to make notes and to write, their reading became much more spasmodic, with shorts bursts of reading interspersed with writing and rereading their own work.

Selective global reading also generated high numbers of fixations during the first six minutes of the task, peaking at around 1400 fixation per minute in minute three. Selective global reading levels also drop after six minutes, but not as far as careful global reading. From minute six onwards, selective global and selective local reading follow a very similar path, with small increases and decreases but overall maintaining a steady decrease from 800 fixations per minute to around 100 fixations per minute by the end of the task. The early peak in selective global reading could be seen as confirmation that fixations classified as selective global represented gist reading in the early stages of the task. After six minutes the fixations classified as selective global are more likely to represent search reading as participants returned to the texts during note taking and writing, looking for relevant content.

Selective local reading reached a peak of 1000 fixations per minute in minute four and then made a steady decline, fluctuating within a 200 fixation per minute range as the task progressed, petering out at 50 fixations per minute at the end of the test. Levels of selective local and selective global reading ran in parallel from minute six onwards, dominating both types of careful reading.

In summary, the first 6-7 minutes of the task saw a large peak in global reading activity. Careful global reading was slightly more in evidence than selective global reading during this time. The researcher used RTA and questionnaires to triangulate the eye tracking data and assist with interpretation of the data. The next section moves on to consider their findings.

### 4.2.8 Students' processes of completing the reading into writing task

This study used two other research methods to assist with the interpretation and triangulation of the eye tracking data. These were the retrospective think-aloud (RTA) described in section (3.6.6 and 3.7.12) recordings made immediately after the participants finished the task and a short cognitive processing questionnaire delivered after the RTA session (3.6.7 and 3.7.13). The cognitive processing questionnaire was exploratory in nature and therefore the results are reported using descriptive statistics. Drawing upon the RTA, the questionnaire and the eye tracking data, the students' processes on the task will now be reported one cognitive process at a time.

#### 4.2.8.1 Task Representation

Items in the questionnaire were rated on a Likert scale with 'strongly agree' rated as five, through to 'strongly disagree' rated as one. Therefore, for the means reported in Table 25, the higher the number, the greater the level of agreement.

In relation to task representation (Table 25), all participants agreed or strongly agreed that they needed to include ideas from the source texts but were more ambivalent about the need to include ideas from their memory their own opinions.

*Table 25 Questionnaire Results: Task Representation*

| | | All participants | | | | | |
|---|---|---|---|---|---|---|---|
| **Question** | **Task Representation** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** |
| **1** | I thought I needed to include a lot of my own opinions | 1 | 4 | 9 | 14 | 2 | **2.60** |
| **2** | I thought I needed to include a lot of ideas from the source texts | 13 | 17 | 0 | 0 | 0 | **4.43** |
| **3** | I thought I needed to include a lot of ideas / information from my memory | 1 | 3 | 5 | 13 | 8 | **2.20** |
| **4** | I used a strategy / strategies to help me decide what to include | 12 | 16 | 0 | 1 | 1 | **4.23** |
| **5** | I returned to re-read the assignment question several times. | 20 | 8 | 0 | 1 | 1 | **4.50** |
| **6** | My understanding of the assignment question changed whilst I was completing the assignment. | 5 | 7 | 4 | 12 | 2 | **3.03** |
| **7** | I set myself additional goals, beyond what was required for the assignment | 2 | 6 | 13 | 8 | 1 | **3.00** |

The use of strategies and returning to the task instructions (assignment question) were also almost universally adopted. Fewer participants reported updating their understanding of the task as they worked or setting themselves additional goals.

During the RTA the following themes emerged.

*Table 26 RTA results: Task representation*

| Total number of comments for **Task Representation** | | **61** |
|---|---|---|
| Sub themes | Genre | 12 |
| | Goal setting | 8 |
| | Recall of task | 14 |
| | Rhetorical function | 26 |
| | Time constraints | 1 |

Despite participants reporting ambivalence about including ideas from their memory, the RTA suggested that during the task, reading the source material did cause participants to make connections with their own experiences and memory for example P12 line 34 ‘I wanted to make sure the goal was truthful based on my own previous experience not just the text I must admit.’ Eight participants made reference to their own experiences. Fewer participants (5) made reference to forming or including their own opinions which perhaps suggests that participants were less confident that their own opinion was required.

The eye tracking data revealed that the task instructions were quite poorly attended to compared to the source texts, however, all the participants read the task instructions several times. During the RTA there were 14 comments related to returning to the task in order for participants to check or remind themselves what the task required. In answer to question five of the questionnaire participants all agreed or strongly agreed that they returned to reread the task instructions several times.

The questionnaire suggested that participants had mixed feelings about whether their understanding of the task changed as they worked (13 recorded a neither agree nor disagree). During the RTA five participants made comments that suggested that their understanding of the task changed as they worked on the task. Therefore, for some participants at least, task representation took a while to develop.

### 4.2.8.2 Macro planning

The questionnaire results for macro planning phase of the task are reported below.

*Table 27 Questionnaire results: Macro planning*

| | | All participants | | | | | |
|---|---|---|---|---|---|---|---|
| **Question** | **Macro-planning** | Strongly agree | Agree | Ambivale | Disagree | Strongly | **Mean** |
| **8** | I formed a plan before starting to read or write | 5 | 13 | 4 | 7 | 1 | **3.47** |
| **9** | My plan was formed / changed as I was reading the source texts | 6 | 12 | 4 | 8 | 0 | **3.53** |
| **10** | My plan was formed / changed as I was writing my answer | 4 | 12 | 5 | 7 | 2 | **3.30** |
| **11** | I thought about how to adapt my writing to suit a particular style (business / academic / journalistic etc.) | 6 | 15 | 5 | 4 | 0 | **3.77** |
| **12** | I thought about how formal / informal to make my writing | 9 | 18 | 2 | 1 | 0 | **4.17** |
| **13** | I had a clear idea of who I was writing to. | 6 | 11 | 8 | 4 | 1 | **3.57** |
| **14** | I adapted my writing to suit the reader. | 2 | 12 | 8 | 6 | 2 | **3.20** |
| **15** | I asked myself whether the information was relevant to the assignment question. | 15 | 14 | 0 | 1 | 0 | **4.43** |
| **16** | I asked myself whether there was enough information to answer all parts of the assignment question. | 10 | 10 | 4 | 6 | 0 | **3.80** |
| **17** | I thought I would follow the same organisation as one of the source texts | 2 | 4 | 4 | 17 | 3 | **2.50** |
| **18** | I thought I would use the assignment question to help me organise my answer | 17 | 9 | 1 | 2 | 1 | **4.30** |

In terms of macro-planning only one participant reported not considering relevance of the content and only three reported not using the assignment question to help them organise their answer. However, in practice, not all participants did organise their answer in response to the question. During the task the researcher noted that several of the participants reported used the key points of the task instructions as a macro plan, making notes from the task instructions and using these as headings for their answer. However, the researcher noted that several participants reported using other sources for their organisation such as a previous essay (P26, line 18) or the structure of the texts (P35, line 93).

*Table 28 RTA results for Macro planning*

| **Total number of comments for Macro-planning** | | **44** |
|---|---|---|
| Sub themes | Finding content | 8 |
| | Structure | 36 |

However, comments such as ‘‘I've taken the question and translated it on to the page as a guide, so I know what I'm looking for.’(P12, line 12) were more common.

The other notable finding relating to macro-planning related to when planning occurred. The responses to questions eight, nine and ten which asked about when macro-plans were formed and whether they changed during the task. Responses ranged across all three options (Plan emerged before starting to read / Plan formed - changed during reading sources / plan formed - changed during writing). During the RTA the researcher noted that some participants skipped reading the task instructions until after they had read the texts (for example P31, P37, P50).

Many of the participants that decided to use the task as a macro-plan seemed to use their macro-plan to help them guide their reading. For example, P32, line 16 used the task headings to plan ‘what bits I need to go for’. This is discussed further in the next section.

### 4.2.8.3 High level reading processes

For some participants the procedure they used appeared well rehearsed and methodical. Upon reading the task instructions some participants immediately began to make notes. These participants tended to use the task instructions as the basis for their macro-plan, extracting key points from the task to form the skeleton or outline of their answer. This approach seemed to reflect a very systematic approach to the task where participants used the task instructions to prime their attention for the reading to come.

The questionnaire results (Table 29) revealed that most participants felt that they had engaged in a variety of forms of selective reading in order to extract the relevant information from the texts.

*Table 29 Questionnaire results for High level reading and connecting and generating*

| Question | High-level reading including connecting and generating | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | Mean |
|---|---|---|---|---|---|---|---|
| **19** | When I wanted to **completely understand** the meaning of an individual sentence I read it slowly and carefully. | 14 | 16 | 0 | 0 | 0 | **4.47** |
| **20** | When I wanted a **deep understanding of the main ideas** in a source text I read the whole of that source text slowly and carefully. | 9 | 14 | 4 | 3 | 0 | **3.97** |
| **21** | When I wanted **to find a particular word, date or specific detail** I quickly scanned through the text, skipping over some sections of the text until I found the sentence I needed. | 23 | 7 | 0 | 0 | 0 | **4.77** |
| **22** | When I wanted to **get the gist (main idea)** of what the text was about I quickly sampled the text to find out what it was generally about. | 16 | 10 | 2 | 0 | 2 | **4.27** |
| **23** | When I wanted to **locate the main ideas,** I searched quickly and selectively through the text for them. | 13 | 12 | 3 | 2 | 0 | **4.20** |
| **24** | I decided **which ideas** were more important than others for this assignment. | 11 | 16 | 2 | 1 | 0 | **4.23** |
| **25** | It was important **to link** new information to what I already knew. | 2 | 9 | 11 | 6 | 2 | **3.10** |
| **26** | I worked out how **the main ideas** in each source text were linked together within that source text. | 6 | 18 | 5 | 1 | 0 | **3.97** |
| **27** | I tried to understand **how** the main ideas in the different source texts related to each other. | 13 | 11 | 6 | 0 | 0 | **4.23** |

In particular, scanning (question 21), careful local reading (question 19) and reading for gist (question 22) were reported. These also emerged as themes from the RTA.

During the RTA, four themes emerged as participants accounted for rereading sections of text. These themes were: to achieve a deeper understanding, to overcome a failure to understand, to extract information and lastly to overcome a lack of engagement. Given this study's focus on reading, each theme will now be explored.

*Table 30 RTA results: High level reading*

| **Total number of comments for High level reading processes** | | **99** |
|---|---|---|
| Sub themes | Deeper understanding | 17 |
| | Extracting info | 12 |
| | Failure to understand | 19 |
| | Gist | 19 |
| | Lack of engagement | 8 |
| | Scanning | 6 |
| | Search reading | 7 |
| | Targeted | 11 |

The first sub-theme of Deeper Understanding was applied to comments which seemed to suggest that rereading sections of the text furthered understanding, such as 'The more I did my reading, the more I could take from the article, the deeper my understanding was.' (P45 transcription line 69).

The second sub-theme 'Failure to understand' could perhaps be seen as another point on the same continuum as achieving a deeper understanding. Comments typical of this group were 'The information hasn't gone in, so I'm reading it over and over again until it does just sink in.' (P42 line 75) or 'I struggled to understand' (P47 line 33).

The third most common reason (12 comments) given for rereading areas of text was that participants were extracting information, usually in the form of note taking or paraphrasing. Comments such as P47 line 83, 'it's just copying what was in the text' are one such example.

The final theme to emerge relating to rereading sections of the text was a lack of engagement. This referred to times when the participants reported their eyes going over a section of text, but subsequently having no idea what it had said. One such example was:

> *Interviewee: Do you know when you're reading something, and you're not thinking about it? You're just reading it and that's what happens to me.*

P14 line 77

In addition to participants reporting some of the types of reading reported in the literature, scanning, search reading and reading for gist, they also reported goal driven reading which the researcher has labelled as targeted reading. This refers comments when participants reported that they had a goal or purpose as they approached the text. They were not searching for a particular piece of information, but they had set themselves a goal / purpose for reading. Examples of this included:

*Interviewee: I wrote down what the key points were, the types of problem causes, solutions. I knew that I was going into it looking for these things.*

P42 line 7

This type of targeting reading suggests that some participants had an awareness of strategic reading and were deploying strategies at the appropriate stage as suggested by Kuzborska (2018).

#### 4.2.8.4 Organising

The questionnaire results for organising are reported in Table 31. Participants reported using a range of organising ideas, of which using a main theme (Q30) was most highly rated. Answers to Q31, 'I wrote about each text, one at a time' varied noticeably. Ten participants agreed and ten participants disagreed.

*Table 31 Questionnaire results: Organising*

| | | All participants | | | | | |
|---|---|---|---|---|---|---|---|
| **Question** | **Organising** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** |
| **28** | The source texts helped me decide what order to put the ideas in. | 9 | 9 | 4 | 8 | 0 | **3.63** |
| **29** | I organised the ideas into an order I thought of myself. | 4 | 11 | 6 | 5 | 4 | **3.20** |
| **30** | I had a main theme that helped me select ideas from the source texts and from my ideas. | 5 | 20 | 3 | 2 | 0 | **3.93** |
| **31** | While writing I wrote about each text, one at a time. | 5 | 10 | 4 | 10 | 1 | **3.27** |
| **32** | While writing I combined information from different texts to support the points I was making. | 5 | 15 | 4 | 5 | 1 | **3.60** |

Participant made relatively few comments in the RTA about organising their ideas. As shown in table Table 32, only one sub-theme emerged from the RTA data relating to organising. Perhaps this was because the task was relatively short and the participants all seemed to adopt one of two overarching plans (to use the structure of the task or to follow the order of the themes as they emerged from the texts).

*Table 32 RTA results for organising*

| **Total number of comments for Organising** | | **9** |
|---|---|---|
| Sub theme | Categorise | 9 |

To some extent the texts reflected the order of the task (problem, followed by solutions) and so perhaps there were limited options for organising the ideas.

#### 4.2.8.5 Micro-planning and translating

The questionnaire did not pose any questions about micro-planning and translating because these processes are often rapid and not always conscious. The researcher did not expect to elicit comments about micro-planning or translating during the RTA however, when asked to talk about writing and composing some participants did report processes which the researcher felt best fitted in this category. These comments were largely related to times when participants paused. Prompting referred to participants using the task instructions, the source texts or their own earlier work to prompt their next sentence.

*Table 33 RTA results for micro-planning and translating*

| **Total number of comments for Micro planning & translating** | | **19** |
|---|---|---|
| Sub themes | Prompting | 4 |
| | Reflecting | 5 |
| | Rewording | 10 |

Reflecting related to times when they paused and stared at part of the screen but then reported not really 'seeing', they were busy thinking about how to express an idea.

#### 4.2.8.6 Monitoring and revising

The question relating to monitoring and revising covered both high- and low-level revisions and also asked about when revisions took placed. Questions 33, 34,38,40 and 41 all relate to low level revisions, questions 35,36,37 and 39 relate to high level revisions. Questions 42 and 43 relate to the timing of revisions. As can be seen from Table 34, on the whole the rates of agreement are higher for low level revisions than for high level revisions. This suggests that participants engaged in greater levels of low-level revisions.

*Table 34 Questionnaire results for Monitoring and revising*

| Question | Monitoring and revising | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | Mean |
|---|---|---|---|---|---|---|---|
| **33** | I corrected spelling mistakes and typing errors. | 20 | 9 | 1 | 0 | 0 | **4.63** |
| **34** | I corrected grammar mistakes. | 17 | 12 | 1 | 0 | 0 | **4.53** |
| **35** | I made changes to the main ideas. | 5 | 5 | 5 | 12 | 3 | **2.90** |
| **36** | I made changes to the order of the paragraphs / ideas. | 7 | 6 | 1 | 9 | 7 | **2.90** |
| **37** | I made changes to ensure my ideas were clearly linked together. | 10 | 12 | 5 | 3 | 0 | **3.97** |
| **38** | I made changes to the vocabulary I had used. | 14 | 12 | 2 | 2 | 0 | **4.27** |
| **39** | I removed ideas that weren't relevant to the question. | 6 | 7 | 7 | 8 | 2 | **3.23** |
| **40** | I checked the quotations were properly formatted and referenced. | 2 | 6 | 9 | 3 | 10 | **2.57** |
| **41** | I checked that I had put the ideas from the source texts into my own words. | 12 | 13 | 2 | 1 | 2 | **4.07** |
| **42** | I edited my work while I was writing. | 11 | 9 | 3 | 5 | 2 | **3.73** |
| **43** | I started editing my work after I had finished most of the writing. | 10 | 9 | 2 | 7 | 2 | **3.60** |

However, as seen in Table 35, there were more references in the RTA data to high level revisions than low level revisions suggesting that the picture is likely to be mixed.

*Table 35 RTA results for monitoring and revising*

| | | |
|---|---|---|
| **Total number of comments for High level monitoring and revising** | | **59** |
| Sub themes | Amount of info | 6 |
| | Coherence / Cohesion | 23 |
| | Improving expression | 30 |
| **Total number of comments for Low level monitoring & revising** | | **36** |
| Sub themes | General | 6 |
| | Grammar | 9 |
| | Plagiarism | 6 |
| | Spelling | 8 |
| | Vocab | 7 |

This concludes the reporting of the RTA and questionnaire data. In summary the RTA and questionnaire data suggested that participants saw the source texts as the main source of information. Participants reported using strategies as they worked, and the researcher speculates that one of these strategies was to use the task instructions as a basis of their macro plan. Reading was most influenced by participants consideration of relevance to the question. The RTA data seemed to suggest that some participants approached reading the source texts with the goal of finding information to meet the key points of their macro plan. Participants reported focusing on their own emerging work to evaluate it, revise it and at times, as a prompt to generate more of their answer.

### 4.2.9 Summary of findings for RQ1

In summary, there were **four key findings** from RQ1 that characterised reading on the reading into writing task. They are as follows:

1. Participants paid different attention to the different AOIs. Participants spent almost **half their time (47 per cent) fixating on their own emerging** text and about **a third of their time reading the source texts.** By comparison the **task instructions were relatively poorly attended to** (6 per cent). Participants fixated on their own work to evaluate and revise their own work. At times, they used their own work to help generate the next part of their answer.
2. The **fixations on the written text were more homogeneous** than fixations across all the AOIs with lower standard deviations. These fixations also reported a **shorter mean fixation duration** and contained much **higher rates of regression** than for reading reported in the literature.

3. Whilst fixating on the source texts, **careful reading accounted for less than 30 per cent of reading**. Other forms of **selective reading were utilised for the remaining 70 per cent** of the time.
4. **Participants fixated more on the more relevant sentences**, although several factors seem to interact to determine total time on each sentence. However, the researcher speculates that fixation on the more relevant sentences may be due to some participants using their macro plan to approach the source texts with clear reading goals. These goals led some participants to target the most relevant areas of the text.

This concludes the analysis of the data in response to research question one. Having identified the characteristics of reading on a reading into writing task, the data was subdivided into two groups (Y1 undergraduates and year three undergraduates / postgraduates) and reanalysed to determine whether any differences were evident between the groups in answer to RQ2.

## 4.3 RQ2 What are the similarities and differences between the Y1 and Y3+ participants

The purpose of RQ2 is to consider whether any differences emerge between first year undergraduates (Y1) and third year undergraduates / postgraduate participants (Y3+). As with RQ1 the researcher considered the number of fixations (both total and averages) per participant or group of participants, the average fixation duration per participant or group of participants, as well as the total duration of all the fixations when added together. Therefore, the data can be summarised as the number of fixations, the average duration per fixation and the total duration of fixations (see Table 10).

*Table 36 Total number of fixations and total fixation duration between the Y1 undergraduate and year three / postgraduate groups*

| Participant Group | No. of fixations | Total fixation duration (ms) | Total fixation duration (hh:mm:ss) |
|---|---|---|---|
| Y1 undergraduates | 100,056 | 22,024,669 | 06:07:05 |
| Year three undergraduates and postgraduates | 114,996 | 27,405,410 | 07:36:45 |
| **Grand Total** | **215,052** | **49,430,079** | **13:43:50** |

Immediately, we can see that the year three / postgraduate group generated more fixations and fixated on screen for longer. Table 37 shows the keys statistics which characterised the data.

*Table 37 Summary of mean, SD and range of fixation data for Y1 / Y3+ groups*

| | Year 1 Under graduates | | | Y3 undergraduates / postgraduates | | |
|---|---|---|---|---|---|---|
| **Number of fixations per participant** | Mean | SD | Range | Mean | SD | Range |
| | 6,670 | 1966 | 5,894 | 7,666 | 1616 | 5,496 |
| **Total Fixation duration (hh:mm:ss) per participant** | Mean | SD | Range | Mean | SD | Range |
| | 0:24:28 | 00:07:36 | 00:24:50 | 00:30:27 | 00:06:59 | 00:22:12 |
| **Fixation duration (milliseconds) per fixation** | Mean | SD | Range | Mean | SD | Range |
| | 220 | 206 | 5,791 | 238 | 229 | 6,207 |

As reported above, the year three and postgraduate group (which shall be referred to as 'Y3+' hence forth) made a greater number of fixations, and spent more time fixating than the Y1 undergraduate group (which shall be referred to as the 'Y1' henceforth). We can also see from Table 37 that the Y1 group had a lower mean number of fixations (6,670) compared to the three plus group (7,666) and also spent less time fixating (24 mins per participant on average) than the three plus group (30 minutes per participant on average). We can also see that whilst the Y3+ group generated most fixations, the Y1 participants had a wider range of number of fixations per participant; there was greater variation between Y1 participants than between Y3+ participants. The same is true for total fixation duration.

Perhaps surprisingly, the Y1 group had a shorter mean duration per fixation (220 milliseconds) than the Y3+ group (238 milliseconds). It might have been expected that the more experienced, Y3+ group would be better readers, and this in turn might have led to them having shorter average fixation times. However, the longer mean fixation duration in the Y3+ group was due to their increased attention on their own work. The reason for this becomes clear in the next section where we discuss how the two groups distributed their fixations on the different AOIs.

In order to facilitate comparisons between the two groups across the different AOIs the data is presented as both actual numbers and percentages.

### 4.3.1 Y1 / Y3+ fixations across different AOIs

This section considers the way in which the participants spent time fixating on different AOIs. The researcher directs the reader back to section 4.2.2 to review the information about how the screen was divided into different areas of interest. The results for fixations on the different AOIs, when subdivided into two groups (Y1 and Y3+), are reported below. The data for the number of fixations are reported first (Table 38) and illustrated in Figure 29.

*Table 38 Y1 / Y3+ fixations across AOIs*

| **Number** of fixations | Y1 undergraduates | | | | Year three undergraduates and post graduates | | | |
|---|---|---|---|---|---|---|---|---|
| AOIs | No. of fixations | Mean | SD | % | No. of fixations | Mean | SD | % |
| 1) Own Work | 43,916 | 2928 | 1055 | 44% | 57,165 | 3811 | 917 | 50% |
| 2) Move Page | 1,503 | 100 | 75 | 2% | 1,863 | 124 | 56 | 2% |
| 3)Screen Instructions | 222 | 15 | 17 | 0% | 323 | 22 | 27 | 0% |
| 4) Task instructions | 5,718 | 381 | 192 | 6% | 6,637 | 442 | 203 | 6% |
| 5) Written text | 40,362 | 2691 | 891 | 40% | 39,022 | 2601 | 883 | 34% |
| 6) Diagrams | 2,866 | 191 | 125 | 3% | 4,333 | 289 | 171 | 4% |
| 7) Outside any AOI | 5,469 | 365 | 228 | 5% | 5,653 | 377 | 151 | 5% |
| Total | 100,056 | 6670 | 1966 | 100% | 114,996 | 7666 | 1616 | 100% |

The data from Table 38 is visualised in Figure 29. It is immediately noticeable that the Y3+ group fixated much more on their own work.

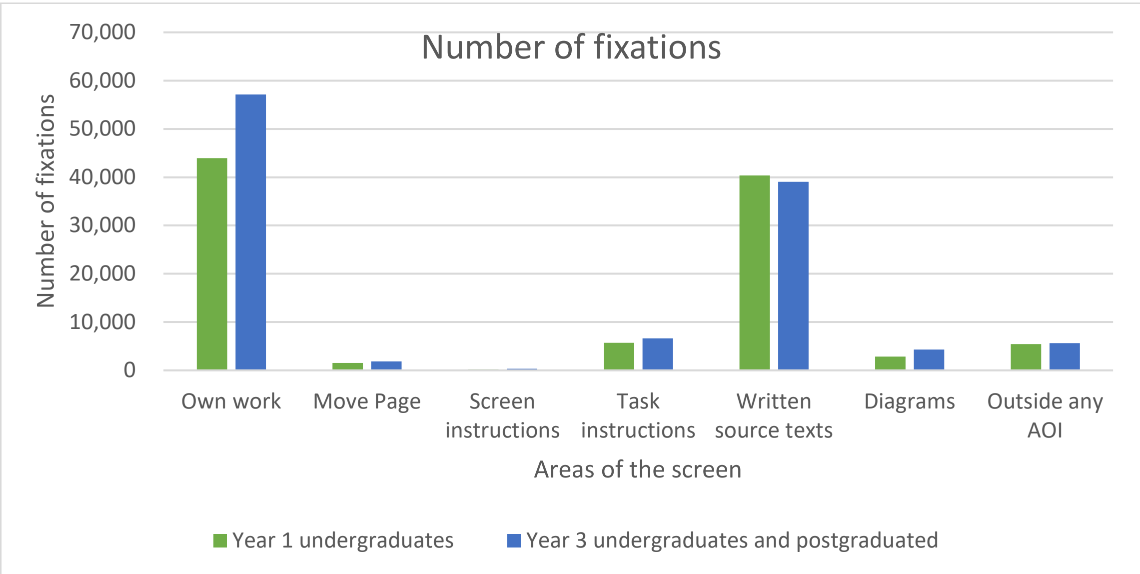



*Figure 29 Comparison of number of fixations across different AOIs for the Y1 and Y3+ groups*

As discussed in 4.2.1, fixations on the different areas of the screen were characterised by different mean fixations and different ranges of fixation lengths therefore, it is important to report not only the number of fixations but also the fixation duration information (Table 39).

*Table 39 Distribution of fixation duration across different AOIs for the Y1 and Y3+ groups*

| Fixation **duration** | Y1 | | | | Y3+ | | | |
|---|---|---|---|---|---|---|---|---|
| AOI | Total duration (hh:mm:ss) | Mean | SD | % | Total duration (hh:mm:ss) | Mean | SD | % |
| 1) Own Work | 03:13:31 | 00:12:54 | 00:05:41 | 53% | 04:28:08 | 00:17:53 | 00:05:06 | 59% |
| 2) Move Page | 00:06:52 | 00:00:27 | 00:00:27 | 2% | 00:09:38 | 00:00:39 | 00:00:26 | 2% |
| 3)Screen Instructions | 00:00:40 | 00:00:03 | 00:00:03 | 0% | 00:00:57 | 00:00:04 | 00:00:05 | 0% |
| 4) Task instructions | 00:18:11 | 00:01:13 | 00:00:36 | 5% | 00:22:10 | 00:01:29 | 00:00:42 | 5% |
| 5) Written text | 02:01:37 | 00:08:06 | 00:02:50 | 33% | 02:01:57 | 00:08:08 | 00:02:49 | 27% |
| 6) Diagrams | 00:09:58 | 00:00:40 | 00:00:25 | 3% | 00:15:21 | 00:01:01 | 00:00:41 | 3% |
| 7) Outside any AOIs | 00:16:15 | 00:01:05 | 00:00:41 | 4% | 00:18:34 | 00:01:14 | 00:00:39 | 4% |
| Total | 06:07:05 | 00:24:28 | 00:07:36 | 100% | 07:36:45 | 00:30:27 | 00:06:59 | 100% |

The data from Table 39 is visualised in Figure 31. From this we can see that whilst the two groups fixated on their own work for different amounts of time, they fixated on the written source texts for an almost identical amount of time.

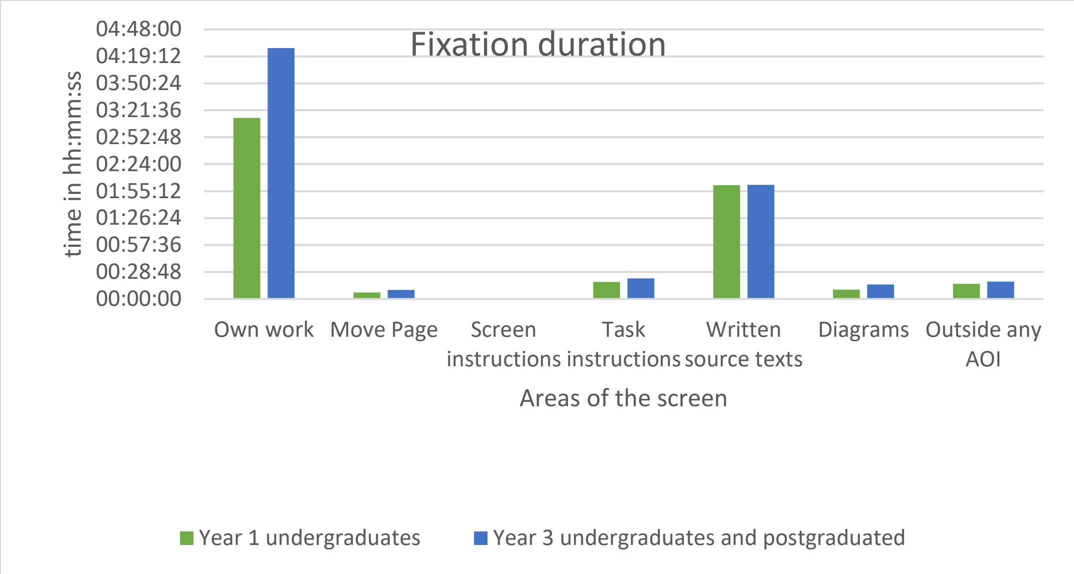



*Figure 30 Distribution of fixation duration across different AOIs for the Y1 and Y3+ groups.*

As we begin to analyse the data from Table 38, Table 39 and Figure 29 and Figure 30 we can see that the majority of the difference between the two groups can be accounted for by the differences in fixations on their own work. 50 per cent of the Y3+ group's fixations were on their own work and 59 per cent of their time was spent fixating on their own work. In comparison, the Y1 group allocated 44 per cent of their fixations and 53 per cent of their time to their own work. This is a key finding and helps to explain the longer mean fixation duration for the Y3+ group. As discussed in 4.2.2.1 fixations on participants' own work included some very long fixations which increased the mean fixation duration for this AOI.

Referring to Figure 29 and Figure 30 we can see that for all other AOIs, the differences between the groups were minimal. Examination of the mean fixation durations helps add further detail to our understanding of the differences between the groups and is reported in Table 40 below.

*Table 40 Mean fixation duration times for different AOIs*

| Mean fixation duration times ( milliseconds) | | | |
|---|---|---|---|
| Area of Screen | Y1 undergraduates | Year three undergraduates and postgraduates | All participants |
| Own work | 264 | 281 | 274 |
| Move page buttons | 274 | 310 | 294 |
| Screen instructions | 180 | 176 | 177 |
| Task instructions | 191 | 200 | 196 |
| Written Sources | 181 | 188 | 184 |
| Diagrams | 209 | 213 | 211 |
| Outside any AOI | 178 | 213 | 188 |
| Mean for all areas of screen | 220 | 238 | 229 |

If we examine the mean fixation duration for each AOI we can see that the Y3+ group had longer mean duration times than the Y1 group for every AOI except for the screen instructions and outside any area of interest.

In summary, the numbers of fixations and total fixation duration were rather similar for both groups across all AOIs *except* for time spent fixating on their *own work*. In total the Y1 group spent 367 minutes fixating on screen and the Y3+ group spent 456 mins fixating on screen; a difference of 89 minutes. The Y1 group spent 193 minutes fixation on their own work compared to 268 by the Y3+ group. A difference of 75 minutes. Therefore 84 per cent of the difference between the two groups can be accounted for by the differences in the time they spent fixating on their own work. The Y3+ group generally had higher mean durations per fixation, although the additional time they spent fixating on their own work, where long fixation durations prevailed, would have accounted for much of the difference between the mean duration per fixation for all fixations.

Having examined how the two groups compared over the different AOIs, we move on to consider the differences between attention devoted to individual sentences of the written source texts.

### 4.3.2 Y1 / Y3+ fixations on different sentences in the written sources.

As reported in section 4.3.1, fixations on the written source texts were similar for both groups. The Y1 group made 40,3621 fixations which added up to two hours, one minute and 37 seconds. The Y3+ group made 40,362 fixations which amounted to two hours, one minute and 57 seconds.

The total fixation duration for each group are amazingly similar, although because the Y1 group made a greater number of fixations, their average fixation duration time is shorter than the Y3+ group. Now we will compare whether those fixations were distributed across different sentences in the written source texts in a similar pattern. The researcher refers the reader back to 4.2.3 for a full explanation of the techniques used to complete this analysis, however, a brief reminder is included here.

By breaking the fixation data down into sentence level areas of interest it is possible to consider which parts of the text attracted more attention and to consider the relevance of each sentence to the reading into writing task that the participants were completing.

As reported in section 3.7.6, the length of each sentence in words was also measured. This enabled the researcher to compare the both number of fixations and total fixation duration per word for each sentence. This process enables easy comparison of the attention given to each sentence irrespective of sentence length.

The table reporting distribution of fixations across the different sentences of the written source texts for both groups is lengthy and cumbersome. Therefore, the researcher has not included the table here. Instead, the data is available in Appendix 11 (number of fixations) and Appendix 12 (total duration).

A graph which emphasises the key findings is included below. By comparing the total fixation duration per word of each sentence it is possible to visualise both the amount of attention that each sentence attracted, when adjusted for length, and compare whether both groups afforded each sentence similar amounts of attention.

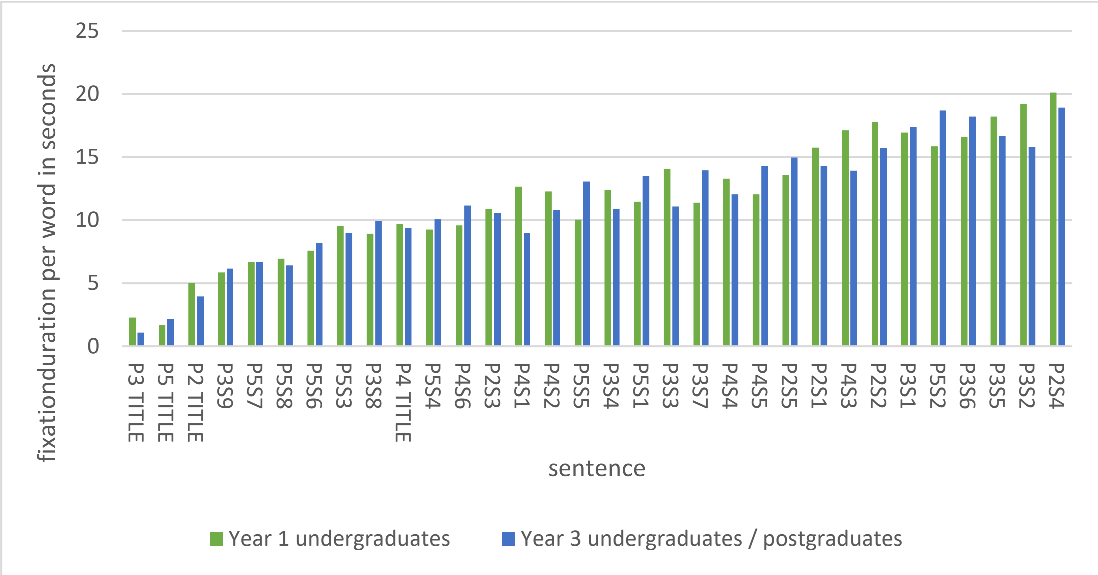



*Figure 31 Fixation duration per word of each sentence for Y1 undergraduate and year three undergraduate/ postgraduate groups*

The first thing to note is the similarity between the general pattern of both groups. Both groups gave very little attention to the titles, and both groups gave page two sentence four the most attention when adjusted for sentence length. However, beyond this the picture was rather difficult to fathom. Therefore, the next stage, when individual sentences were grouped according to relevance proved more useful in identifying differences between the groups.

### 4.3.3 Y1 / Y3+ fixations on sentences according to sentence relevance.

For a full explanation of the analysis of the distribution of attention across source text sentences according to sentence relevance please refer to section 3.7.7 however, a brief reminder is included here. To identify which sentences could expect to receive the most attention, the sentences were rated by three expert raters according to their relevance to the writing task. The sentences were rated using the following categories: 4-highly relevant (core ideas); 3-relevant (additional details); 2-less relevant for the task; 1-irrelevant. The total number of fixations made by Y1 undergraduates on sentences rated as highly relevant was divided by the total number of words in the sentences rated as highly-relevant. This procedure was repeated for each category of sentence and then the same calculations were performed for the year three / postgraduate group. The results are illustrated in Figure 32. The number of fixations analysed in the next section is slightly lower than the numbers reported in Table 38 because fixations on the titles (Y1: 410 fixations, Y3+:337 fixations) were excluded for reasons outlined in 3.7.7.

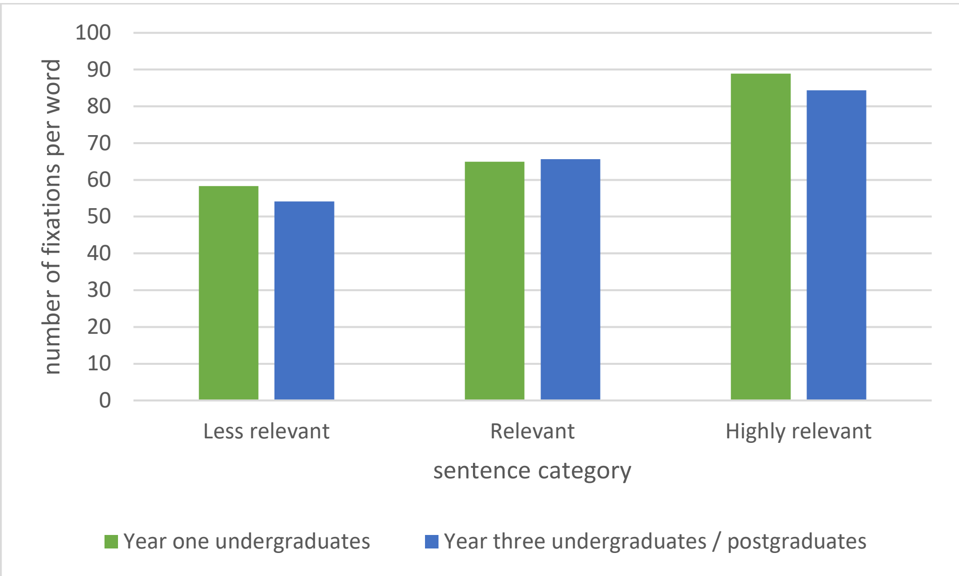



*Figure 32 Comparison of Y1 / year three -postgraduate fixations per word for each category of sentence relevance*

Figure 32 shows that both groups allocated consistently more fixations per word of text as the relevance of the sentences increased. The Y1 group made slightly more fixations (39,953 fixations) compared to the Y3+ group (38,685), therefore it is also useful to consider the number of fixations and total duration as percentages across the relevance categories as percentages.

*Table 41 Distribution of fixations across sentence relevance categories for Y1 undergraduates and Y3+*

| | **Number of fixations** | |
|---|---|---|
| | **Y1 undergraduates** | **Year three undergraduates /postgraduates** |
| Less relevant | 28% | 27% |
| Relevant | 41% | 42% |
| Highly relevant | 31% | 31% |
| **Total** | **100%** | **100%** |
| | **Fixation duration** | |
| | **Y1 undergraduates** | **Year three undergraduates /postgraduates** |
| Less relevant | 28% | 27% |
| Relevant | 41% | 42% |
| Highly relevant | 31% | 31% |
| **Total** | **100%** | **100%** |

We can see from Table 41 that the percentage of attention given to highly relevant sentences was the same for both groups. There is a very small difference between the two groups for the less relevant and relevant categories. The Y1 group allocated 1% more of their attention on the less relevant sentences and 1% less of their time on the relevant sentences than the Y3+ group. In order to assess the significance of the correlation between the relevance score and the fixation duration Pearson's correlation was run. There was a moderate correlation between relevance score and fixations duration per word for both groups although the Y3+ group (r=.503) reported a slightly higher correlation than for the Y1 group (r=.484).

In conclusion, **both the Y1 group and the Y3+ groups devoted their attention according to relevance**. The more relevant the sentence was, the more attention both groups focused upon on. This suggests that both groups were able to identify the areas of the text that contained the information required to successfully answer the question set by the task. The correlation between fixation duration and relevance was moderate for both groups but slightly stronger for the Y3+ group.

Having compared the way in which the two groups distributed their attention the across the written source text sentences, consideration will now be given to whether the characteristics of the fixations were similar for both groups.

### 4.3.4 Y1/ Y3+ movements between fixations

As described in section 3.7.8 of the methodology, fixations on the source texts can be classified according to whether they have moved forward through the text or regressed to an earlier part of the text when compared to the previous fixation. The distance between the previous fixation and the current fixation can also be a useful measure. When compared, these classifications can provide a quantifiable overview of the differences between the group's reading behaviour. Table 42 below, reports the fixations, categorised by distance and direction for the Y1 and Y3+ groups.

*Table 42 Comparison of characteristics of fixations for Y1 and year three / postgraduate groups*

| | Y1 undergraduates | | | | | Y3+ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixation type** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration / fixation (ms)** | **SD of duration / fixation** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration / fixation (ms)** | **SD of duration / fixation** |
| Short forward | 20,470 | 01:04:05 | 51% | 188 | 100 | 18,591 | 00:59:58 | 48% | 194 | 108 |
| Short regression | 12,340 | 00:36:35 | 31% | 178 | 93 | 11,819 | 00:36:47 | 30% | 187 | 106 |
| Long forward | 6,054 | 00:16:40 | 15% | 165 | 86 | 6,891 | 00:20:08 | 18% | 175 | 85 |
| Long regression | 1,498 | 00:04:17 | 4% | 172 | 88 | 1,721 | 00:05:03 | 4% | 176 | 97 |
| **Total** | **40,362** | **02:01:37** | **100%** | | | **39,022** | **02:01:57** | **100%** | | |

From Table 42 we can see that although the pattern of duration distribution is similar for both groups, there are small differences. For both groups, short forward moving fixations accounted for the largest proportion of their fixations. However, short forward moving fixations accounted for a greater proportion of the Y1 group's fixations (51 per cent) than for the Y3+ group's fixations (48 per cent). The Y1 undergraduates also made more short regressive fixations than the Y3+ group (31 per cent to 30 per cent respectively). However, the Y1 group made fewer long forward moving fixations than the Y3+ group (15 per cent to 18 per cent respectively). Both groups had four per cent of long regressive fixations.

When the results for two groups were compared using the Mann-Whitney U Test no statistically significant differences between the Y1 and Y3+ groups emerged in terms of the amounts of the different types of movement between fixations.

*Table 43 Results of Mann-Whitney U Test for reading type when comparing Y1 and Y3+ groups*

| Group | Number | Statistical Measure | Short Forward | Long Forward | Short Regression | Long Regression |
|---|---|---|---|---|---|---|
| Y1 | 15 | Median | 51.65% | 13.77% | 31.52% | 3.52% |
| Y3+ | 15 | Median | 46.71% | 18.92% | 31.60% | 4.16% |
| Mann-Whitney U | | | 83 | 63 | 101 | 82 |
| z | | | -1.224 | -2.053 | -.477 | -1.265 |
| *p* | | | .221 | .040 | .633 | .206 |
| *r* | | | .32 | .53 | .12 | .33 |

Therefore, in summary, the Y3+ groups engaged in a greater proportion of fixations which moved more than 16-character spaces forward through the text (long forward). Therefore, because this type of fixation is most likely to be associated with the expeditious reading outlined by Khalifa and Weir (2009:46) as readers make long jumps through the text, sampling sections in the case of scanning or searching for specific words, it suggests that the Y3+ group engaged in more expeditious reading than the Y1 group. The higher proportion of short forward moving and short regressive fixations amongst the Y1 participants suggests that they engaged in slightly more careful reading than the Y3+ participants. To determine more accurately whether this was the case, the patterns formed by fixations were compared.

### 4.3.5 Y1/Y3+ amount and type of reading

As described in 3.7.11 the patterns formed by fixations were analysed and each fixation was classified according to the type of reading pattern it fell within. For a full explanation of the classification procedure refer to sections 3.7.11.1 to 3.7.11.4 of the methodology chapter. The results of the classification of fixations were compared for the two groups and are summarised in Table 44.

*Table 44 Fixations classified by reading type for Y1 undergraduates and Y3+*

| | Y1 undergraduates | | | | | Y3+ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixation type** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration /fixation (ms)** | **SD Duration per fixation** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration /fixation (ms)** | **SD Duration per fixation** |
| Careful Local | 4,172 | 00:13:26 | 10% | 193 | 99 | 3,287 | 00:10:46 | 8% | 197 | 106 |
| Careful Global | 8,500 | 00:28:00 | 21% | 198 | 103 | 7,015 | 00:23:58 | 18% | 205 | 108 |
| Selective Local | 14,065 | 00:40:23 | 35% | 172 | 92 | 14,356 | 00:43:03 | 37% | 180 | 102 |
| Selective Global | 13,625 | 00:39:48 | 34% | 175 | 92 | 14,364 | 00:44:10 | 37% | 185 | 100 |
| Total | 40,362 | 02:01:37 | | | | 39,022 | 02:01:57 | | | |

When the results for two groups were compared using the Mann-Whitney U Test no statistically significant differences in the amounts of the different types of reading between Y1 and Y3+ participants were revealed.

*Table 45 Results of Mann-Whitney U Test for reading type when comparing Y1 and Y3+ groups*

| Group | Number | Statistical Measure | Careful local | Careful Global | Selective Local | Selective Global |
|---|---|---|---|---|---|---|
| Y1 | 15 | Median | 11.07% | 24.86% | 33.25% | 33.97% |
| Y3+ | 15 | Median | 8.42% | 14.12% | 39.63% | 36.41% |
| Mann-Whitney U | | | 80 | 73 | 83 | 76.5 |
| z | | | -1.348 | -1.638 | -1.224 | -1.493 |
| *p* | | | .178 | .101 | .221 | .135 |
| *r* | | | .35 | .42 | .32 | .39 |

Whilst the number and duration of fixations generated by both groups were similar, we can see from Table 44 that the patterns of eye-movements did vary between the groups.

The Y1 group invested a greater proportion of their time in both types of careful reading than the Y3+ group. To understand whether these differences were spread over the course of the whole task or concentrated in one phase of the task the frequency of fixations classified by reading type was plotted minute by minute as the task progressed.

The graph for selective global reading showed marked differences between the groups and is, therefore, presented first.

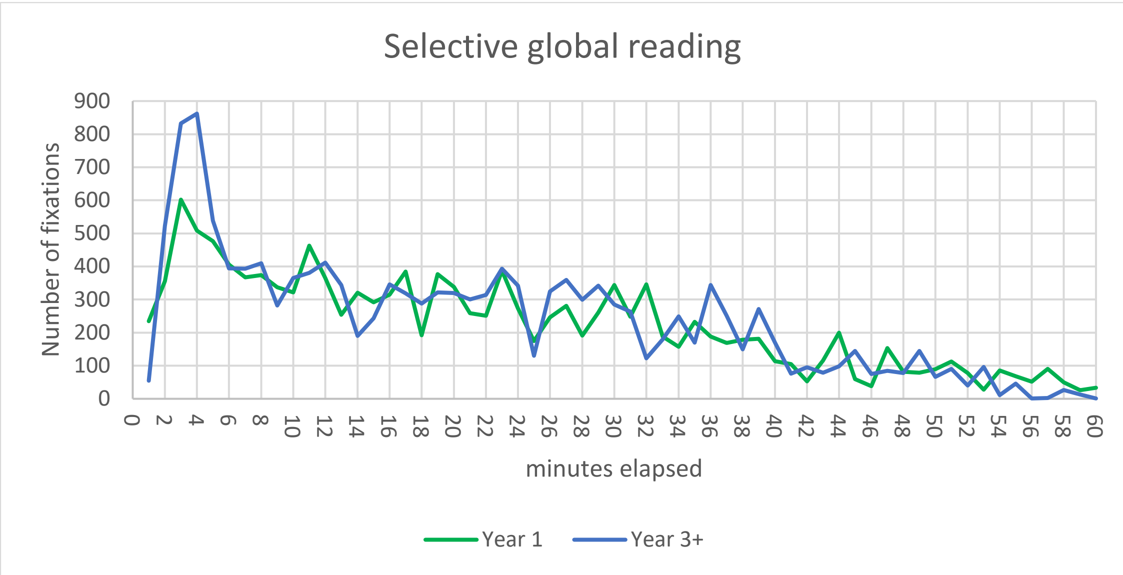



*Figure 33 Selective global reading during the task for Y1 and Y3+ groups*

The Y1 group reported 34 per cent of number and close to 40 minutes of fixation duration compared to 37 per cent and just over 44 minutes respectively for the Y3+ group. The Y3+ group (blue line) have a much higher number of fixations classified as selective global reading in minutes one to six of the task (Figure 33).

In contrast the  Y1 group had higher levels of fixations classified as part of the careful local reading during the early stages of the task (Figure 34).

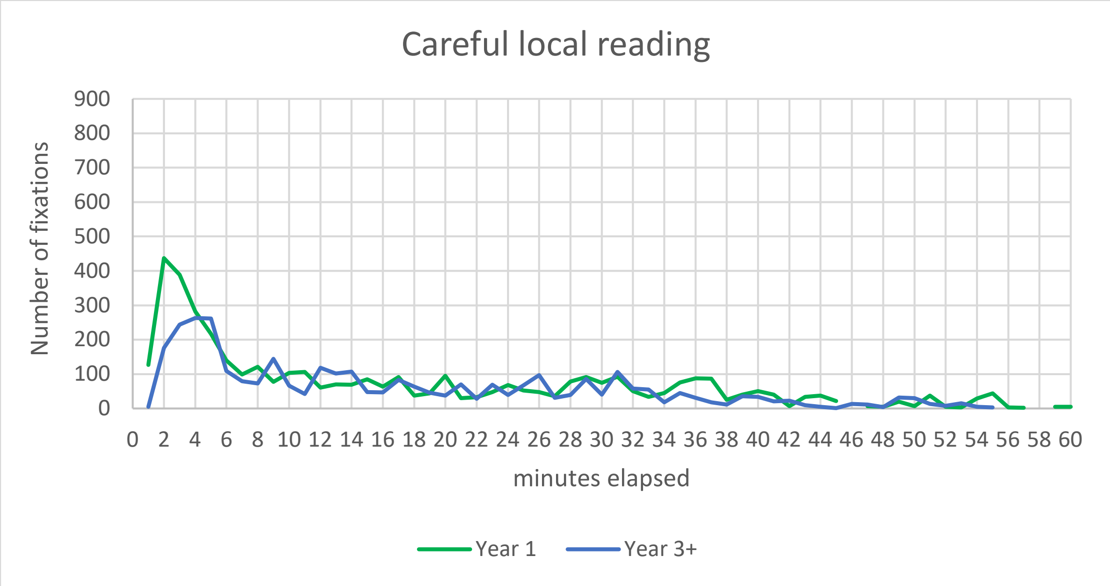



*Figure 34 Careful local reading during the task for Y1 and Y3+ groups*

For Y1 group 10 per cent of fixations and 13 and a half minutes of total duration were accounted for by careful local reading. This compared to 8 per cent and nearly 11 minutes for the Y3+ group.

The differences in careful local reading during the opening few minutes are not as marked as the differences between selective global reading during the same period. However, if the graphs for careful global reading (Figure 35) and selective local reading (Figure 36) are reviewed it becomes apparent that, for these types of reading, there are only small differences between the groups throughout the task. Therefore, the differences in selective global reading and careful local reading are the most notable differences between the two groups and account for the majority of the difference.

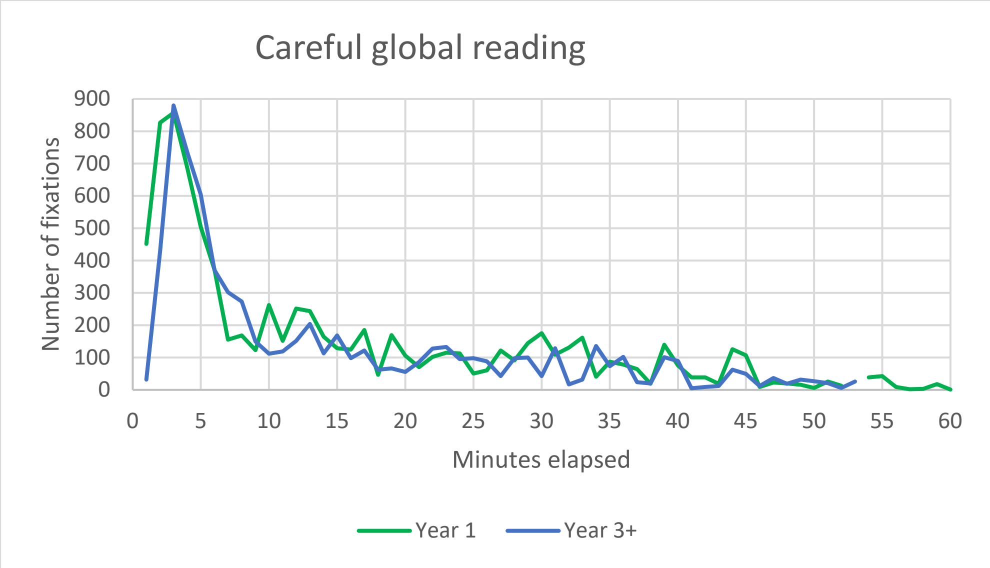



*Figure 35 Careful global reading during the task for Y1 and Y3+*

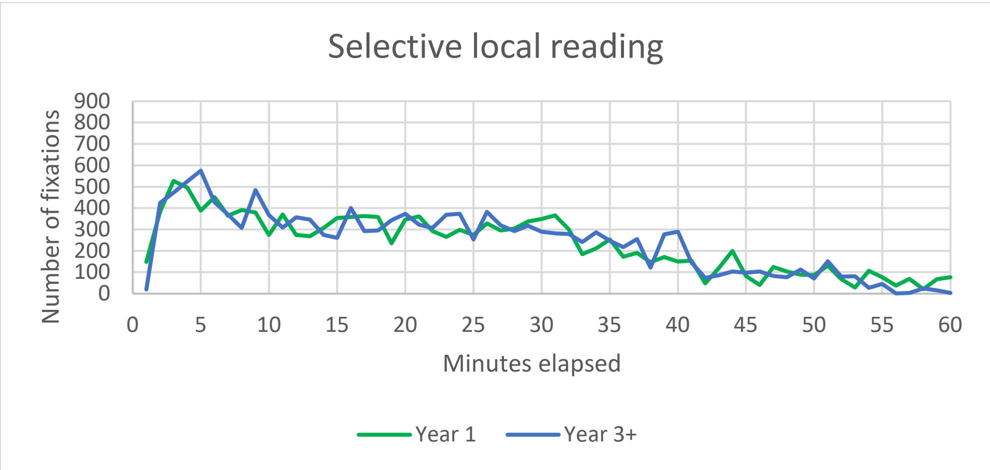



*Figure 36 Selective local reading during the task for Y1 and Y3+*

To summarise, the Y3+ group engaged in slightly more (5 per cent) expeditious reading than the Y1 year. The Y1 group engaged in 5 per cent more careful reading than the Y3+ group. These differences were largely due to differences during the early stages of the task. These are key findings.

Having considered the main differences between the two groups in terms of eye tracking data now the differences in the retrospective think aloud data and the questionnaire data will be reviewed.

### 4.3.6 Y1/Y3+ triangulating the eye tracking data

This study used two other research methods to assist with the interpretation and triangulation of the eye tracking data. These were the retrospective think-aloud (RTA) recordings made immediately after the participants finished the task and a short cognitive processing questionnaire after the RTA.

The RTA and questionnaire data for all the participants has been reported in sections 4.2.8.1. to 4.2.8.6. When reanalysed as two separate groups some of the data did not reveal any differences between the two groups. Therefore, here the researcher will only report the RTA and questionnaire data that shed light on the differences that emerged between the two groups. Appendix 15 includes a breakdown of the RTA data between the two groups and Appendix 16 includes a breakdown of the questionnaire data.

There were no notable differences between the questionnaire responses of the two groups. There were a few notable differences between the RTA data for the two groups. As there was an imbalance in the quantity of RTA data between the Y1 group and the Y3+ group (due to a failure of recording equipment) percentages are reported to facilitate easy comparison. The broad findings are illustrated in Figure 37.

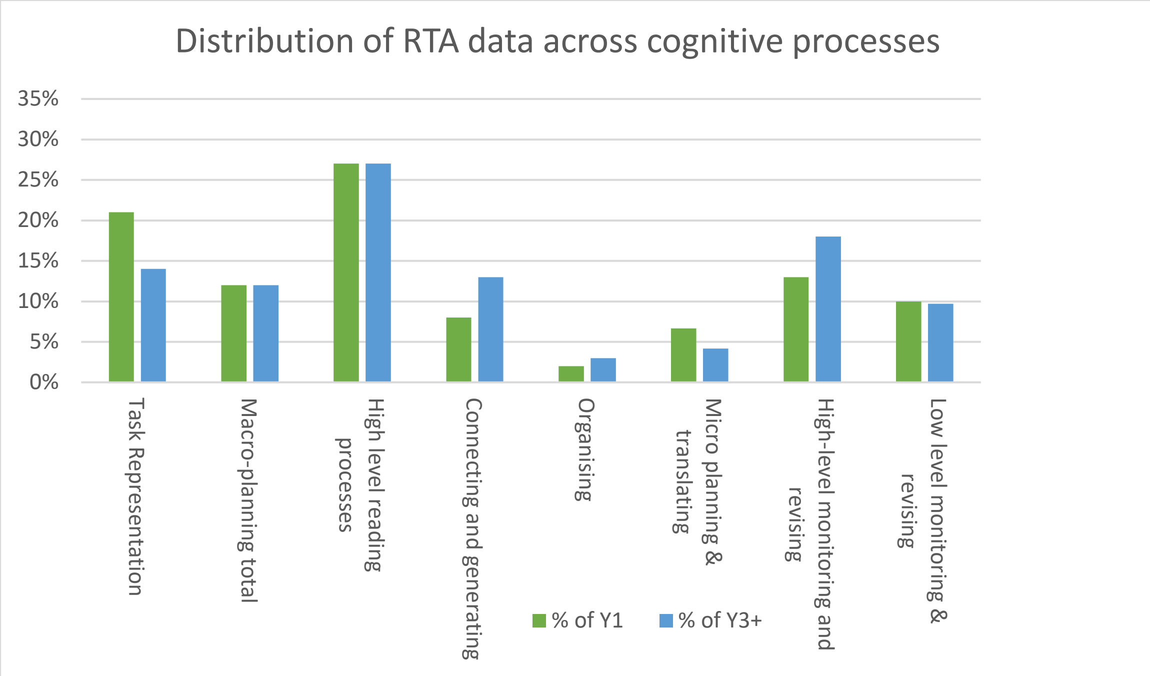



*Figure 37 Distribution of RTA data across cognitive processes*

The Y1 group made proportionally more references to task representation (21 per cent versus 14 per cent). Much of this difference related to the Y1 participants reporting that they set themselves goals *in addition* to the goals set out in the task instructions. These were goals such as P26 reporting that she wanted to see if her level of English had continued to improve.

Although both groups reported similar levels of comments on high level reading, the Y1 participants made more comments about reading to deepen their understanding than the Y3+ group (Y1:7 per cent, Y3+:3 per cent). The sub-category of Deeper Understanding was applied to comments which seemed to suggest that rereading sections of the text furthered understanding, such as 'The more I did my reading, the more I could take from the article' P45, line 69. The Y3+ group made more comments about reading for gist (7 per cent) than the Y1 group (3 per cent). However, as the researcher commented 4.2.8.3, these comments could have been grouped with failure to understand. Interestingly, the Y3+ participants made more comments about failing to understand than the Y1 participants. In conjunction with the Y1 participants making less comments about 'deepening understanding' this suggests that the Y1 participants felt they understood the texts better than the Y3+ participants. The other notable difference between the two groups in the high-level reading section was in relation to reading for gist. The Y3+ participants made more than twice as many references to reading for gist (7 per cent) compared to the Y1 participants (3 per cent). This suggests that the more experienced, Y3+ group were more aware of reading for gist as a strategy.

The Y3+ group also made more references to connecting and generating (13 per cent) than the Y1 group (3 per cent). The biggest difference was in the opinions sub-theme and included comments such as 'so I have a vague idea. I can start forming my own opinions if I needed to' P42, line 27.

Y1 participants made more references to micro-planning and translating (7 per cent versus four per cent) whilst the Y3 group made more references to high level revising. The differences in high level revisions largely centred around cohesion and coherence (Y1: three per cent, Y3+: 6 per cent).

This concludes the finding for the RTA and questionnaire data and therefore concluded the findings in relation to RQ2. The finds for RQ2 are summarised next.

### 4.3.7 Summary of the findings for RQ2

There were **two** key findings relating to differences between the way the Y1 and Y3+ groups tackled the task.

The first was that the **Y3+ group spent much more time fixating on their own work** than the Y1 group (4 hours 28 minutes versus three hours 13 minutes).

The second was that **the Y3+ group engaged in more selective** reading than the Y1 group (74 per cent versus 69 per cent). This inevitably meant that the Y3+ group engaged in less careful reading than the Y1 group (26 per cent versus 31 per cent). The increased levels of selective reading may have contributed to the greater attention that the Y3+ group devoted to the more relevant sentences.

Having reviewed the data relating to the differences between the Y1 and year three groups it suggests that whilst there are differences between the groups, they are quite small. When the participants tasks were scored, on average the Y3+ group scored more highly than the Y1 group. However, the results were not clear cut. Some Y1 participants scored very highly, whilst some Y3+ participants scored rather poorly. Therefore, the researcher decided to compare the five highest performing and five lowest performing participants. If the differences between the Y1 and year three groups were reflected and indeed increased between the highest and lowest scoring participants, this would assist with interpretation of the differences.

## 4.4 RQ3 What are the similarities and differences between the way high and low scoring participants tackle the task

For RQ3 the researcher went on to analyse all the information again, but this time to limit analysis to the five participants that achieved the highest marks and the five participants that were awarded the lowest marks on the reading into writing task. The high scoring group consisted of three participants from the Y3+ group and two from the Y1 group. The low scoring group consisted of three participants from the Y1 group and two from the Y3+ group. By comparing the highest and lowest scoring participants, the researcher hoped to clarify whether the differences between the Y1 and Y3+ group were related to experience of performing this type of academic task or were linked to academic ability more generally. If the differences between the highest and lowest scoring participants reflected similar, but more sizable differences than those seen between the Y1 and Y3+ groups this would seem to confirm that the differences related to reading into writing ability, rather than simply experience of academic reading into writing tasks.

The participants' responses to the test task were marked by the researcher and another experienced language assessor. The participants' responses were marked using the assessment criteria and rating scales for the task, a copy of which is included in Appendix 5. The rating scales are analytical and include three categories: a score for content, a score for organisation and a score for language. Each category is scored out of three, giving a maximum possible score of nine.

Both raters marked a batch of sample answers (none of the participant's responses were included in this sample). The marks awarded were then compared. Where differences emerged the interpretation of the assessment criteria was discussed and score revisions agreed. Once both raters felt comfortable using the rating scales the participants' responses were scored by both raters independently.

The level of agreement between the total scores awarded by the raters was calculated using Spearman correlation coefficient (because the scores were not normally distributed). The agreement for the content scores was $r = .710$, $p = 0.01$, agreement for the organisation scores was $r = .433$, $p = 0.05$, and agreement for the language scores was $r = .636$, $p = 0.01$. When added together the overall scores showed a strong positive correlation, $r = .830$ $p = 0.01$.

The scores from the two raters were averaged and the results are reported in Table 46. The average score was used to identify the five highest and five lowest scoring participants. Therefore, the data for P14, P32, P34, P36 and P53 were analysed for the low scoring group and the data for P23, P28, P46, P48 and P52 were analysed for the high scoring participants.

*Table 46 Scores awarded to participants' essays*

| Participant | Av of rater 1&2 Scores | Score category | Participant | Av of rater 1&2 Scores | Score category |
|---|---|---|---|---|---|
| P14 (Y1U) | 5 | Low scorers | P37 (Y1U) | 8 | |
| P32 (Y3+) | 6 | | P47 (Y3+) | 8 | |
| P34 (Y1U) | 6 | | P50 (Y3+) | 8 | |
| P36 (Y1U) | 6 | | P38 (Y3+) | 8.5 | |
| P53 (Y3+) | 6.5 | | P41 (Y3+) | 8.5 | |
| P11 (Y1U) | 7 | | P42 (Y3+) | 8.5 | |
| P26 (Y1U) | 7 | | P45 (Y3+) | 8.5 | |
| P55 (Y1U) | 7 | | P49 (Y3+) | 8.5 | |
| P15 (Y1U) | 7.5 | | P51 (Y3+) | 8.5 | |
| P31 (Y1U) | 7.5 | | P54 (Y3+) | 8.5 | |
| P40 (Y3+) | 7.5 | | P23 (Y1U) | 9 | High scorers |
| P12 (Y1U) | 8 | | P28 (Y1U) | 9 | |
| P25 (Y1U) | 8 | | P46 (Y3+) | 9 | |
| P29 (Y1U) | 8 | | P48 (Y3+) | 9 | |
| P35 (Y1U) | 8 | | P52 (Y3+) | 9 | |

The mean, standard deviation (SD) and ranges for both low and high scoring participants are set out in Table 47 below.

*Table 47 Summary of fixation information for low / high scoring participants*

| | Low scorers | | | High scorers | | |
|---|---|---|---|---|---|---|
| | Mean | SD | Range | Mean | SD | Range |
| No. of fixations | 5549 | 1287 | 3071 | 7397 | 1977 | 4859 |
| Duration per fixation (ms) | 273 | 296 | 5791 | 225 | 215 | 6207 |
| Total duration (hh:mm:ss) | 00:25:13 | 00:06:56 | 00:15:13 | 00:27:41 | 00:09:36 | 00:23:13 |

The number of fixations generated by each group are illustrated Figure 38. There is considerable variation in the number of fixations (low:27,744, high:36,985) but less relative difference in cumulative fixation durations (low: two hours and six minutes, high: two hours and 18 minutes). This suggests that the high scoring participants made a greater number of short fixations than the low scoring participants, as illustrated in Figure 38.

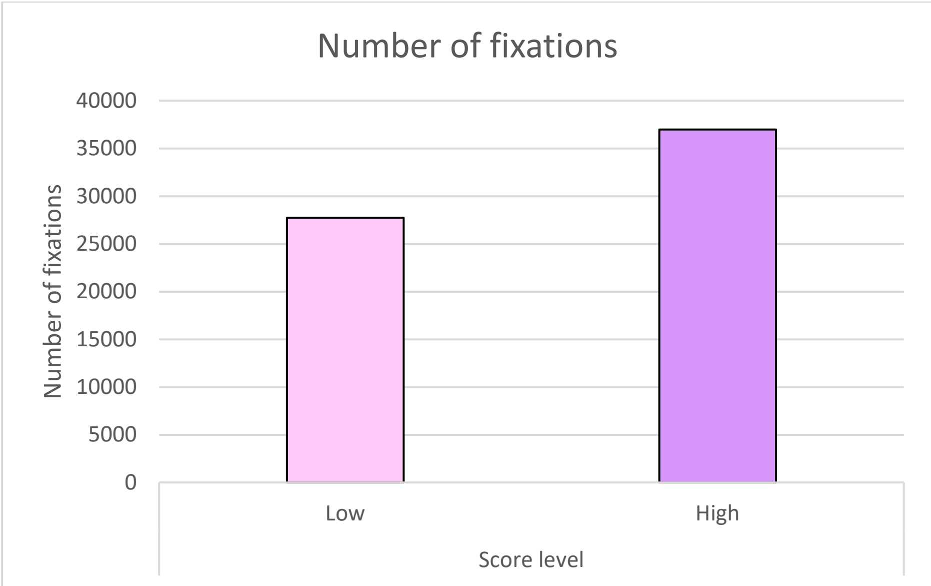



*Figure 38 Comparison of the number of fixations generated by low scoring / high scoring participants*

The low scoring participants also had a considerably longer mean fixation duration time (273 milliseconds) compared to the high scoring participants (225 milliseconds), as illustrated in Figure 39 below.

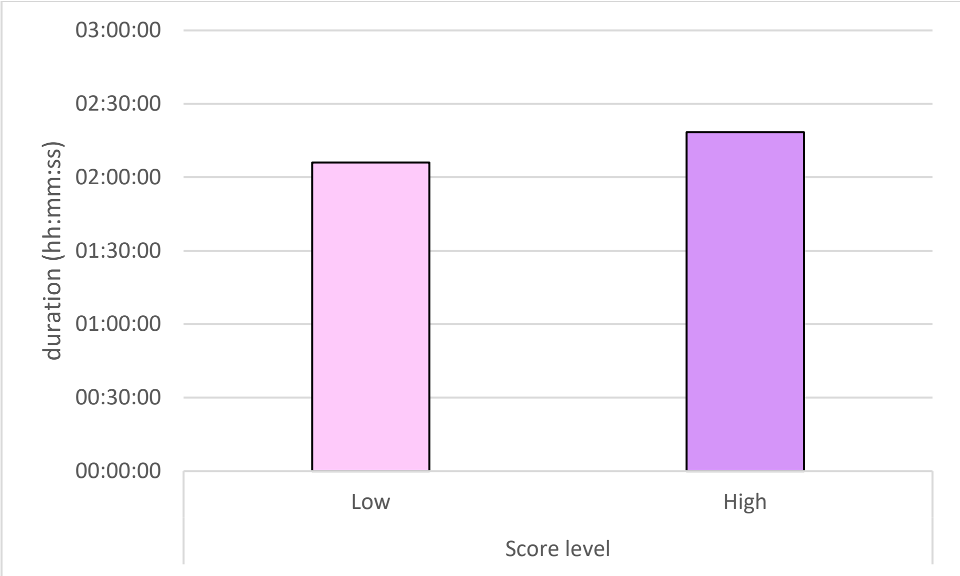



*Figure 39 Comparison of the total duration of fixations generated by the low / high scoring participants*

The differences between low and high score participants do not consistently continue the trends set by the Y1 and Y3+ groups.

*Table 48 Comparison of number of fixations, fixation durations and means for Y1, Y3+, low scoring and high scoring groups.*

| **Participant Group** | **No. of participants in group** | **No. of fixations** | **Total fixation duration (hh:mm:ss)** | **Mean fixation duration** | **Mean number of fixations / participants** | **Mean duration / participant hh:mm:ss** |
|---|---|---|---|---|---|---|
| Y1 undergraduates | 15 | 100,056 | 06:07:05 | 220 | 6670 | 00:24:28 |
| Year three undergraduates and postgraduates | 15 | 114,996 | 07:36:45 | 238 | 7666 | 00:30:27 |
| Lowest scoring participants | 5 | 27,744 | 02:06:05 | 273 | 5549 | 00:25:13 |
| Highest scoring participants | 5 | 36,985 | 02:18:25 | 225 | 7397 | 00:27:41 |
| **All participants** | **30** | **215,052** | **13:43:50** | 230 | 7168 | 00:27:28 |

There are larger differences between the mean fixation per participant for the low and high scoring groups than for the Y1 and Y3+ groups. The same is true of the mean duration per participant with larger differences between the Y1 and Y3+ group (24 mins 28 sec and 30 mins 27 secs respectively) than between the low and high scorers (25 mins 13 secs and 27 mins 28 secs respectively).

The mean fixation duration for the low and high scoring groups reverses the trend set for the Y1 / Y3+ groups. The high scorers have a lower mean fixation duration than the low scorers, but the Y3+ group have a higher mean fixation duration than Y1 group. The lower mean fixation duration could be an indication that the high scorers are generally better readers than the low scoring group, although because the fixations at this stage serve a wide range of purposes (reading own work, source texts, instructions, looking at diagrams) it would be too much of a generalisation to make at this point.

Having considered the number of fixations and fixation durations we move on to consider how those fixations were distributed across different AOIs.

### 4.4.1 Low and high scorers' fixations across different AOIs

Figure 47 reports the results for the five lowest scoring participants (Low scorers) and the five highest scoring participants (High scorers).

*Table 49 Number of fixations across broad AOI for low / high scoring participants*

| | Low scorers | | | | High scorers | | | |
|---|---|---|---|---|---|---|---|---|
| | No. of fix | mean | SD | % | No. of fix | mean | SD | % |
| Total | 27,744 | 5,549 | 1,287 | 100 | 36,985 | 7,397 | 1,977 | 100 |
| Own work | 12,911 | 2,582 | 1,347 | 47 | 17,470 | 3,494 | 1,193 | 47 |
| Move Page | 385 | 77 | 53 | 1 | 695 | 139 | 63 | 2 |
| Screen Instruction | 71 | 14 | 20 | 0 | 67 | 13 | 12 | 0 |
| Task instruction | 1,458 | 292 | 90 | 5 | 1,926 | 385 | 114 | 5 |
| Written source texts | 10,550 | 2,110 | 597 | 38 | 13,627 | 2,725 | 934 | 37 |
| Diagrams | 828 | 166 | 150 | 3 | 1,455 | 291 | 148 | 4 |
| Not on any AOI | 1,541 | 308 | 238 | 6 | 1,745 | 349 | 165 | 5 |

Figure 40 illustrates that whilst high scoring participants made a much greater number of fixations (low: 27,744, high: 36,985), when calculated as a percentage, those fixations were similarly distributed across the different AOIs.

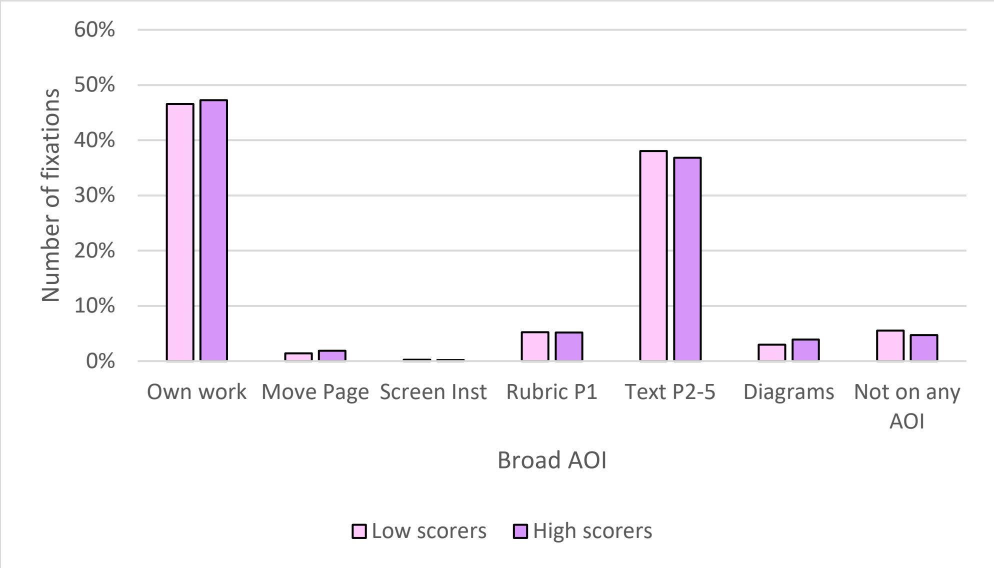



*Figure 40 Distribution of low / high scoring participants' fixations across broad AOIs*

The data for the number of fixations of low and high scoring participants does not continue to reflect the patterns that emerged for the Y1 and Y3+ groups as can be seen in Figure 41 .

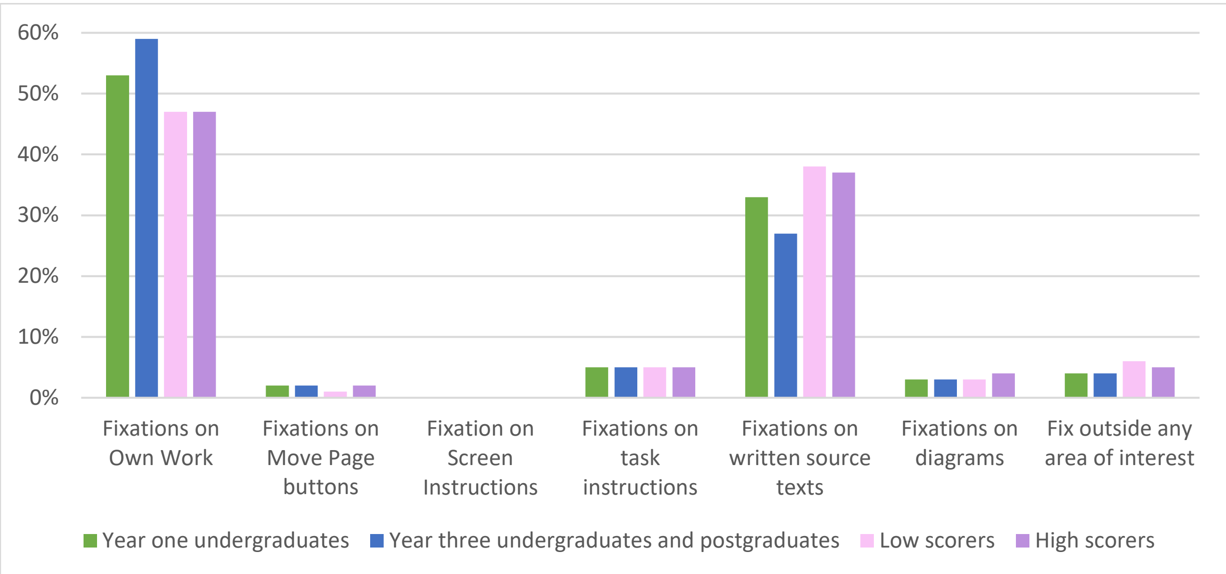



*Figure 41 Comparison of distribution of fixations across different AOIs for Y1, Y3+, low scoring and high scoring groups*

Indeed, if we review Figure 41, we can clearly see that the differences between the Y1 and Y3+ group (blue and green bars) show much greater disparity than between the low and high scorers (pink and purple bars). The same is true for the distribution of the fixation duration across the different AOIs when viewed as percentages. The researcher will not report it here for reasons of brevity, but the data can be found in the Appendix 17.

In summary, the trend seen in the Y1 and year three groups for the Y3+ students to spend a greater proportion of time fixating on their own work is not reflected by the lowest and highest scoring participants. Therefore, it is reasonable to conclude that fixating more on your own work (probably in the form of rereading their own work) does not necessarily lead to an improved scored in reading into writing tasks.

Having reviewed the broad distribution of attention across the different AOIs by the high and low scoring participants their fixations on the sentences of the written source texts is reported next.

### 4.4.2 Low and high scorers' fixations on sentences according to relevance.

To remind the reader, when considering the distribution of attention across individual sentences of the written source texts for the Y1 and Y3+ groups (section 4.3.3) the researcher found it difficult to draw conclusions from the data at individual sentence level. Instead the researcher found the analysis of sentences when grouped by sentence relevance (section 4.3.4) more revealing. Therefore, we will skip the reporting of attention on individual sentences and go directly to comparing the low and high scorers' distribution of attention across source text sentences when grouped by sentence relevance. The data reporting the distribution of high and low scorers' fixations at individual sentence level can be found in Appendix 18 (number of fixations) and Appendix 19 (total duration).

*Table 50 Comparison of fixation data for high and low scoring participants when written source text sentences are grouped by relevance*

| | | Low scorers | | | High scorers | | |
|---|---|---|---|---|---|---|---|
| | No. of words | No. of fixations | Fixation duration | Mean fixation duration | No. of fixations | Fixation duration | Mean fixation duration |
| Less relevant sentences | 191 | 3036 | 00:10:37 | 210 | 3676 | 00:11:27 | 187 |
| Relevant sentences | 252 | 4668 | 00:15:32 | 200 | 5802 | 00:17:11 | 178 |
| Highly relevant sentences | 140 | 2735 | 00:09:39 | 212 | 4053 | 00:12:09 | 180 |
| Grand Total | 583 | 10439 | 00:35:48 | | 13531 | 00:40:48 | |

*Note: fixations on titles are excluded from this analysis*

From Table 50 we can see that the low scoring participants made significantly fewer fixations on the written source text sentences and hence had a lower total fixation duration than high scoring participants. We can also note that the low scoring participants have a longer mean fixation time on all categories of sentence than the high scoring participants. This contrasts with the trend for the Y1 group / Y3+ group where the Y1 group had a shorter mean fixation duration (181 milliseconds) compared to the Y3+ group (187) when fixating on the written source texts.

If we compare the amount of attention devoted per word across the relevance categories we can see that both the low and high scorers devoted more attention per word as the sentences became more relevant (see Figure 42 and Figure 43).


*Figure 42 Comparison of fixations per word for low and high scoring participants*

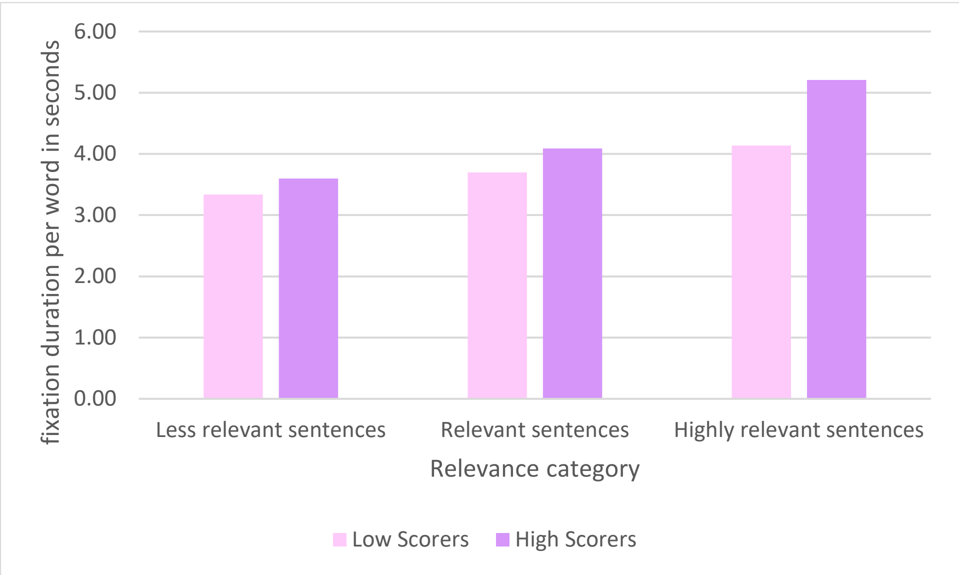



*Figure 43 Comparison of fixation duration per for low and high scoring participants*

In Figure 42 and Figure 43 we can see that as the sentence relevance increases, the difference between the low and high scoring participants seems to increase. In order to verify this the researcher compared the attention given to each sentence category as a percentage of the groups' total attention. See Table 51

*Table 51 Comparison of Y1 / Y3+ / low scorers / high scorers' attention to sentences grouped by relevance.*

| | Y1 under-graduates | Year three under-graduates /postgrad. | Low scorers | High scorers |
|---|---|---|---|---|
| **Number of fixations** | | | | |
| Less relevant | 28% | 27% | 29% | 27% |
| Relevant | 41% | 42% | 45% | 43% |
| Highly relevant | 31% | 31% | 26% | 30% |
| **Total** | **100%** | **100%** | 100% | 100% |
| **Fixation duration** | | | | |
| Less relevant | 28% | 27% | 30% | 28% |
| Relevant | 41% | 42% | 43% | 42% |
| Highly relevant | 31% | 31% | 27% | 30% |
| **Total** | **100%** | **100%** | 100% | 100% |

From Table  we can see that whilst there is no difference between the Y1 group and Y3+ group in terms of percentage of fixations and percentage of duration on the highly relevant sentences, this is not the case for the low and high scoring groups. The high scorers devote four per cent more fixations and three per cent more fixation duration to the most relevant sentences than the low scorers. This trend is reversed for the less relevant sentences, with the high scorers devoting two per cent less fixations and duration than the low scorers. These figures are not adjusted for sentence length, therefore the relevant sentences (252 words in total) receive more attention than either the less relevant (191 words) or the highly relevant (140 words) sentences. In order to check the strength of the relationship between sentence relevance and fixation duration (per word of each sentence) the researcher calculated the Pearson's correlation coefficient for the low and high scoring groups. For the high scoring group, the correlation was moderate (r=.411) but for the low scoring group there was only a small correlation (r=.165). Because the numbers of participants in each group was very small (five), individual differences are likely to have had a greater influence than for the

Y1 and Y3+ groups or for the entire group. Nevertheless, the difference between the low and high scoring groups is marked.

In summary, the data suggests that the high scorers devoted more of their attention towards the more relevant sentences than the low scorers. The differences between the correlation coefficients of the low and high scoring groups are more marked than the differences between the Y1 and Y3+ groups. This suggests that although sentence relevance alone is alone is not responsible for scoring well, an inability to spot the more relevant sentences inhibits scoring well.

Having compared the data relating to which parts of the written text the high and low scorers fixated on we turn to the characteristics of the fixations on the written source texts.

### 4.4.3 Low and high scorers' movements between fixations

As described in section 3.7.8 of the methodology, fixations on the source texts can be classified according to whether they have moved forward through the text or regressed to an earlier part of the text when compared to the previous fixation. The distance between the previous fixation and the current fixation can also be a useful measure. When considered together, these classifications can provide a quantifiable overview of a participants reading behaviour. Table 52 compares characteristics of the fixations on the source texts for the low and high scoring participants.

*Table 52 Summary of the low / high scoring participants fixations classified by direction and distance*

| | Low scoring participants | | | | High scoring participants | | | |
|---|---|---|---|---|---|---|---|---|
| | Number of fixations | | | | | | | |
| **Fixation type** | **Total No.** | **Mean** | **SD** | **%** | **Total No.** | **Mean** | **SD** | **%** |
| **Total** | **10,550** | **2,110** | **597** | **100%** | **13,627** | **2,725** | **934** | **100%** |
| Short forward | 5,080 | 1,016 | 391 | 48% | 5,902 | 1,180 | 5,902 | 43% |
| Short regression | 3,526 | 705 | 297 | 33% | 2,570 | 514 | 2,570 | 33% |
| Long forward | 1,626 | 325 | 78 | 15% | 4,492 | 898 | 4,492 | 19% |
| Long regression | 318 | 64 | 11 | 3% | 663 | 133 | 663 | 5% |
| | Fixation duration | | | | | | | |
| **Fixation type** | **Dur (hh:mm:ss)** | **Mean** | **SD** | **%** | **Dur (hh:mm:ss)** | **Mean** | **SD** | **%** |
| **Total Dur** | **00:36:09** | **00:07:14** | **14:41:02** | **100%** | **00:41:03** | **00:08:13** | **06:42:08** | **100%** |
| Short forward | 00:18:35 | 00:03:43 | 00:27:00 | 51% | 00:18:11 | 00:03:38 | 13:55:55 | 44% |
| Short regression | 00:11:38 | 00:02:20 | 09:02:00 | 32% | 00:13:44 | 00:02:45 | 18:39:45 | 33% |
| Long forward | 00:04:56 | 00:00:59 | 13:36:34 | 14% | 00:07:16 | 00:01:27 | 18:15:17 | 18% |
| Long regression | 00:00:59 | 00:00:12 | 23:55:51 | 3% | 00:01:53 | 00:00:23 | 15:27:05 | 5% |

From Table 52 we can see that whilst the high scoring group generated 29 per cent more fixations than the low scoring group, the high scoring group only generated 14 per cent more fixation duration. This suggests that the high scoring group are making a greater proportion of short duration fixations. To check this, the researcher plotted frequency of fixations in each duration group for each fixation type (please refer back to section 4.2.2.5for a full explanation of this process).

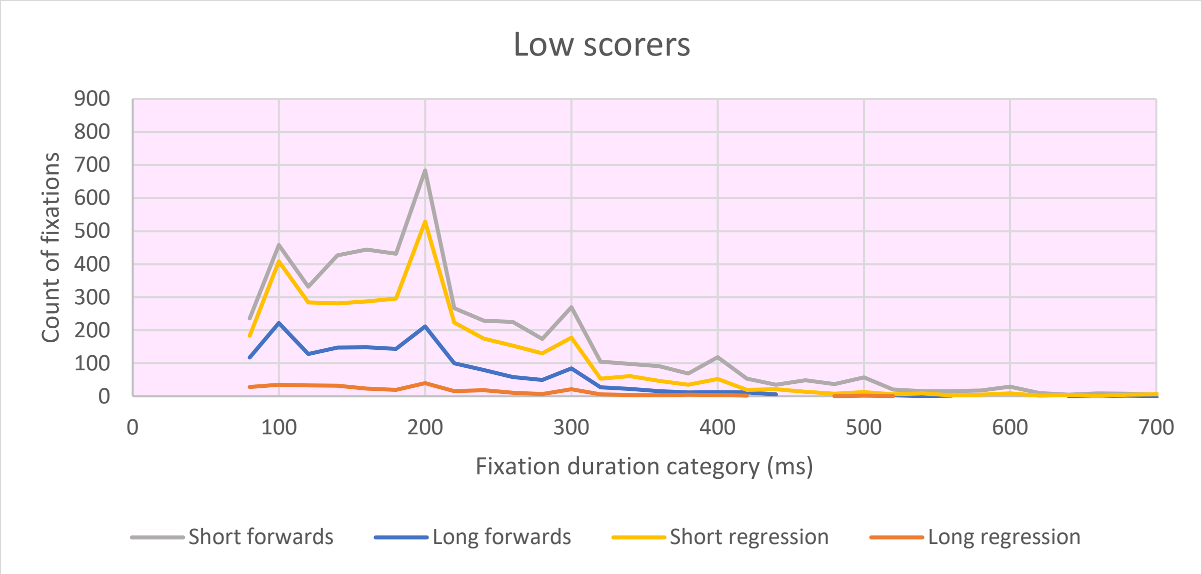



*Figure 44 Low scoring participants: The frequency of fixations grouped by directions/distance when categorised by fixation duration*

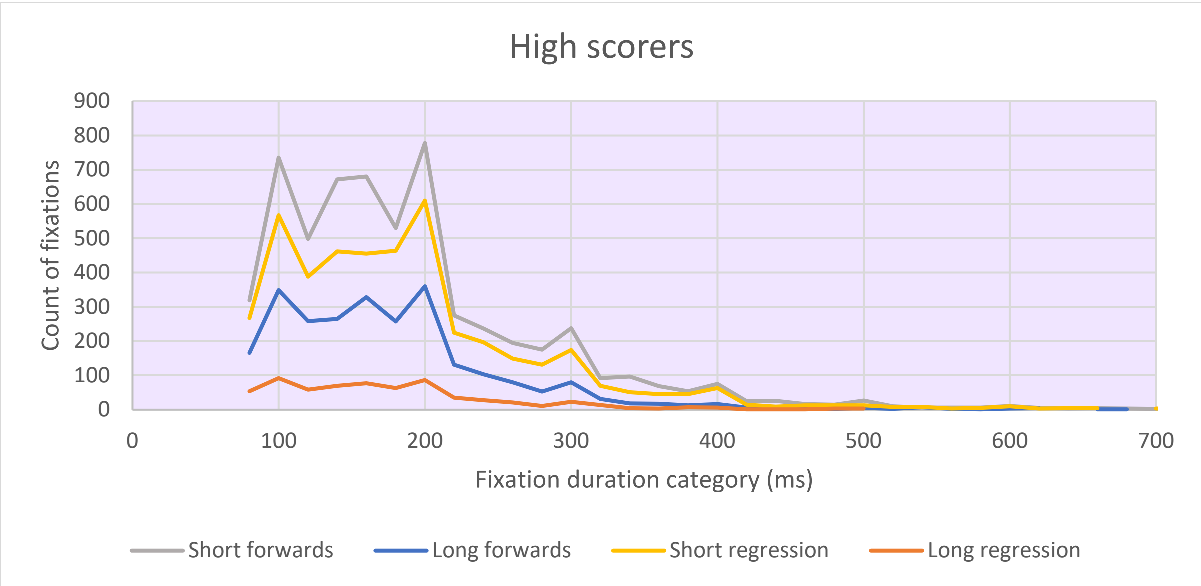



*Figure 45 High scoring participants: The frequency of fixations grouped by direction/distance when categorised by fixation duration*

By reviewing Figure 44 and Figure 45 we can see that the high scoring participants generated a greater number of short duration fixations across every type of fixation.

In order to more easily review the percentages of fixations / fixation duration across the different categories of fixation and compare them to the results for the Y1 and Y3+ groups the researcher has summarised the percentage in Table 53.

*Table 53 Comparison of Y1 / Y3+, low and high scorers' percentages of fixation characteristics*

| | Percentage of fixations | | | | Percentage of duration | | | |
|---|---|---|---|---|---|---|---|---|
| | Y1 under-graduates | Year three undergrads & postgrads | Low scorers | High scorers | Y1 under-graduates | Year three undergrads & postgrads | Low scorers | High scorers |
| **Total** | **100%** | **100%** | **100%** | **100%** | **100%** | **100%** | **100%** | **100%** |
| Short forward | 51% | 48% | 48% | 43% | 53% | 49% | 51% | 44% |
| Short regression | 31% | 30% | 33% | 33% | 30% | 30% | 32% | 33% |
| Long forward | 15% | 18% | 15% | 19% | 14% | 17% | 14% | 18% |
| Long regression | 4% | 4% | 3% | 5% | 4% | 4% | 3% | 5% |

From Table 53 we can see that for all four groups the most common type of fixation was a short forward moving fixation. The tendency of the Y1 group to engage in more short forward moving fixations did not increase, in fact both the low and high scoring groups had lower percentages of short forward moving fixations compared to the Y1 and year three groups respectively. However, the high scoring group had a greater percentage fall than the low scoring group, leading to a larger percentage difference between the low / high than between the Y1/Y3+ groups. This pattern was also reflected in the duration of short forward moving fixations.

The picture for short regressive fixations is different. Here there is no consistent trend with differences between the Y1 and Y3+ groups (2 per cent) in terms for the number of fixations, disappearing for the low and high scoring groups. The fixation duration for short regressive fixations confounds this pattern with no difference between the Y1 and Y3+ group but a difference of one per cent between the low and high scoring groups.

Long forward fixations follow a similar pattern to the short forward fixations. Differences that emerge between the Y1 groups and Y3+ groups (the Y3+ group engage in three per cent more long forward; both number of fixation and fixation duration) increase to four per cent more for the high scorers compared to the low scorers.

There is no disparity in number or duration of long regressive fixations between the Y1 and Y3+ groups. A difference of two per cent emerges with the high scorers making more long regressive fixations than the low scorers.

In summary, both the Y1 and low scoring groups engaged in more short forward moving fixations than their partner group. The three per cent difference between the Y1 group and the Y3+ group increased to a 5 per cent difference between the low and high scoring participants. The tendency of the Y3+ group to engage in three per cent more long forward moving fixations than their partner group becomes more pronounced with the high scoring group producing four per cent more long forward moving fixations than the low scorers.

As discussed in the conclusion of section 4.3.4 because long forward moving fixations are more likely to be associated with the expeditious reading outlined by Khalifa and Weir (2009:46) as readers make long jumps through the text, sampling sections in the case of scanning or searching for specific words, it suggests that the high scorers engaged in more selective reading than the low scorers. The higher proportion of short forward moving fixations amongst the low scoring participants suggests that they engaged in slightly more careful reading than the high scoring participants.

## 4.4.4 Low and high scorers' amount and type of reading

Table 54 provides a summary of the fixation data for the fixations classified as belonging to different types of reading.

*Table 54 Summary of number and total duration of fixations by reading type for high / low scoring participants*

| | Low scorers | | | | | High scorers | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixation type** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration /fixation (ms)** | **SD Duration/fixation** | **No. of fixations** | **Total duration (hh:mm:ss)** | **Percentage of fixations** | **Mean duration /fixation (ms)** | **SD Duration /fixation** |
| Careful Local | 952 | 00:03:32 | 9% | 223 | 140 | 1,051 | 00:03:17 | 8% | 188 | 103 |
| Careful Global | 2,150 | 00:08:23 | 20% | 234 | 133 | 2,028 | 00:06:40 | 15% | 197 | 107 |
| Selective Local | 3,684 | 00:11:41 | 35% | 190 | 109 | 5,130 | 00:14:58 | 38% | 175 | 106 |
| Selective Global | 3,764 | 00:12:33 | 36% | 200 | 122 | 5,418 | 00:16:07 | 40% | 178 | 94 |
| Total | 10,550 | 00:36:09 | 100% | | | 13,627 | 00:41:03 | 100% | | |

Table 54 shows that the higher scoring participants engaged in a lower percentage of careful reading than the low scoring participants (careful local – low scorers: 23 per cent fixations and 26 per cent total duration versus high scorers: 18 per cent and 19 per cent respectively. Careful global – low scorers: 6 per cent fixations and 7 per cent total duration versus high scorers: 5 per cent and 6 per cent respectively). Instead the high scoring participants engaged in more selective reading activities (selective local- low scorers: 43 per cent fixations and 26 per cent total duration versus high scorers: 46 per cent and 30 per cent respectively. Selective global – low scorers: 27 per cent fixations and 41 per cent total duration versus high scorers: 31 per cent and 45 per cent respectively).

The difference between the low and high scorers' levels of selective global reading is clearly illustrated in Figure 46 where we can see an early peak in levels of selective global reading by the high scorers (purple line) during the first ten minutes of the task. Between six and ten minutes the high scoring participants' level of selective global reading drops away dramatically and during minutes 16 to 22, the low scoring participants engage in slightly more selective global reading. However, by minute 24 the high scorers are once again engaging in more selective global reading than their counterparts and this trend continues for the rest of the task.

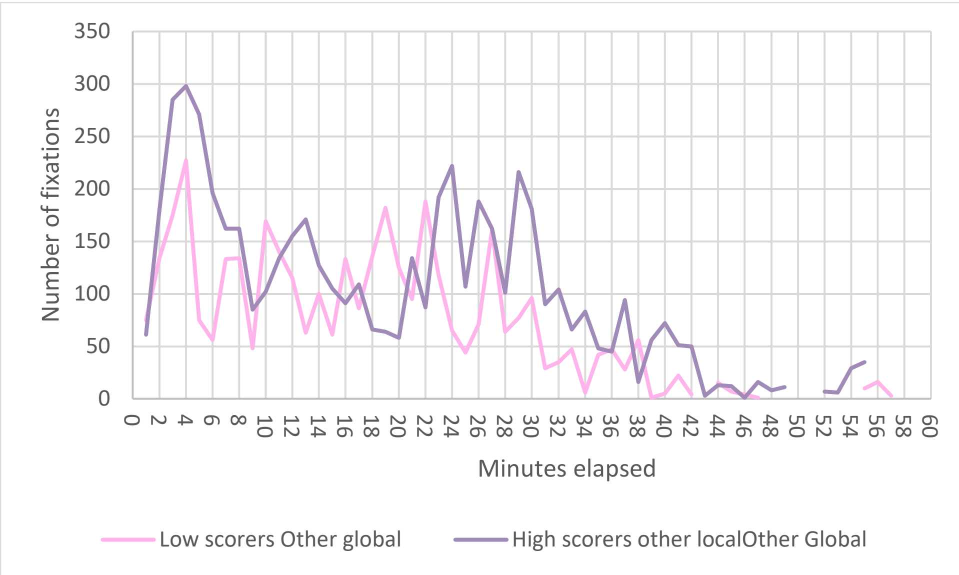



*Figure 46 Comparison of low scoring and high scoring selective global reading during the course of the task*

If we compare the percentages of fixation duration for the high and low scoring groups with the percentages for the Y1 and Y3+ groups (Figure 47), we can see that the small differences between the Y1 and year three groups increase slightly for the low and high scoring groups.

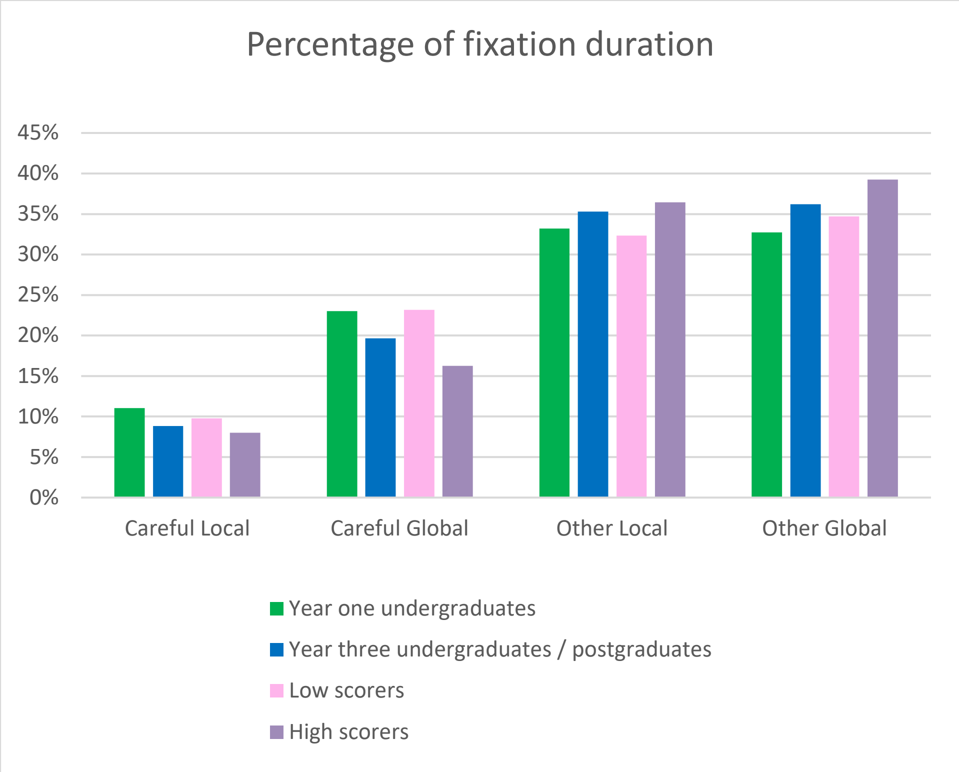



*Figure 47 Comparison of percentage of fixation duration classified by reading type for Y1. Y3+, low scoring ang high scoring groups.*

In summary, the high scorers produced a much larger number of fixations and this inevitably resulted in a longer fixation duration total than for the lower scorers. However, this reverses the trend seen in Y1 and Y3+ participants (with Y1 participants generating a lot more fixations) therefore the number of fixations on written source texts does not appear to be linked to scores achieved.

The tendency for more experienced (Y3+) participants to engage in more selective reading (selective local and selective global reading) than their less experienced counterparts (Y1) increases when we consider the high / low performing groups. The high performers spend more time engaged in selective reading whilst, inevitably, engaging in less careful reading. Thus, it seems reasonable to conclude that higher levels of selective reading could be linked to improved scores.

### 4.4.5 Triangulating the eye tracking data

This study used two other research methods to assist with the interpretation and triangulation of the eye tracking data. These were the retrospective think-aloud (RTA) recordings made immediately after the participants finished the task and a short cognitive processing questionnaire after the RTA.

The RTA and questionnaire data for all the participants has been reported in section 4.2.8 However, with only five participants in each of high and low scoring groups limited conclusions can be drawn from the data. Therefore, here the researcher will only report the RTA and questionnaire data that shed light on the differences that emerged between the two groups. Appendix 20 includes a breakdown of the RTA data between the high and low scoring groups and Appendix 21 includes a breakdown of the questionnaire data.

The number of RTA comments per cognitive process for the high and low scoring groups are illustrated in Figure 48 alongside the data from the Y1 and Y3+ groups. Although the trends in the Y1 / Y3+ data are reflected and generally increase in the categories for task interpretation, macro-planning, organising, micro planning and translating and high-level monitoring and revising, the researcher suggests that the very small number of participants in each group (5) makes generalising the data unsafe.

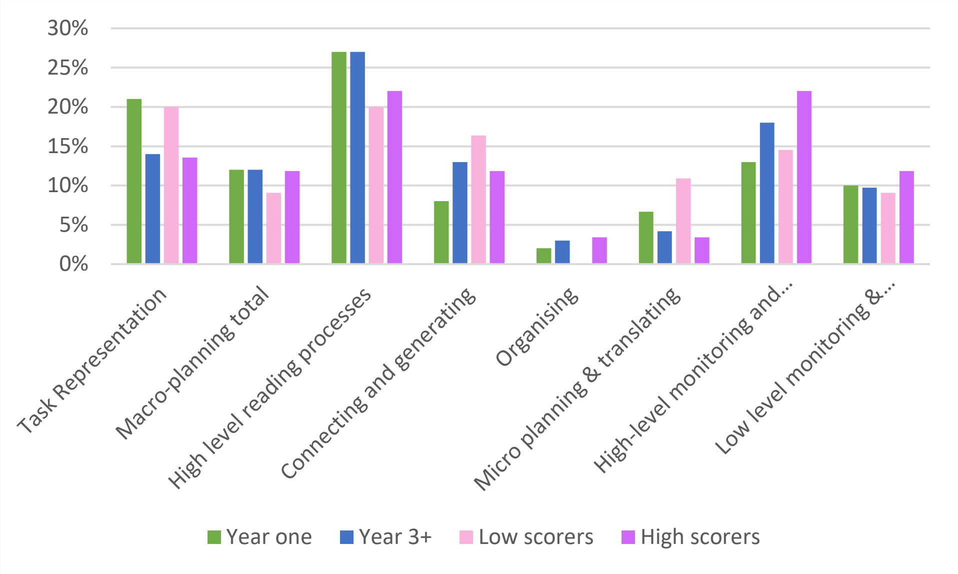



*Figure 48 Distribution of RTA data for Y1, Y3+, low scoring and high scoring participants across cognitive processes*

However, when reviewing the comments made by the individual participants, the researcher noted that whilst three of the five high scoring participants mentioned using the task instructions to guide their selection of information from the source texts, this was not mentioned by any of the lowest scoring participants. In the questionnaire some of the largest differences between mean scores related to task representation. Question 3, 'I thought I need to include a lot on information from my memory' perhaps suggested that the low scorers were more inclined to draw from their memory than seeing the source texts as the main source of information (low 3.0, high 1.2). Question 6, 'My understanding of the question changed whilst I was completing the task' suggested that the high scoring group continued to revaluate what the question asked them to do as they worked (high 3.6, low 2.8).

This theme of evaluation emerged again from the RTA comments concerning revising their work. The high scoring participants made a lot more comments about high level revisions (low 15 per cent, high 22 per cent). Their comments seem to suggest that they actively reflected on the demands of the task and considered the extent to which they had met the task demands whilst they were reading and writing. For example, participant 48 line 153 reports reflecting and making changes to his conclusion to try and be more decisive about his conclusions. In contrast the low scorers' remarks related more to ensuring they had accurately paraphrased the written sources. The results of the questionnaire seemed to reinforce the theme that the high scoring participants evaluated their own work more. One of the largest differences in the means of the two groups (high 4.2, low 3.0) was for question 43 which read 'I started to edit my work after I had finished most the writing'. This suggests that they took time to carefully review what they had written once their answer was largely complete.

### 4.4.6 Summary of the findings for RQ3

This findings for RQ3 have reported where the data for the high and low scoring groups suggest key differences in the way participants tackled the task. As a result, the key finding for RQ3 is that the high scoring participants, in common with the Y3+ participants, engaged in **more selective reading.** For the high scorers this selective reading seems to have been devoted to the more relevant sentences in the source texts. Conversely, **the low scoring participants were much less effective at focusing their attention on the most relevant sentences.**

This concludes the findings of RQ3 and therefore, concludes the findings of all three research questions. The next chapter goes on to discuss the broader implications of the findings reported in this chapter.

# 5 Discussion

## 5.1 Introduction

This chapter discusses the findings reported in Chapter four and attempts to draw the findings from all three research questions together into broader understanding of the way the participants tackled the reading into writing task. The discussion will follow the order of the research questions. To remind the reader, the research questions are 'what are the characteristics of reading during an academic reading into writing test task?' (5.2), 'what are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?' (5.3) and finally, 'what are the similarities and differences between the way high and low scoring participants tackle an academic reading into writing test task?' (5.4).

The findings offer insight into the way participants interacted with the task instructions, the way participants read the written source texts and the way participants interacted with their emerging answers (own work). Each of these areas of interest has clear links to different stages of the reading into writing process. Therefore, for each research question, the discussion will be structured as follows:

**Task instructions** - discussion of the insight offered for task representation and macro-planning processes

**Written source texts** - discussion of the insight offered for the high-level reading, organising, connecting and generating processes

**Own work** - discussion of the insight offered for the micro planning and monitoring and revising processes.

Before discussing the findings of the individual research questions the researcher will briefly comment on the role of individual differences.

### 5.1.1 Individual differences

One of the findings, in line with the literature (Pichert and Anderson,1977, Anderson, Pichert and Larry, 1983, Rayner et al., 2012, Rothkopf and Billington, 1979), was that there was a large amount of individual variation. This variation took the form of variation of number of fixations (mean 71068, standard deviation 1839) and mean fixation duration (mean 230 milliseconds, standard deviation 219). The large differences between participants in terms of number of fixations and mean durations are to be expected as there are several factors which can interact to influence fixations patterns. As long ago as 1979 Rothkopf and Billington concluded that differences in individual eye-movement patterns were marked and it is widely accepted that reading skill influences patterns (Rayner et al., 2012) with more able readers making shorter fixations and longer saccades. It has also been demonstrated that readers vary their rate of reading and the information they pay attention to according to their purpose for reading (Pichert and Anderson, 1977, Anderson, Pichert and Larry, 1983). However, in a large-scale study Fisher (1983) concluded that individual differences remain consistent across tasks, for example when changing from careful reading to search reading.

Given the factors discussed above, large variations between participants in the number and duration of fixations are inevitable. However, it is important to explain the implications of this individual variation for this study. Firstly, this study is interested in establishing if there are metrics or patterns of behaviour that apply broadly across the group, irrespective of individual differences. Secondly, this study is interested in comparing groups of participants (for example first year and third year / postgraduate participants). By converting the group data to percentages and comparing percentage differences between groups, individual differences are ameliorated. In conclusion, although there were large differences between participants in terms of both the number of fixations generated and total duration, this is to be expected and does not obscure the findings.

Despite the individual differences, there were notable processes and trends that emerged from the data. These are now discussed as they relate to the process of completing a reading into writing task.

## 5.2 Discussion of the findings for RQ1

RQ1 is undoubtedly a broad and rather open-ended question. However, because quantitative process research that suggests how reading might operate on a reading into writing task is missing in the literature, it was deemed appropriate to begin the research with a broad investigation of how reading on a reading into writing task operated. RQ1 produced four key findings (4.2.9).

The first key finding was an understanding of how attention was divided between the different parts of the task (the task instructions, the source texts and the participants' own work).

The second key finding was that fixations on the written source texts were much more homogenous than fixations across the different areas of interest. These fixations on the written source texts differed from eye movement data for reading in the literature because they had a shorter mean fixation duration and included much higher rates of regression.

The third key finding related to the amount of selective reading undertaken (70%) compared to the amount of careful reading (30%).

The last key finding related to the way participants interacted with the source texts, namely that different types of reading emerged from the data and that more attention was devoted to the most relevant sentences of the source texts.

The discussion of the first key finding spans all three sections covering the interaction with the task instructions 5.2.1, the interaction with the source texts 5.2.2 and the interaction with the participants' own work 5.2.3. Discussions of the second, third and fourth key findings are embedded within the section on interaction with the source texts. The homogeneity of the fixations on the written source texts is discussed in section 5.2.2.1. The different types of careful and selective reading are discussed in sections 5.2.2.2 to sections 5.2.2.7 and the attention on the most relevant sentences is discussed in section 5.2.2.8

The broadest findings, as reported in section 4.2.1, revealed that on average the participants spent about half of the time on task (55 minutes was the average time to complete the task) fixating on the screen (27 mins). Limited conclusions can be drawn from the fixation data from this stage of the analysis because the fixations on screen served a variety of purposes: reading the task instructions, reading the source texts, looking at diagrams as well as looking at emerging answers as they were typed. Therefore, the next stage of the analysis considered how participants divided their attention between different parts of the screen.

Participants devoted almost half of their attention (47 per cent of fixations) to their emerging work, just over a third of their attention (37 per cent) to the written source texts and just six per cent of their fixations to the task instructions. Initial analysis of the data suggested the data was not homogeneous, participants interactions with different parts of the task (the task instructions, the written source texts, the participants' own work) seemed to be characterised by differences in the key eye-movement metrics. The discussion begins with the interaction with the task instructions.

### 5.2.1 Interaction with task instructions

Whilst the low percentage of fixations (six per cent) on the task instructions seemed to suggest that the task instructions played a rather brief role in the process, the RTA data suggested that the participants interaction with the task instructions laid the foundations for their interaction with the source texts, having a profound influence on later reading activities. Therefore, this section starts by discussing the implications of how participants interacted with the task instructions.

The analysis of attention on the different areas of interest (4.2.2) indicated that participants fixated only half as many times on each word of the task instructions as on each word of the written texts, although the task instructions were read more slowly. The RTA data, the questionnaire data and the researchers own observations suggested that there were considerable differences between participants in terms of task representation and macro-planning. The researcher noted that several participants seemed to adopt a very structured, methodical approach which suggested that these participants had a well-rehearsed procedure which they adopted in order to tackle the task. This procedure tended to involve using the task instructions as the basis for their macro-plan, extracting key points from the task to form the outline or macro plan of their answer. Creating an outline of their answer before reading seemed to prime participants attention for the reading to come, as information was extracted from the texts to 'fill in the gaps' in the outline plan.

By contrast, some other participants opted to fixate briefly on the task instructions, not making any notes about the key demands of the task before going on to read the source texts. For example, P31 and P37 only made a cursory inspection of the task instructions before going on to read the source texts. In these cases, the participants seemed to delay engaging in task representation until after reading the source materials. Plakans (2008: 114) suggests that 'less skilled writers plan less initially and stop more often for local planning'. The researcher suggests that planning for reading into writing may also involve 'planning' *reading* or at least having a clear reading goal before starting to engage with the source texts. The researcher suggests that early development of a macro-plan helps clarify the purpose for reading, allowing the writer to engage in more goal driven reading.

Therefore, the researcher suggests that whilst only a small proportion of the participants' fixations were devoted to the task instructions, the participants' task representation process had a major influence on how the participants went on to fixate on and engage with the written source texts. Participants with a well-developed procedure for tackling such tasks developed macro-plans at the outset which set the agenda for their reading of the source texts. This suggests that there may be benefit in developing students' awareness of task representation, both in terms of visualising the finished product and reflecting on or giving explicit attention to the process that they will need to go through en route to creating the finished product.

Having discussed the interaction with the task instructions, the discussion moves on to consider the interactions with the source texts.

### 5.2.2 Interaction with the source texts

Participants devoted 37 per cent of their fixations to the written source texts. As reported in section 4.2.1 the eye tracking data on the written source texts provided a wide range of information about the way participants engaged with the source texts. All of this information suggested that reading during reading into writing tasks differs significantly to the careful reading described in the literature in two key respects. The first is that selective forms of reading account for more than 70 per cent of reading undertaken on the source texts. The second is that the more relevant parts of the texts are singled out for greater attention. The researcher suggests that these two factors are inextricably linked, not only with each other but as suggested above, with the task representation and macro-planning stages of reading into writing.

Chan (2013:73) proposes that following macro-planning the writer moves on to high-level reading using *'careful reading to create textual and or intertextual representations of the text'.* The researcher suggests that this study can add more detail to the process that takes place as writers engage with the text. The findings from this study suggest that to some extent, during the macro-planning stage, some writers have already developed preconceptions or criteria about the type of information they are looking for. Therefore, some participants used different forms of selective reading to establish what information the source texts offered and to extract a detailed understanding of the most relevant parts of the texts.

This section, discussing the findings in relation to participants' engagement with the source texts is organised as follows: firstly, the researcher will discuss the ways in which the eye tracking data in this study broadly differed from that described in the literature for careful reading and what this signifies. After that the researcher proposes some forms of selective reading that may be represented by periods classified as selective reading. Finally, the way in which the most relevant parts of the texts of texts were targeted will be discussed.

#### 5.2.2.1 Homogeneity of fixations on the written source text

The fixations of the written source texts were more homogeneous than fixations on other parts of the screen except for the screen instructions (4.2.2). Fixations on the written source texts had a low standard deviation of fixation duration (100 milliseconds) compared to mean fixation duration (184 milliseconds). Only the screen instructions had a lower standard deviation (SD: 87 milliseconds, mean 177 milliseconds). This illustrates that the fixations on the written source texts were much more homogeneous, with a much smaller range of fixation durations than fixations on, for example, the participants own work. The researcher speculates that the homogeneity of fixations on the written source texts could suggest that participants engaged in a narrower range of reading processes than when participants engaged with their own work where the range of processes might have included reading to evaluate, reading to revise and reading to help compose. This aspect is discussed further in section 5.2.3on participants' engagement with their own work.

With a mean fixation duration of 184 milliseconds, the fixations on the written source texts in this reading into writing task are shorter than the mean fixations reported in the literature for reading. For example, Rayner, Pollatsek et al. (2012) reported mean fixations durations of around 200 milliseconds for light fiction, through to a mean of 260 milliseconds for more complex scientific texts. Therefore, the mean of 184 milliseconds in this study is surprisingly short considering that the texts were constructed to represent the norms for higher education in the UK (3.6.1).

The researcher suggests that two key factors might have influenced the mean fixation duration, making it short by comparison to those reported in the literature for careful reading. The first factor may be the high amount of short regressions (30 per cent of all fixations on the written text). The mean fixation duration for short regressions in this study was 182 milliseconds, compared to 190 milliseconds for short forward moving fixations. Given that this study has reported much higher levels of regressions than in the literature and those regressive fixations have shorter durations this has the effect of reducing the mean fixation duration overall. There were no references to mean durations of regressive fixations in the literature therefore the researcher acknowledges that this suggestion might be speculative. However, 20 of the thirty participants in this study had shorter mean fixation durations for short regressive fixations than for short forward moving fixations, therefore this may be generalisable to the broader population.

The second factor might be the high rates of fixations after long saccades, which also had lower than average fixation durations. In this study, fixations after long saccades (both forward and backward through the text) accounted for 20 per cent of all fixations on the written source text. These fixations would all have had a saccade length in excess of 16-character spaces. The data reported by Rayner et al. (2012:95) suggests that, on average, only eight per cent of saccades are in excess of 16-character spaces. This supposition also relies on the assumption that fixations after long saccades, like regressions, generally have a lower mean fixation duration. All 30 participants in this study had a lower mean fixation duration for long forward than short forward, and 21 of the 30 participants had a lower mean fixation duration for long regressions compared to short regressions. In summary, the participants' short forward moving fixations had the longest mean fixation duration and accounted for a lower proportion of fixations in this study compared to those mentioned in the literature, resulting in a lower mean fixation duration for this study.

In addition to a shorter mean fixation duration, the fixations on the source texts also included much higher rates of regression than any reading reported in the eye tracking literature on reading. The literature, much of it generated by Rayner and Pollatsek and detailed in Rayner et al. (2012) (see the literature review section 2.4.1for a full review of eye-movement metrics in reading), generally characterises reading by fixation duration, saccade length (usually expressed in characters), regression rate and sometimes words read per minute.

The researcher was unable to calculate the word per minute for the participants eye-movements on the written source texts but was able to compute and compare all the other metrics. The results of the comparison can be seen in Table 55

*Table 55 Comparison of key eye tracking metrics for reading from the literature and this study*

| Source of eye tracking data | Fixation duration (ms) | Saccade length (in character spaces) | Regression percentage |
|---|---|---|---|
| Mean from Rayner et al.'s table (2012:96) | 231 | 7.8 | 11 |
| Mean for all fixations on written source texts in this study. | 184 | 8.8 | 34 |

The disparity between the means for fixation duration and regression rate from this study and those reported by Rayner et al. (2012) in Table 55 reinforce that participants were engaged in very different reading behaviour than that reported by Rayner. The very high rate of regressions, in conjunction with a short average fixation duration are at odds with the patterns suggested by Rayner et al. The saccade length is the only metric whose mean is similar to those quoted by Rayner et al. (2012).

This high regression rate may be explained, in part, by the data from the retrospective think aloud which suggested that many of the regressions were due to participants deliberately rereading sections of the text to achieve or deepen understanding or memorise parts of the source texts, rather than being the predominately subconscious regressions described in the literature (2.4.1). The researcher suggests that in order to write about the information contained in the source texts, participants needed to achieve a full understanding of the sources, hence the repetitive reading behaviour.

The other factor that may also have contributed to high regression rates is that reading into writing necessitates regular searching of the source texts. Search reading, as defined by Khalifa and Weir (2009 and discussed in 2.3) is when readers search through the text for information on a predetermined topic. The researcher suggests that when readers search they may well ‘overshoot’, going too far through the text necessitating a backtrack to the part they are searching for.

Long regressions remained constant, at a very low level, throughout the task. This suggests that long regressions play a limited role in both careful and selective reading. During careful reading, Rayner et al. (2011) suggest that long regressions result from a reader’s conscious return to an earlier part of the text to resolve a misunderstanding. In selective reading the researcher can only speculate that long regressions may simply be the result of readers moving the focus of their search to an earlier part of the text.

In summary, the broad differences between the eye-movement metrics in the literature for careful reading and the metrics produced by this study were symptomatic of the presence of more long saccades (both backward and forward) and more regressive fixations as participants selectively reread some sections of the text or searched the texts for information.

Having established that careful, systematic reading of the texts was in a minority (4.2.7) that is not to say that careful reading does not have an important role to play. After participants reviewed the task instructions, some participants more thoroughly, some less thoroughly, participants turned their attention to the source texts. During the first 6-7 minutes of the task there was a large peak in both careful global reading and selective global reading (Figure 28), suggesting that at the outset the text was parsed progressively and systematically. When reviewing the data for individual participants (for an example graph see Appendix 22) the data does not suggest that some readers engaged exclusively in careful global reading whilst others engaged in exclusively selective global reading. The picture is subtler and more complex, with all readers engaging in a mixture of reading as they processed the texts, probably utilising different types of reading to meet different needs or goals. Careful global reading played an important part in the opening 6-7 minutes of the task, probably utilised more by the participants with a slightly slower pass through the texts to identify the mains of the written source texts.

Although the study is quite clear about the type of reading that was taking place when fixations were classified as careful reading, the type of reading classified as selective reading also need to be accounted for.

#### 5.2.2.2 The importance of different types of selective reading in reading into writing

As reported in section 4.2.7.1 of the findings, the algorithm developed by the researcher seemed to be able to distinguish a form of careful reading from other types of selective reading. The periods classified by the algorithm as careful global reading are, by virtue of the parameters of the algorithm, periods when the participants' eye-movements progressed forward through the texts in small steps (less than 16 characters) with a limited amount of regressions (18 per cent on average, but not exceeding 25 per cent), not skipping forward or missing out sections of the text. This type of reading broadly accords with the careful global reading described by Khalifa and Weir (2009) and read to learn (Cohen and Upton, 2006, Grabe and Stoller, 2011, Rouet, 2006) as discussed in the literature review (2.3).

However, periods classified as selective reading, when the eye-movements at that time fell outside the parameters set for careful reading, could represent a range of types of reading. The researcher proposes that four other types of reading may be occurring during periods classified as selective reading:

1. Reading for gist / skim reading
2. Intensive reading of selected parts
3. Search reading
4. Scanning for information

Each of these types of reading will now be discussed in turn.

#### 5.2.2.3 Reading for gist

The high rates of selective global reading in the first 6-7 minutes of the task (Figure 28) could represent reading for gist. Khalifa and Weir (2009) suggest that the purpose of skimming, or reading for gist, is to 'build a macro structure of the whole text (the gist) based on careful reading of as little of the text as possible' (2009:57). Khalifa and Weir also suggest that the reader is likely to use 'selective eye-movements which attempt to locate sentences within the text stating the major issues'. Duggan and Payne (2011) proposed that when reading for gist readers may start reading a section or paragraph sequentially, but as the quantity of information about the key point or proposition begins to fall the reader will abandon reading and advance to the next page, section or paragraph and start again. Therefore, it would seem logical that gist reading as described by Khalifa and Weir and Duggan and Payne would include fixations following long forward movements (long forward), interspersed with bouts of short forward moving fixations.

The researcher speculates that if, as suggested in the Duggan and Payne study, the reader begins to realise that the rate of information being gleaned from a section of the text is declining, the distance between fixations may increase and the fixation duration may begin to fall signalling the reader 'disengaging' with the text before jumping to the next area.

During the RTA, the researcher asked some participants about how much of the text they felt they had understood after the first reading of the sources, for example:

> ***Interviewer:*** *Okay. We can see here you're reading through the text in a quite progressive pattern. How much of the information are you able to take?*
>
> ***Interviewee****: Not a lot. I just kind of get what it's about, I suppose.*
>
> *(P34 line 69).*

In these cases, participants reported achieving a broad understanding of the text, rather than a full understanding from their first read through of the source texts, which seems to accord with the concept of reading for gist.

However, when reviewing the data for the individual participants mentioned above, the researcher noted that all these participants made a systematic pass through the texts in the opening few minutes, without skipping any large chunks of the text. In fact, 24 of the 30 participants started with a systematic pass through the entire written source texts (about 600 words) without skipping large chunks. This can be evidenced by the drop in long forward in mins 2-6 in Figure 25. This suggests that most participants in this study engaged in a systematic pass through the written texts. For some it may have been a slower, more careful reading and for others a slightly faster read through. However, as illustrated in Figure 25, short regressions also peaked during minutes 2-6, therefore, for some participants short regressions played a significant role, perhaps prompted by a failure to parse sentences. Participants used this systematic pass through the texts to achieve an outline understanding rather than employing selective reading of the text (as described by Duggan and Payne, 2011). Perhaps this is because participants felt they had time to read the whole text and indeed, knew that at some point they would be required to read the whole text and so were disinclined to skip sections. In the Duggan and Payne study participants had approximately five minutes to read a single text of just over 3000 words, a far more time pressured task, impossible to achieve without skipping some areas. It is unlikely that the participants in Duggan and Payne's study had time to make many regressions, unlike the participants in this study.

The researcher concludes that participants did not engage in the type of reading for gist suggested by Khalifa and Weir and Duggan and Payne which involved skipping sections of the text. Instead they engaged in a slower more methodical parsing of the texts that did not involve skipping any sections of the text. In some cases, it did not seem to afford the participants a thorough understanding of the texts.

Therefore, although selective global reading peaked in the first 6-7 minutes of the task and some participants reported only achieving an outline understanding or a failure to understand, this reading does not appear to represent reading for gist in the ways suggested by Khalifa and Weir and Duggan and Payne, due to the lack of skipping sections.

#### 5.2.2.4 Intensive reading

For some participants a form of intensive reading may have played a part in the opening few minutes of the task. The high peaks of selective global reading may represent another type of reading proposed in the literature. Lorch, Lorch and Klusewitz (1993) reported that when college students described reading undertaken in preparation for examination they reported reading more slowly, rereading frequently, paying attention to detail, memorising information and testing their understanding. There are perhaps, similarities between Lorch, Lorch and Klusewitz's reading before an exam and the intense careful reading to learn described by Cohen and Upton (2006), Grabe and Stoller (2011) and Rouet (2006) in section 2.3. Both would involve frequent rereading as part of a serial progress through the text.

Perhaps participants in this study, knowing that they would shortly have to include the information from the texts in their answer, engaged in the types of processes reported in the previous paragraph; utilising frequent rereading of the text, reflecting on their reading, making connections to their existing knowledge.

If participants did engage in the processes suggested by Lorch et al. (1993) and the proponents of reading to learn, engaging in frequent bouts of rereading, this would have resulted in high levels of regressions. Such behaviour would break the 25 per cent limit on regressions for careful global reading and result in a classification of selective global reading. Taking time to reread frequently and to reflect may account for the slow progress through the text (generally 3-4 minutes for 600 words) despite the short mean fixation durations for selective global reading.

The researcher notes that, overall, periods classified as selective global reading also included 25 per cent of long forward moving fixations. As the participants did not skip any sections of the text (as discussed above in reading for gist), the presence of these longer forward jumps also needs to be accounted for. Rayner (1998:376) suggests that after a regression, readers are more likely to make a longer saccade forward to 'place the eyes ahead of where they were prior to making the regression'. The researcher notes that 7,176 of the 12,945 long forward s remained in the same sentence as the previous fixation. This suggests that the majority of long forward s were more likely to relate to skipping a word after a regression than skipping a chunk of the text in the case of reading for gist.

Therefore, the researcher suggests that reading classified as selective global during the first six or seven minutes of the task may represent reading to learn or the type of reading suggested by Lorch et al. (1993). This type of reading might represent a more intense form of careful reading; ultra-careful reading or intensive reading perhaps, rather than a form of selective reading.

#### 5.2.2.5 Search reading

Khalifa and Weir (2009) use the term *search reading* to describe a reader's purposeful hunt for information on a predetermined topic where the reader looks for semantic links to the topic or theme (2.3). During this type of reading the reader remains alert to the meaning conveyed by the text because they are required to infer a link to the topic, they are not limiting their search to a specific word or fact. Once a semantic link has been established, the reader will then engage in careful reading in the sentence or sentences around the semantic link.

During the RTA there where 30 references to search reading, extracting information or participants targeting their reading for example: P34, line 67, 'I couldn't remember it but I knew there were some examples in there' or P52, line 58 'it's just a matter of just picking out the right information from the articles' or P41 line 91 'I then go back into the article and then find what they are saying the problem is'. This type of behaviour seems to accord with Khalifa and Weir's search reading because the participants' reading was driven by a desire to locate or identify certain information that they knew was in the text having already read it.

These comments were reported at times when participants eye-movements regressed back or jumped forward to an area of text which they then examined in detail, going back and forth within a sentence often with a high rate of regressions. Such behaviour resulted in fixations being labelled as selective global or selective local reading. The high rate of regressions could suggest that the reader was having difficulty parsing the sentence however, the lower average fixation duration (98ms) of selective local reading compared to careful local reading (103ms) perhaps suggests that participants were more likely to be engaged in search reading because difficultly parsing a sentence might be expected to produce longer fixation durations as the reader slowed down whilst they attempted to resolve any comprehension problems.

Therefore, it seems that some periods classified as selective global reading could represent search reading, as the reader hunts for information on a predetermined topic and selective local reading could represent periods after the reader has found information and examines that area of the text more closely.

The researcher suggests that search reading (as defined by Khalifa and Weir 2009:57) is likely to generate similar fixation patterns to skim or gist reading however, skim or gist reading would logically seem to be very sequential in nature, whereas perhaps in search reading the reader may jump to an area where they think the information is most likely to be located. This may not necessarily be near the beginning of the text, and if the reader is wrong and does not locate the information they are looking for, they may well jump backward to an earlier point in the text to search.

#### 5.2.2.6 Scanning for information

Khalifa and Weir (2009: 59) describe scanning as a targeted search for a specific word or piece of information. They suggest that the reader will not necessarily progress through the text in a linear fashion. Instead the reader will engage in a hunt for a particular word, ignoring the syntax of the text; not bothering to parse sentences for meaning.

Grabe and Stoller (2011:7) also describe a form of reading which they, rather confusingly, call *search* reading. In Grabe and Stoller's 'reading to search for simple information' the reader is also moving rapidly through the text to find a specific piece of information, although it seems that Grabe and Stoller may be referring to both scanning for a particular word or searching for a semantic match. Khalifa and Weir's (2009) *scanning* refers to reading in which the reader *does not* engage with the meaning of the text but rather performs a word or text matching exercise. To be clear, this research adopts Khalifa and Weir's term *scanning* to describe this type of word matching behaviour and uses the term *search reading* to apply to Khalifa and Weir's purposeful hunt for information on a predetermined topic where the reader *does* engage with the meaning of the text (as discussed in section 5.2.2.5).

There were some references in the RTA to this kind of reading, for example P41 line 114 'I knew where certain words I wanted to use were' or P42 line 57 'I scanned until I found it' although there were only six references in total. Once again, the researcher was unable to locate any research which detailed the eye-movement patterns involved in scanning or lexical search but logically it would seem to involve high numbers of consecutive long saccades (possibly forward and backward) with few or no short saccades until the 'match' is found. Such patterns of eye-movements would be categorised as selective global reading by the algorithm used in this study. Accurate detection of such reading behaviour would seem to depend on an absence of short forward moving fixations and the presence of consecutive long saccades, either backward or forward. Periods classified as selective reading could represent skimming.

#### 5.2.2.7 Summary of reading types

In summary, the eye tracking data in conjunction with the RTA data would suggest that participants in this study did not engage in selective reading for gist in the way suggested by Khalifa and Weir and Duggan and Payne. Instead, between minutes two – seven of the task, the participants tended to read through the texts without skipping any sections. Although the first pass through the text generally lasted about four minutes, which suggests quite a slow rate of reading, the average fixation durations during the first six or seven minutes of the task were quite short compared to those in the literature for careful reading. Some participants reported achieving a broad understanding of the texts, more in line with reading for gist (5.2.2.3). For other participants, the first read through may have afforded them a more comprehensive understanding of the texts, more in line with careful reading. In summary, this suggests that during the minutes two to seven of the task participants engaged in a first pass through the written source texts. Some participants, those using longer saccades and therefore likely to have more have periods classified as selective global reading, achieved a broad understanding of the text. Others, those with shorter saccades and therefore more likely to have periods classified as careful reading, achieved a slightly more comprehensive understanding. However, the reality is that whilst some participants tended towards selective global reading and others tended towards careful global reading, overall there was a mixture with participants falling somewhere on that spectrum.

Most of the participants then engaged in a second read through of the texts starting from minute 5- seven onwards. This pass tended to be less sequential; involving intense rereading of some sections whilst other sections were skipped. This may reflect intensive reading discussed in section 5.2.2.4. After this, as participants began to write their answers they often returned to the text, searching for information, checking facts or extracting information perhaps using search reading 5.2.2.5 and scanning 5.2.2.6

Table 56 below summarises the types of reading discussed above, identifies the fixation patterns that might distinguish them and how those fixation patterns would have been categorised by the researcher's algorithm for careful reading.

The researcher concludes that the algorithm has enabled a broad division between some forms of progressive careful reading and other forms of selective reading but suggests that this technique could be developed and refined to offer greater discrimination between types of reading.

*Table 56 Classification of different types of reading and associated eye movement charactericstics*

| Type of Reading | Purpose | Proposed by | Characteristics of eye-movements | Classified as | Crosses sentence boundaries |
|---|---|---|---|---|---|
| **Careful reading** | Understand main points in preparation for writing | Khalifa and Weir (2009) | Mainly short forward with regressions up to 25%, resulting in a sequential pass through the text, with no skipping of sections | Careful Global | Yes, but must be sequential |
| | Memorise information and test understanding | Lorch, Lorch and Klusewitz (1993) | Still largely sequential but will involve frequent rereading potentially resulting in very high rates of regression. Sections that are of less interest or relevance may be skipped. | Selective Global / Selective Global / Careful local | Yes, but must be sequential |
| **Gist / Skim reading** | Investigate text to identify general theme / main point | Duggan and Payne (2011) / Khalifa and Weir (2009) | Very sequential pass through the text using short forward with very low rates of regression until a point / theme is identified. | Data suggests it was not present in this study as participants failed to skip sections of text in the initial minutes of the task. | Yes, but must be sequential until jump |
| | Move to next section | | Long forward saccades to next section | | Yes, but must move the reader forward to next paragraph / section |
| **Search reading** | Look for semantic match | Khalifa and Weir (2009) | Sequential reading using short forward with quite low rates of regression (as not attempting to resolve misunderstandings) until a semantic link is found | Selective Global | Yes, not necessarily sequential |
| | Gather information after a match has been found | | Careful reading around point of match. High rates of regression may occur | Selective local / Careful local | Confined to a single sentence |
| **Scanning for specific word / detail** | Find an exact word | Grabe and Stoller (2011) Khalifa and Weir (2009) | Very fast pass through the text using longer saccades not necessarily linear or sequential. | Selective Global | Yes, not necessarily sequential |
| | Extract / check fact | | Careful reading at point of match | Careful local | No |

So far, the discussion of the interaction with the source texts has covered the broad characteristics of the fixations and the different types of reading that might be represented by fixation patterns. The discussion now moves on to the way in which attention was directed towards the most relevant sentences.

#### 5.2.2.8 Participants fixated more on the most relevant sentences

The analysis of sentences according to their relevance to the reading into writing task showed that, per word, highly relevant sentences received the most attention, relevant sentences received slightly less attention and less relevant sentences received the least attention. This implies that most of the participants were able to identify the most important material in the source texts to some extent. Pearson's correlation coefficient was also calculated for the relevance score per sentence and the average fixations per sentence when adjusted for sentence length. The results suggested that there was a moderate correlation (r=.531) between sentence relevance and the number and duration of fixations. However, the full picture is likely to be more complex.

The literature on eye-movements in reading suggests that word recognition issues such as word frequency and predictability (Inhoff and Rayner,1986, Reichle, Rayner and Pollatsek, 2003, Rayner and Duffy, 1986), and sentence comprehension issues such as ambiguity of meaning (Clifton, Staub and Rayner, 2007, Boland, 2004) have a role to play in influencing fixations durations and saccade lengths. During the retrospective think aloud data four themes emerged when participants were asked to account for periods when they engaged in repeated rereading of sentences (4.2.8.3).

Two of the explanations suggested that additional attention was allocated to more relevant sentences as the result of a deliberate strategy, whilst the other two explanations suggested necessity rather than a strategic decision.

The two meta-cognitive strategies were ‘rereading to deepen understanding’ and ‘rereading to extract information’. The two necessities were ‘rereading due to a failure of understanding’ and ‘rereading due to a lack of engagement with the reading process’. This refers to times when the participants’ eyes moved over the text but participants’ reported realising that they had no recollection of what they had just ‘read’ and therefore needed to reread.

With the conscious strategies, it is logical that participants identified the sentences as important to the task and therefore targeted them, ensuring understanding and extracting information. With the ‘unconscious’ strategies it is more difficult to explain. In the case of a failure to understand participants may spot key words or phrases that they recognise as being significant, despite not being able to understand the sentence and therefore invest effort in reparsing the sentence. In the case of lack of engagement with ‘reading’, logic would suggest that this is arbitrary and not related to sentence relevance. Therefore, three out of the four reasons that participants gave for needing or wishing to reread sentences were motivated by the perceived relevance of the sentence.

The researcher also noted that several participants reported finding one of the solutions provided in the source text (the ‘management standards approach’) difficult to understand. Appendix 23 details the text of the two articles contained in the written source texts, marked to show the sentence reference and relevance rating. Sentence four on page two reads

*The Health and Safety Executive (HSE) has designed the Management Standards approach to help employers manage the causes of work-related stress*

and was rated as highly relevant. P3S5 is also rated as highly relevant and reads

*The Resolve Staff Support Service offers another organisationally-focused solution to help reduce sickness absence resulting from work-related stress.*

If we refer back to Figure 23 we can see that P2S4 received the highest amount of attention per word (220 fixations per word) and P3S5 received less (just under 190 fixations per word). The researcher suggests that in terms of semantics and lexis P3S5 is slightly more challenging and complex than P2S4. However, participants reported that the management standards approach mentioned in P2S4 was abstract and difficult to understand (P51 line 65, P54 line 60, P47 line 33). Therefore, the researcher suggests that the abstractness or concreteness of ideas contained in the texts, regardless of syntactic complexity or ambiguity, may also have a role to play in determining the amount of attention sentences receive. The role of abstractness could prove a fruitful area for further research. If abstractness proves to increase comprehension difficulty this is a factor that will need to be taken into consideration by teachers and test developers.

This concludes the discussion on the interaction with the written source texts. The discussion now moves on to the way participants interacted with their own work.

### 5.2.3 Participants' interaction with their own work

The participants devoted almost half of their fixations (47 per cent) to their own work and, because fixations on their own work tended to be longer than fixations on any other AOI, 56 per cent of fixation duration.

Much of the literature on writing recognises the role of reading the emerging text to facilitate evaluation (Alamargot et al., 2010, Hayes 1996, Hayes and Flower 1980, Wengelin et al, 2010) revision (Alamargot et al., 2010, Hayes 1996, Hayes and Flower 1980, Wengelin et al, 2010) and the production of what comes next (Alamargot et al., 2010, Wengelin et al, 2010).

The large proportion of time that participants devoted to focusing on their own work seems to confirm that the emerging text performs a vital role in the reading into writing process. The mean fixation duration (274ms) was significantly longer than the mean fixation duration on the written source texts (184ms) and the RTA data suggested long fixation times were due to participants engaging in evaluation, revision, micro planning or translating as they fixated on their work (see sections 4.2.8.5 and 4.2.8.6).

After evaluating their work, the RTA data suggested that when participants made higher level revisions to their work they were frequently focusing on coherence and cohesion and trying to improve expression. Lower level revisions included spelling, grammar and changes to vocabulary. Participants also decided to add more information after reading to evaluate and finding their work lacking. Periods when participants made frequent saccades between their own work and the source texts were usually to ensure that they were not plagiarising the sources. Again, a type of evaluation of their work.

Several studies have already used eye tracking to investigate the role of reading the writer's emerging work during a general writing task (Alamargot et al., 2010, Wengelin et al, 2010), but none have investigated the role of the emerging work in a reading into writing task. This study was able to analyse number and fixation duration of participants' fixations on their own work, but unable to analyse the participants' fixations on different parts of their own text due to technical difficulties. The researcher believes that with work, these technical issues could be overcome. This is certainly an area where eye tracking could be used fruitfully, perhaps in conjunction with keystroke logging and RTA to further investigate the role of reading the emerging work. The amount of time that participants spent fixating on their work suggests that any model of reading into writing would have to take account of the significant role played by the emerging work.

### 5.2.4 Summary of the discussion for RQ1

The researcher draws a number of conclusions about academic reading into writing. Firstly, although reading the task instructions only accounts for 6 per cent of fixations the researcher suggests that the task representation process has a key role to play in shaping not only the answer but also the way the reader engages with the source materials and their evaluation of their own work.

Secondly, when engaging with the source texts, reading during a reading into writing task differs significantly from the careful reading reported thus far in the eye-movement literature. Fixations on the written source texts were characterised by far higher rates of regressive fixations than reported in the literature for careful reading due to the participants deliberate rereading of sections of the text to deepen their understanding, overcome a lack of understanding or search for information. Although careful reading only accounted for about 30 per cent of fixations on the written source texts it still played an important role in helping participants develop a clear understanding of the source materials, especially during the first 6-7 minutes of the task. The researcher suggests that in order to produce their answers (own work) participants needed to develop a much deeper understanding of the most relevant parts of the source texts. The need to focus on certain parts of the term perhaps explains the high levels of selective reading which accounted for 70 per cent of fixations. The selective reading was in the form of reading to achieve an overview, intensive rereading, search reading and scanning for information.

Thirdly, when reading the written source texts participants fixated more on the most relevant sentences, in part as the result of conscious / deliberate strategies, in part as a result of a failure to understand. However, relevance was not the only factor exerting an influence. The abstractness of ideas, word frequency / predictability and syntactic complexity may all have a role to play in determining the amount of attention devoted to each sentence.

Finally, as with writing more generally, the writers' emerging work plays a critical role in the process, facilitating evaluation, revision and generation.

Having discussed the findings in relation to RQ1, we move on to the implications of the results for RQ2.

## 5.3 Discussion of the findings for RQ2

RQ2 was 'what are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?' RQ2 wanted to investigate whether the way that students tackled a reading into writing task changed as they became more experienced. Therefore, the researcher compared the data generated by the Y1 participants with the data generated by the Y3+ participants.

In answer to RQ1 the researcher has highlighted the key findings that characterised the reading behaviour of participants as they engaged in the reading into writing task (5.2.4).

Any differences between the two groups in relation to these key findings from RQ1 might indicate whether greater experience of completing academic reading to write tasks influenced the way the participants read during the task. Therefore, the discussion of the findings for RQ2 will follow that same pattern as RQ1; interaction with the task instructions, interactions with the source text (including types of reading and attention on the most relevant sentences) and interactions with participants' own work.

### 5.3.1 Y1 / Y3+ interaction with the task instructions

The Y3+ group spent slightly more time attending to the task instructions (6,637 fixations / 22 mins ten secs compared to 5,718 fixations / 18 mins 11 secs) however, as a percentage of the time spent fixating on the whole task the percentages were very similar (6%). This could suggest that there was little difference between Y1 and Y3+ interaction with the task instructions however, the RTA data revealed other differences between the groups. These differences related to the way in which participants tended to engage with task instructions rather than for how long.

During the RTA the researcher noted that Y3+ participants often started making notes when they read the task instructions. These notes were subsequently used as a macro-plan for their answer. In total, eleven Y3+ participants made it clear that they were using the task as a basis for their macro structure. This compared to just three Y1 students who made similar comments. However, five Y1 participants reported either using something else as their macro structure (P26: adapting the structure of a previous essay or P29, P31, P28, P35: using their notes of the articles as a macro structure). No Y3+ participants reported using anything other than the task as a basis for their macro structure.

The preference to use the task instructions to generate a macro-structure suggests a more sophisticated task representation on the part of those participants. Using the task instructions to shape the macro-plan ensures that the key demands of the task are central to the final essay. It also ensures that the key points are covered in the same order as the task instructions making it easier for the reader (the fictional lecturer referred to in the task instructions) to understand the structure of the essay and identify the content relating to each point.

In conclusion the researcher suggests that the more experienced Y3+ participants developed a more sophisticated task representation as a result of their interaction with the task instructions than the Y1 participants. The differences in task representation had a knock-on effect on the way the two groups of participants engaged with the source texts, which will now be described in more detail.

### 5.3.2 Y1 / Y3+ interaction with the source texts

The initial broad eye tracking metrics suggested that there were similarities in the ways both groups had engaged with the texts. Total fixation duration in particular was, almost identical (Y1: two hours, one minute, 37 seconds, Y3+: Y1: two hours, one minute, 57 seconds), whilst number of fixations differed by about 1300 fixations (Y1:40,362, Y3+: 39,022). The Y1 group had a slightly shorter mean fixation duration (181 milliseconds) compared to the Y3+ group (188 milliseconds). The standard deviation of fixation duration was also rather similar (Y1: 96, Y3+:103). This suggests that both groups included a very similar range of fixation durations.

However, as the data was analysed, key differences emerged.

### 5.3.3 Y1 / Y3+ broad differences between reading on this task and reading in the literature

Some of the broad differences that marked reading in this study out from reading in the literature did not apply equally to the Y1 and Y3+ data. To remind the reader, the broad differences were a lower mean fixation duration, higher rates of regression and increased rates of long saccades.

The Y1 and Y3+ group had very similar levels of regressive fixations on the source texts. However, they Y3+ group made more long saccades. The two groups had similar levels of long regressive saccades (4 per cent) but fixations after a long forward saccade accounted for 18 per cent of the Y3+ fixations compared to 15 per cent of the Y1 group.

It is tempting to suggest that the Y3+ group were simply better readers, but the other metrics do not substantiate this. If the Y3+ group were better readers in general you could expect to see a shorter mean fixation duration and lower levels of regressive fixations than their count parts, neither of which is the case. Therefore, the differences in saccade length are more likely to be reflective of differences in the types of reading engaged in, as the next section discusses.

#### 5.3.3.1 Y1 / Y3+ types of reading

As suggested above by the differences in saccade lengths, in general terms the Y3+ group engaged in more selective reading (74 per cent) compared to the Y1 group (69 per cent). Inevitably, this meant that the Y1 group engaged in more careful reading (31 per cent) than the Y3+ group (26 per cent). It suggests that the Y3+ group adopted a more selective approach to reading.

#### 5.3.3.2 Y1 / Y3+ attention on the most relevant sentences

The broad fixation metrics were very similar for both groups, but small differences between the groups began to emerge as attention across sentences when grouped according to sentence relevance was considered.

As reported in Chapter 4.2.4, it proved difficult to draw conclusions regarding the amount of attention paid to individual sentences of the written source texts, therefore the sentences were grouped according to their relevance to the reading into writing task and analysed again. The results showed that both groups devoted their attention according to relevance with the Y3+ group allocating fractionally more of their attention towards most relevant sentences. Analysis of the correlation between attention per sentence (after adjusting for sentence length) and sentence relevance (as scored by the raters) showed that whilst both groups were able to identify the areas of the text that contained the information required to successfully answer the question set by the task, the Y3+ group data represented a slightly stronger correlation than the Y1 group (Y3+: $r$=.503, Y1: $r$=.484). Therefore, the Y3+ group were slightly better at focusing their attention on the most relevant areas of the text.

#### 5.3.3.3 Y1/Y3+ reading during the first 6 / 7 minutes of the task

The differences in saccade length appear to reflect the higher rates of selective reading by the Y3+ group and higher rates of careful reading by the Y1 group. It is useful to consider whether these differences were as the result of subtle differences over the course of the whole task, or whether they were concentrated in certain periods of the task.

If we review Figure 33, it becomes clear that the majority of the difference between the two groups occurred in the first few minutes of the task. Two factors combine to suggest that the differences reported above are as a result of the Y3+ group engaging in more reading for gist. Firstly, the results of the RTA data were quite marked. The Y3+ group made 14 comments regarding reading for gist, compared to just four comments from the Y1 group. Secondly, the large spike in Y3+ selective global reading occurred in minutes two to four of the task and this very early stage of the task is when logic would suggest reading for gist takes place.

Although in the discussion of RQ1 the researcher concluded that participants did not skip sections of the text, as the literature suggests might be expected when reading for gist, the researcher suggested that participants only achieved an outline understanding of the text from the first sequential pass through the source texts. The high number of references to reading for gist by the Y3+ group in combination with the high number of Y3+ fixations classified as selective global reading in minutes two to four of the task suggests that the Y3+ group used a quick sequential pass through the texts in minutes two to four to achieve an outline understanding of the texts. Perhaps reading for gist but without skipping any sections. As Figure 33 illustrates, after minute four the number of fixations classified as selective global reading fell rapidly, so that by minute six levels of selective global reading were very similar for both groups and continued at similar levels for the remainder of the task.

During the same period, minutes two to four, almost twice as many of the of the Y1 group’s fixations were classified as careful local reading compared to the Y3+ group. Once again by minute six their levels were approximately the same and both groups’ rates of careful local reading followed a similar pattern for the rest of the task. Therefore, the Y1 group tended to read sentences individually, using short saccades of less than 16-character spaces.

This suggests that they tackled the texts in a more piecemeal approach, perhaps making less connections between main ideas and progressing through the texts more slowly. The Y3+ group, on the other hand, used a more holistic approach to create an overview with less understanding of the detail, but more understanding of how the ideas related to each other.

### 5.3.4 Y1 / Y3+ interaction with their own work

When the figures are broken down to consider how fixations were divided across the different AOIs in section 4.3.1 we can see that the Y3+ group made many more fixations on their own work than Y1 group (Y1=44% of fixations / Y3+=50% of fixations) . As noted in section 5.2.2.8 reading the emerging written texts forms an import role in the writing process and the additional time spend fixating on their own emerging text suggests that the more experienced participants reread their own work more often.

The RTA data suggested that the additional time on focusing on the participants own work was in due to the Y3+ group engaging in more evaluation and revision (the Y3+ group made a total of 60 comments relating to evaluating and revising compared to just 35 comments from the Y1 group). Comments relating to the micro planning and translation phases, which would seem to be related to the production of what comes next were almost identical (Y1: ten comments, Y3+: nine comments). Therefore, the researcher concludes that the majority of additional time the Y3+ group spent looking at their own texts related to evaluating and revising it. From the outset the Y3+ group seem to have had a more complete task representation, not only in terms of what the product would look like but also in terms of the process that they would go through in order to achieve it. The researcher speculates that because the Y3+ participants had a more developed concept of what was required for the finished product, they evaluated their writing more to see if it met the demands of the task. Therefore, the researcher suggests that good task representation prompted the more experienced Y3+ participants to evaluate their work more and subsequently make more high-level revisions to improve their answers.

### 5.3.5 Summary of the discussion of RQ2

There appear to be four main differences between the way the Y1 and Y3+ group tackled the task. The first difference emerged as participants read the task. Although the Y1 and Y3+ group both devoted a similar proportion of their fixations to the task instructions (6%) according to the RTA data, the Y3+ group appeared more likely to use the task instructions to guide the macro-structure of their answer. This difference in initial task representation led to differences in the way the groups tended to engage with the texts.

Secondly, when reading the written source texts, the Y3+ group engaged in 5 per cent more selective reading than the Y1 group. The majority of this difference was accounted for in the first 6-7 minutes of the task when the Y3+ group appeared to read more quickly and progressively through the entire source texts, perhaps creating a better overall understanding of how the key themes or main ideas in the texts related to each other. The RTA data certainly suggests that the Y3+ group were more aware of reading for gist as a strategy and felt they had engaged in reading for gist. Conversely the Y1 group seem to have read the texts more carefully, one sentence at a time. Perhaps rereading each sentence before progressing on to the next (suggested by the classification of more local reading than global reading). Few of the Y1 students referred to reading for gist during the RTA.

The third difference related to identifying the most relevant information in the source texts. The Y3+ group spent slightly more of their time on the more relevant sentences as indicated by the slightly higher correlation (Y3+: *r*=.503*,* Y1: *r*=.484) between time dedicated to each sentence (when adjusted for length) and sentence relevance score. This difference may reflect the difference in levels of selective reading. If selective reading is used to direct attention to the most relevant areas, less selective reading would result in less attention on the most relevant sentences.

The final difference related to the greater amount of time Y3+ students spent rereading their own work (Y3+: 57,165 fixations / 50% of fixations compared to Y1: 43,916 / 44% of fixations). The RTA data suggested that this time reading their own work was devoted to evaluating and revising their work. One again this tendency to engage in more evaluation of their own work may be linked to Y3+ participants having a better task representation. If the Y3+ had a clearer understanding on the demands of the task, this suggests they would also have a clearer vision of what the finished product should include, how it should be organised and so on. This may have led to them spending more time making high level revisions to try and achieve this.

## 5.4 Discussion of the findings for RQ3

Perhaps inevitably, the Y3+ group scored better on the whole than the Y1 group (Y1:7.40, Y3+:8.17). However, there were two Y1 students who scored at the highest level and two Y3+ students that fell in the lowest five scoring participants. This suggests that although experience of performing this type of academic task has a role to play, other factors also have a role to play. In order to establish which, if any, of the differences in the way the Y1 and Y3+ participants tackled the task were the result of greater experience of tackling academic reading into writing tasks, the same metrics were compared for the five highest and five lowest scoring participants. If the difference in a particular metric decreased when the highest and lowest scorers were compared this suggests that the difference related to experience of performing this type of academic reading into writing task. Conversely, if a difference observed between Y1 and Y3+ participants increased when the five highest and five lowest scoring participants were compared this suggests that that metric is more strongly related to reading into writing ability than experience.

As a result, the metrics relating to the four key differences between the Y1 and Y3+ groups were analysed again for the five highest and five lowest scoring participants. Once again, the key findings from RQ3 are discussed in terms of interaction with the task instructions, interaction with the source texts and finally, interaction with the participants own work.

### 5.4.1 High / low scorers' interaction with the task instructions

The results of RQ2 suggested that whilst the Y1 and Y3+ group fixated for a similar percent of time on the task instructions (both Y1 & Y3+ = 6%), according to the RTA data, the Y3+ group appeared more likely to use the task instructions to guide the macro-structure of their answer. Both the high and low scorers fixated slightly less on the task instructions (both high and low scorers = 5%) than the Y1 and Y3+ group. With regards to the RTA data, two of the low scoring participants suggested that they had used the task instructions to guide their macro plan compared to three of the high scoring participants.

As there is no difference between the fixation percentages for the Y1 and Y3+ groups, nor any difference in percentage between the high and low scoring groups these results are rather inconclusive. The RTA data for the high and low scoring groups suggested that the high scorers, in common with the Y3+ participants, were more likely to use the task instructions to guide the macro-structure of their answer. Because of the low numbers of participants in the high and low scoring groups caution needs to be exercised but this may suggest that academic reading into writing ability as well as experience led participants to focus on the task instructions and use them to guide their macro planning.

### 5.4.2 High / low scorers' interaction with the source texts

When reading the written source texts, the Y3+ group engaged in 5 per cent more selective reading than the Y1 group. The overwhelming majority of this difference emerged in the first 6-7 minutes of the task and seemed to relate to the ways in which participants engaged with the texts in the first progressive pass through the texts.

The high scoring group also engaged in more selective reading than the low scoring group, 7 per cent more selective reading. Unlike the Y1 / Y3 groups, where the differences were confined to careful local and selective global reading, the differences between the high and low scoring groups extended over all four different types of reading. However, in common with the Y1 and Y3, the largest differences could be seen in the first 6-7 minutes of the task (review Figure 33 to Figure 36).

The RTA data relating to different types of reading data was, on the whole, limited to one or comments per type of reading. It was therefore difficult to draw any clear conclusions from this. However, the increase in the differences between careful and selective reading between the High and Low scoring groups suggests that ability as well as experience is related to higher levels of selective reading.

### 5.4.3 High / low scorers' focus on the most relevant sentences.

The third finding from RQ2 related to identifying the most relevant information in the source texts. The Y3+ group spent slightly more of their time on the more relevant sentences as indicated by the slightly higher correlation (Y3+: *r*=.503*,* Y1: *r*=.484) between time dedicated to each sentence (when adjusted for length) and sentence relevance score. In contrast to the small difference in correlation between the Y1 and Y3+ groups, the differences between the Low and High scoring groups was marked. For the high scoring group, the correlation was moderate (r=.411) but for the low scoring group there was only a small correlation (r=.165). Because the numbers of participants in each group was very small (five), individual differences are likely to have had a greater influence than for the Y1 and Y3+ groups or for the entire group. Nevertheless, the difference between the low and high scoring groups is marked.

### 5.4.4 High / Low scorers' interaction with their own work

Y3+ participants spent more time rereading their own work and the RTA data suggested that this time was devoted to evaluating and revising their work, rather than as a prompt for what to write next. Although the High scoring participants made more fixations than the Low scoring participants (High: 17,470 fixations, Low: 12,911), both groups allocated 47 per cent of their fixations to their own work (Table 49). Therefore, this suggests that devoting time to reading their own work is linked more strongly to experience of completing academic reading into writing tasks than to ability.

### 5.4.5 Summary of the discussion of RQ3

This section will compare the results for the low and high scoring participants to the results for the Y1 and Y3+ participants in terms of the four key findings from RQ2.

The first difference emerged as participants read the task. Whilst both the Y1 and Y3+ groups devoted 6% of their fixations to the task instructions differences emerged in the way the groups used the task instructions. According to the RTA data, the Y3+ group appeared more likely to use the task instructions to guide the macro-structure of their answer than the Y1 group. This pattern appeared to remain consistent with the high and low scoring groups. This suggests that using the task instructions to guide the macro structure of the answer is linked to both experience of completing academic tasks and academic reading into writing ability.

The second difference related to selective reading of the written source texts. The Y3+ group engaged in 5 per cent more selective reading than the Y1 group, most of the difference emerged in the first 6-7 minutes of the task. This difference increased to 7 per cent between the Low and High Scoring groups. Whereas, for the Y1/ Y3 groups the differences had been confined to careful local and selective global reading, the differences spanned all four different types of reading for the Low and High Scoring groups. These differences were also largely confined to the first 6-7 minutes but some differences later in the task were also evident. This suggests that both experience and academic reading into writing ability exert an influence on selective reading, meaning that better performing, more experienced students utilise selective forms of reading such as reading for gist, searching, scanning and intensive study of some sections of the text. Low scoring participants were less likely to read selectively, relying more on careful reading of the texts.

The third difference related to time spent fixating on the most relevant sentences. Whilst the High Scoring participants did not dedicate more of their time to reading the more important sentences in comparison to the more experienced Y3+ students (slightly less in fact), the Low scoring participants dedicated only a small amount of time more to the more relevant sentences. Therefore, the difference between the Low and High scoring groups was more marked than the differences between the Y1 and Y3+ groups. This seems to suggest that a failure to be able to identify the more relevant content has an important impact on scoring ability. That is not to say that participants that identified the most relevant sentences always scored highly, they did not, and it is likely that other factors also have a role to play in ensuring a high scoring answer.

The final difference related to the greater amount of time Y3+ participants devoted to reading their own work. The RTA data seemed to suggest it related to evaluating and revising their work rather than functioning as a prompt for what to write next. Both the low and high scoring participants devoted less time to reading their own emerging work. This suggests that although the more experienced participants spent more time fixating on their own work, there is not a direct relationship between more time spent reading your own work and scoring highly.

## 5.5 Summary of discussion

To help draw together the main points discussed in this chapter this section will summarise the points discussed in relation to the key findings.

### 5.5.1 The role of reading the task instructions in task representation

Participants began by reading the task instructions. The task instructions were attended to for only half as long per word as the source texts. That said, the fixations on the task instructions were generally longer suggesting that the task instructions were read more slowly and carefully than the written source texts.

The more experienced participants often made notes of the key points from the task and used these as an outline for their writing, developing their macro plan from the points in the task. The RTA comments also suggested that some participants used the task instructions to set themselves reading goals, or to establish what information they were looking for when reading. This was apparent among the more experienced Y3+ participants and to some extent among the high scorers too. This suggests that some participants were able to establish a more comprehensive and sophisticated task representation at the outset which influenced the way they tackled the task. The researcher suggests that this would certainly be a fruitful area for further research.

### 5.5.2 Reading of the written source texts

The amount of reading of the written sources peaked at very high levels during the opening 6-7 minutes of the task. This reading often took the form of a very sequential pass through the texts, not omitting or skipping any sections of the text, as might be expected during reading for gist. However, participants often reported only gaining a broad outline understanding from the first pass through the text. This was often followed by a slower more intense pass through the texts where rereading of sentences / sections was common.

The most notable finding was that not all sentences were afforded the same amount of attention. There was a moderate correlation between the relevance rating of the sentences and the amount of attention they received. Because some sentences were afforded more attention, this suggests a pattern that deviated from a uniform, sequential progression through the text, reading each sentence in turn. Indeed, over 70 per cent of the reading of the written source texts did not appear to reflect the progressive careful reading reported in the literature.

Two eye-movement metrics seemed to differentiate the reading in this study from that reported in the literature. These were the mean fixation duration and the percentage of regressions. These two metrics are inter linked with each other and with the key findings regarding reading of the source texts; that selective reading dominated careful reading and that the most relevant sentences were afforded more attention.

The relationship between regression rate and mean fixation duration is quite straight forward. In this study, fixations that occurred after a regression reported a lower mean fixation duration than fixations after a forward saccade. This study reported higher levels of regressions than careful reading in the literature, therefore a lower mean fixation duration is to be expected.

The cause of the high rates of regression were two-fold. Firstly, participants reported deliberately rereading sentences or sections of text in order to understand, or further their understanding of the ideas reported in the text. The researcher proposes that this is type of reading is 'intensive' reading. Secondly, the participants engaged in forms of reading, such as search reading or scanning which generate high levels of regressions.

The eye tracking data suggests that the type of careful reading described in the literature accounted for less than 30 per cent of the reading of the written source texts. Over 70 per cent of the reading of the written source texts was therefore selective. The forms of selective reading suggested by interpreting the eye tracking data, in conjunction with RTA data, were a form of reading for gist, search reading, scanning for information and intensive reading.

As participants began to write their answers the levels of reading on the written source texts dropped substantially (between minutes seven and 10). After this time all forms of reading the written sources declined gradually as the task progressed.

### 5.5.3 Reading the emerging text (own work).

That is not to say that reading stopped as the task progressed. Rather, instead of reading the source texts participants began to spend more time reading their own emerging texts. Overall, almost half of participants fixations were on their own emerging text. The RTA data suggested that participants read their own texts for three reasons. The reasons were to help generate the next part of the text, to evaluate their work and to revise their work. Revising and evaluating were cited as reasons much more frequently than generating.

### 5.5.4 The role of experience in reading into writing

When considering the key findings of the study for RQ1 four differences emerged between the Y1 and Y3+ groups. Firstly, although both the Y1 and Y3+ participants devoted about 6% of their time to reading the task instructions the RTA data suggested differences in the way the groups used the task instructions. The RTA data suggested that the Y3+ group used the task more to help them form a macro plan before they started reading and used the plan to guide their reading. More experienced participants read with a purpose or goal for reading in mind.

Secondly, perhaps as a result of better task representation, the Y3+ group focused slightly more of their attention on the most relevant sentence. The Y1 group also focused more attention on the more relevant sentences but to a lesser extent, although the difference between the groups was quite small.

Thirdly, the Y3+ group engaged in more selective reading and less careful reading than the Y1 group. The difference equated to approximately 5 per cent of total fixation time on the written sources. The differences between the groups were most notable during the first 6-7 minutes of the task when the Y3+ group seemed to read more quickly to achieve an outline understanding of the text. By comparison the Y1 group seemed to engage in a slower, more careful first pass through the written source texts.

The fourth and last difference was that the Y3+ group devoted a lot more time to their emerging texts (own work) as they worked. The RTA data suggested that this was largely due to evaluating and revising their own work.

### 5.5.5 The differences exhibited between high and low scoring participants

When the data for the five lowest and five highest scoring participants is compared to the data for the Y1 and Y3+ groups one of the differences remains consistent, two of the differences increase, and one is confounded.

The Y1 and Y3+ group devoted the same percentage of their (6%) of time to the task instructions, the same can be said for the high and low scorers, albeit slightly less time (5%). The RTA data for the Y3+ group suggests that these participants were more likely to use the task instructions to guide help form a macro-plan. The same can be said for the high scoring group although, caution needs to be exercised when extending this finding to this group as the number of participants in the high / low scoring groups was very small.

The Y3+ group engaged in 5 per cent more selective reading than the Y1 group. The differences were confined to careful local reading and selective global reading. The High scoring group engaged in 7 per cent more selective reading than the Low scoring group and the differences ranged across all four types of reading. This seems to suggest that selective reading has a role to play in scoring well on reading into writing tasks.

Attending to the most relevant sentences also seems to play a role in scoring well. However, where there were slight differences between the Y3+ group and the Y1 group, there were stark differences between the low / high scoring groups. The high scorers attended more to the more relevant sentences but at rates slightly below both the Y3+ and Y1 groups. However, the low scorers seemed to struggle to distinguish the most relevant sentences, producing only a weak correlation between attention and sentence relevance. This suggests that whilst attending to the most relevant sentences is not the only factor in scoring well, a failure to attend to the most relevant sentences is strongly associated with scoring poorly.

The increased attention on the emerging text (own work) that was seen for the Y3+ group did not continue for the high scorers. Although reading their own work accounted for almost half of all participants' fixations, the trend which saw more experienced participants devoting an even larger proportion of their time to reading their own work was not reflected in the figures for the high scorers. This suggests that participants spending more time reading their own work is not linked to scoring highly.

This concludes the discussion of the research findings. The next and final chapter offers a brief over of the study and draws the conclusions together to suggests what the findings from this study might add to our understanding of the way students tackle academic reading into writing tasks more generally.

# 6 Conclusion

This chapter aims to offer a brief review of the whole study, reminding the reader of the issues which led to the framing of the research questions (6.1) before going on to briefly summarise the data collection and analysis procedures. The findings of the research are summarised in section (6.2) and the limitations are acknowledged in section 6.3. Section 6.4 outlines the implications for teaching and testing academic reading into writing and section 6.5 reflects on the usefulness of eye tracking as a methodology. Finally, section 6.6 makes recommendations for further research.

## 6.1 Introduction

British higher education has seen radical changes over recent decades with wholescale changes to funding arrangements. This has, in part, been responsible for an increase in student numbers which in conjunction with the widening participation agenda, has seen a diversification of the student body.

However, greater diversity brings challenges in the form of supporting a wide range of students, with varying skills and needs, to achieve to the best of their ability whilst at university and in the world beyond. Academic literacy lies at the heart of succeeding in higher education. Whilst academic literacy incorporates a whole range of skills, reading and writing are at the heart.

Several researchers (Carson, 2001; Carson and Leki, 1993; Flower, 1990; Grabe, 2001, 2003; Johns, 1993; Leki and Carson, 1994, 1997; Lenski and Johns, 1997) have suggested that reading-into-writing ability is essential for academic success not least because reading-to-write is one of the commonest forms of assessment in higher education (Bridgeman and Carson, 1983; Hale et al., 1995; Rosenfeld, Leung, and Oltman, 2001).

The literature (Asension-Delaney, 2008, Cumming et al.,2005, Gebril, 2010) suggests that reading into writing elicits from students the same skills and processes as they are likely to engage in when completing written course assignments in higher education. Therefore, a thorough understanding of the cognitive processes involved in reading-to-write would seem to be essential to helping students develop the skills they need to thrive academically at university in the UK.

Some aspects of reading into writing have been explored and valuable contributions to a model of reading into writing have been made (Chan, 2013, Plakans and Gebril, 2012, Spivey and King, 1989, and Yu, 2008), however, there remains a gap in the literature in relation to how different types of reading are utilised in the process of reading to write.

Research on reading using eye tracking has suggested that eye-movement metrics such as fixation duration, saccade length and scan paths alter in response to the readers' cognitive processes and purposes for reading.

Therefore, this study aimed to use eye tracking to record students' eye-movements as they engaged in a short reading into writing task to establish the types of reading students engaged over the course of the task. In addition, the study investigated whether there were discernible differences in the reading patterns of students with more experience of performing reading into writing tasks and those students with less experience. Finally, the study investigated whether there were differences in the way high and low scoring participants tackled the reading in the task.

30 participants (15 Y1 undergraduates and 15 Y3+) were recruited from both the University of Bedfordshire (11 participants) and a range of other UK universities (19 participants).

The participants eye-movements were recording whilst they completed an hour-long academic reading into writing task. The researcher was able to observe the participant's reading and writing behaviour during the task and participants' eye-movements were recorded using eye tracking software for later replay and analysis.

After the participant completed the task the researcher conducted a retrospective think aloud session (RTA) in which the participant was asked to recollect their thought processes during the task prompted by watching a replay of their reading and writing activity. Finally, participants were asked to complete a short questionnaire.

The three research questions were

1. What are the characteristics of reading during an academic reading into writing test task?
2. What are the similarities and differences between the way first year undergraduates and third year or postgraduate students tackle an academic reading into writing test task?
3. What are the similarities and differences between the way high and low scoring participants tackle an academic reading into writing test task?

In order to answer the research questions the eye tracking data was analysed according to:

- How fixations were divided between six different areas of the screen (screen instructions, task instructions, written source texts, diagrams, move page controls and won work: the area for typing the answer)
- How fixations were divided between different sentences of the written source texts (This was calculated for individual sentences and for groups of

sentences when sentences were grouped according to their relevance to the task)

- The direction of movement and distance between fixations. Fixations were divided according to whether they had advanced less than 16 characters through the text, advanced more than 16 characters through the text, regressed a short distance (within the same line or to the line above) or regressed a long distance (more than one line above).
- The detection of patterns which suggested the amount and type of reading that participants engaged in. Reading was categorised according to whether the pattern suggested careful reading described in the literature or whether the pattern suggested a form of selective reading. Both careful and selective reading where then subdivided again, according to whether the episode of reading remained within a single sentence or progressed beyond a single sentence.

Eye tracking data generated by the eye tracker was triangulated with the RTA and questionnaire data. The same data was used to answer all three research questions however, for RQ1 the data for all participants was analysed. For RQ2 the data was split into two groups: first year undergraduate participants and third year undergraduate / postgraduate participants and the data for the two groups was compared. For RQ3 the data was limited to the five highest scoring and the five lowest scoring participants. The data for the high scoring group was compared with the data for the low scoring group.

## 6.2 Summary of the main findings

This research made five key findings in relation to reading into writing, each of which will now be discussed in turn.

1. Participants spent almost **half their time (47 per cent) fixating on their own emerging** text and about a third of their time reading the source texts. By comparison the task instructions were relatively poorly attended to (6 per cent). Participants fixated on their own work to evaluate and revise their own work. At times, they used it to help generate the text part of their answer.
2. The **fixations on the written text were more homogeneous** than fixations across all the AOIs with lower standard deviations. These fixations also reported a shorter mean fixation duration and contained much higher rates of regression than for reading reported in the literature.
3. Whilst fixating on the source texts, **careful reading accounted for less than 30 per cent of reading**. Other forms of **selective reading were utilised for the remaining 70 per cent** of the time. The predominance of selective reading appears to be as a result of participants **targeting their reading to spend more time on the more relevant sentences**, although several factors seem to interact to determine total time on each sentence. However, the researcher speculates that fixation on the more relevant sentences may be due to some participants using their macro plan to approach the source texts with clear reading goals. These goals led some participants to target the most relevant areas of the text.
4. When differences between the Y1 undergraduates (Y1) and year three undergraduates / postgraduates (**Y3+) groups were examined it emerged that the Y3+ group spent much more time fixating on their own work** than the Y1 group (4

hours 28 minutes versus three hours 13 minutes). It also emerged that the **Y3+ group engaged in more selective** reading than the Y1 group (74 per cent versus 69 per cent). This inevitably meant that the Y3+ group engaged in less careful reading than the Y1 group (26 per cent versus 31 per cent). The increased levels of selective reading may have contributed to the greater attention that the Y3+ group devoted to the more relevant sentences.

**5.** When the results for the five highest and five lowest scoring participants were compared it emerged that **the low scoring participants were much less effective at focusing their attention on the most relevant sentences.**

These findings offer an opportunity to draw some conclusions about the way the different cognitive processes interact during reading into writing. The researcher suggests that task representation can have a dramatic impact not only on the writer's final product but also in the way that the writer interacts with the source texts. Some participants began to develop their macro plan as they were reading the task instructions. They appeared to have a well-rehearsed procedure for accessing the demands of the task and developing a plan in response to the task.

Those participants that had a detailed macro-plan before they started to read (predominately Y3+ / high scoring participants), approached the reading with goals firmly established. Like a shopper going to the supermarket with a list, participants with a macro-plan approached the reading with a 'shopping list' of information that they needed to retrieve from the texts. Whilst the participants with the well-developed macro-plan could not know what content the texts would contain, they already had a clear plan of action. These participants went sequentially through the written source texts, picking up pieces of information and inserting them into their macro-plan at the appropriate point. By the time they had finished their first pass through the texts they had an idea of which parts of the task they had information for and roughly where that that information was. They are more likely to have reflected on the relevance of the different ideas to the task in hand. The ideas were already in an appropriate order and all that remained was to go back to the texts to re-read the most relevant ideas to achieve a full understanding of the information. Therefore, because more Y3+/ high scoring participants had a better developed macro-plan, this led to more selectively reading amongst the Y3+/high scoring participants and more sustained attention on the more relevant parts of the text.

Having extracted the relevant information, the 'early macro planners' have the outline of their essay in place and can begin elaborating. At this point these participants do not have too many demands on their working memory. They have a clear overview of the themes and issues that their essay includes and are more likely to have 'spare capacity' in their working memory to engage in knowledge transformation. They also have time to evaluate their texts more holistically in light of the task instructions and this leads to more high-level revisions and ultimately leads to a better product.

In contrast the participants with a less well-developed macro-plan (which included more Y1/ low scoring participants) approached the reading with less defined goals. The researcher suggests that in cases where participants began to read the source texts without first developing a macro-plan, participants were more inclined to make notes from the source texts in the order of reading and then later attempt to reorganise these notes to suit the task (sometimes successfully, sometimes less successfully). They would make an assessment of what the 'supermarket' (the text) had to offer and would select the best looking 'produce' (key ideas from the text). This process can still result in a nicely balanced 'meal' (essay), but it is a riskier process, a more cognitively demanding process. Participants who do not have a plan, progress through the text, picking out ideas. There is a tendency to pick out all the ideas contained in the text as they do not have a clear criterion, in the form of a macro-plan, to guide their selection process. Additionally, the ideas tend to be noted in the order in which they were retrieved rather than in the order required by the task.

These participants now have the task of not only going back to the text to develop their understanding of the ideas but also need to reorganise the ideas into the appropriate order and assess their relevance to the task. Under these circumstances it is more likely that participants waste time developing an understanding of points which are less relevant.

This reorganising of ideas is likely to be demanding and make demands on working memory. These participants were less likely to engage in knowledge transformation (Bereiter and Scardamalia, 1997 (section 2.2) as they are too occupied with reordering existing ideas rather than developing new ideas based on links or relationships between main ideas (connecting and generating). A less well developed understanding of the task may also mean that these participants are less likely to review their work holistically, instead they focus on low level revisions.

This description is, of course, an over simplification of the process. Participants will not have fallen into one camp or the other. Instead there will be a range of possibilities, with some participants with a clear macro plan failing to spot the most relevant information or failing to make links between key points. A good task representation will not guarantee that knowledge transformation takes place, it will just make it a little more likely. Conversely, a participant that started reading without a macro plan may have be able to pinpoint the most relevant sentences and develop a clear understanding of the task as they worked. But this process would have been more demanding, increasing the likelihood of encountering problems en route.

Having discussed the interpretation of the findings, the next section goes on to consider the limitations of this study.

## 6.3 The limitations of this study

For practical reasons, the task used in this study represented a very short reading into writing task compared to most course work assignments which take many hours to complete. Therefore, this task may not have elicited the full range of processes demanded of students completing genuine course work tasks. For example, the source texts may not have been long enough to elicit reading for gist in the way it operationalised in more time pressured conditions.

In the wider academic world students would not have their texts pre-selected. They would need to search through a variety of sources to locate and evaluate source materials. This was not the case for this task. Under such circumstances the types of reading engaged in may differ or occur in very different proportions to those reported for this task. There may even be other types of reading that were not used at all in this task.

The limited range of scores on this task (no participants scored less than five and the maximum score was nine, 19 participants scored eight or nine) suggest that the task may have been too easy to fully differentiate between some of the participants. In addition, the task may not have been complex enough to afford many opportunities for knowledge transformation to take place.

In common with many studies, practical reasons also restricted the number of participants in this study. The number of participants in eye tracking studies in the literature vary. Some elect to adopt a case study approach (Alamargot, 2010) and make detailed observations from five to ten participants. Several other studies in the literature (Alamargot, Chesnet, Dansac and Ros, 2006, Brunfaut and McCray, 2015) have drawn conclusions using 25 participants. Although there are some studies that have used more than 30 participants; Bax (2013) used 38 and Ashby, Rayner and Clifton (2005) used 44, these studies are quite rare. The practicalities of recruiting, recording and analysing the large quantity of data generated generally restricts the number of participants in eye- tracking studies to 50 or less. Therefore, although the findings from this study are informative about the reading processes during a short academic reading into writing task, caution needs to be exercised about generalising the findings to a larger population.

Rayner et al. (2012) point out that eye-movements do not equate to cognitive processes and researchers need to exercise caution in interpreting results. However, they also suggest that 'eye-movements are by far the best tool for understanding normal silent reading'. The researcher has tried, where possible, to triangulate the eye tracking data with RTA data to inform her interpretation but must acknowledge that some of the conclusions represent the researcher's interpretation of the data.

It also needs to be acknowledged that reading on screen may differ to reading on paper, however most of the data from the literature data from this study was compared to was also generated from reading on screen. The researcher suggests that reading on screen is the norm for many students today.

The researcher also acknowledges that in order to ensure that fixations on one line of text could be distinguished from fixations on the line above or below, the lines of text were quite widely spaced. This limited the amount of text on each 'page' creating a rather short page. The algorithm 'reset' the reading to local every time a page change occurred. As a result, levels of global reading may be under estimated.

## 6.4 Implications for testing and teaching

The findings of the research clearly show that selective reading skills perform a vital role in academic reading into writing tasks. As has been discussed at length, the researcher suggests that task representation can help generate a clear set of goals or criteria for reading and selective reading skills perform a vital role in helping to find and identify relevant information from the texts. However, the findings also suggested that careful reading then needs to be deployed once the relevant parts of the texts have been located. In order for students to successfully tackle reading into writing tasks, they are likely to need both careful and selective reading skills. If students have no selective reading skills, they are unlikely to be able to cope with the volume of texts encountered. Without careful reading skills, they will be unlikely to achieve a thorough understanding of the relevant parts of the text. Therefore, both careful and selective reading skills are essential, in conjunction with an understanding of when and where to deploy them.

However, the researcher is not aware of any university admissions tests that focus on selective reading skills. For many students, the bulk of their assessments will be in the form of reading into writing tasks (Bridgeman and Carson, 1983; Hale et al., 1995; Rosenfeld, Leung, and Oltman, 2001). Therefore, it would seem particularly important to be able to assess potential students in terms of their selective and careful reading skills. The development and application of tests of selective reading, in conjunction with tests of careful reading, would result in positive washback with students developing an awareness of the different types of reading and when and how they are utilised before they embark on their academic careers.

For those students, unlikely to encounter an admissions test before entry to higher education, it would seem imperative to teach selective reading skills, in addition to careful reading skills, to undergraduates at an early stage in their academic career. This may help them achieve a better grade point average from the start of their courses.

The researcher also suggests that task representation should be considered not just as a process but as a skill and taught to new undergraduates. As students become familiar with the routine of performing reading into writing task, they may well develop meta-cognitive strategies for task representation, but they may not. Therefore, it would seem imperative to assist students to develop an awareness and give instructed practice of task representation.

Inevitably, this leads the discussion full circle, back to the task. Given the critical role of task representation in the production a high-quality product, universities have a responsibility to ensure that their assignment briefs are well constructed, clearly written and unambiguous about the criteria for scoring well.

## 6.5 The usefulness of eye tracking as a methodology

Eye tracking has been widely used to successfully reveal some of the fine detail of how the brain controls the movements of the eye during reading (for example the SWIFT model by Engbert et al., 2005; E-Z Reader by Reichle et al., 1998). It has also been used very successfully to investigate reading at sentence level and below, for example, syntactic parsing, lexical access and word recognition. These studies are of value because they provide detailed descriptions of the eye-movement characteristics of careful reading. Whilst there are a few studies mentioned in the literature review which have used eye tracking to explore the higher-level processes involved in global text processing this study is the first, to the researcher's knowledge, to study text level reading processes on a reading into writing task.

The researcher suggests that several obstacles combine to make studying the higher-level processes involved in global text processing very challenging. First and foremost, there is the vast quantity of data that is generated. Even short periods of eye tracking generate large volumes of data, so eye tracking participants over longer periods of time exacerbates the problem exponentially. Handling large quantities of data is daunting and fraught with problems, slowing down well specified personal computers and introducing the risk of errors. Much of the data generated seems meaningless in and of itself and a lot of interpretation is required to translate the data into meaningful results. The researcher found it to be a long a rather steep learning curve.

Other factors are the issues described in the methodology section relating to the physical set up of the eye-tracker during data collection. Not all eye-types track well, eye colour and the thickness of eye lashes are likely to interact with the specifications of some eye-trackers to render some participants unsuitable. It can be very frustrating when a researcher has worked hard to recruit participants (not an easy job when participants are being asked to give up hours, rather than minutes of their time) only to find that a satisfactory calibration cannot be achieved. The researcher also found that some types of spectacles interfered with the signal as participants' gaze neared the bottom of the screen. Combined with other factors such as finding the appropriate setting (a large heavy table, on a concrete floor in a quiet room where light levels can be controlled) add to the difficulties when attempting to make recordings of a suitable quality.

However, all of these frustrations aside, once collected eye tracking data can provide a very rich seam of information, which if triangulated with RTA recordings can offer a valuable quantifiable insight into participants reading behaviours.

## 6.6 Recommendations for further research

Whilst eye good quality tracking data can be time difficult to collect and time consuming to analyse it offers a wide range of possibilities in terms of researching reading. The researcher suggests that care and attention need to be exercised when considering the practical issues of where and when to collect eye tracking data. The physical setting (the lighting, the stability of the furniture, the size of the computer screen) can interact with the data collection procedure to reduce the quality and reliability of the data. The display of reading materials on the screen also needs to be given careful consideration as font size and spacing between lines can affect the researcher's ability to analyse the data at word or sentence level.

It took the researcher a long time to develop the method of identifying reading from the eye tracking data and whilst the researcher's algorithm successfully identifies careful reading, it is not as successful in identifying the different types of selective reading suggested by the data. This is an area where, the researcher suggests, improvements could be made. The researcher used Microsoft Excel to process the data according to the rules of her algorithm. In reality, the choice of software was guided by the researcher's familiarity with the product, rather than a technical appraisal of the most suitable tool for the job in hand. The researcher suggests that, collaboration with colleagues from the computer science department would enable the development of a more user-friendly method of applying the algorithm. The addition of a metric for reading speed could also prove extremely useful.

In terms of further research using eye tracking to study selective reading the researcher suggests that Duggan and Payne's (2011) theory of satisficing in skim reading could be tested further by tracking the eye-movements of students whilst they skim reading and searching for increases in saccade length, decreases in fixation durations and increases in reading speed just before the reader abandons a section of text and jumps to the next section.

This research has provided a little insight into the way that selective reading interacts with the relevance of the text, but the researcher suggests that there is a lot more work to be done in this field. For example, as with Duggan and Payne's theory of satisfying, the researcher wonders whether readers' fixations become longer, and saccade length reduces as the reader begins to release that the sentence being parsed is relevant to their present reading goal. The researcher would also suggest that eye tracking could be usefully deployed to study the effects of abstractness / concreteness of text and its effect on the reader.

The researcher found the RTA data invaluable in triangulating and assisting with the interpretation of the eye tracking data and would suggest that this combination of research methods would be particularly useful for studying task representation. Technical limitations prevented in depth analysis of the reading of the task instructions on this study however the researcher suggests that eye tracking participants as they read task instructions, perhaps in conjunction with concurrent think aloud may be an area where eye tracking could be successfully used to provide further insight into the cognitive processes involved in task representation.

## 6.7 Closing comments

In conclusion the researcher reminds the reader that although caution needs to be exercised in interpreting the data because cognitive processes can only be inferred, as Rayner et al. suggest 'the work of the eyes in reading is an invaluable tool for understanding the process of reading' (Rayner et al.,2012:92).

Reading for reading into writing is very different to the careful reading described in the literature. It demands a wide range of selective reading skills and strategies in addition to careful reading skills. It also combines many different forms of reading. During the reading into writing process reading is constantly shifting in relation to the reader/writer's goals. At times reading is driven by the reader's goals (searching for relevant information in the text). At other times reading is used to push the composing process forward (for generation of ideas / connections) whilst at other times it is driven by the demands of the emerging text (searching for a word or fact to support an idea / point). Some of these processes are more meta cognitive than others but it appears that experience of performing the tasks results in a more structured goal driven approach at the outset. Some less experienced students arrive at the same process but seemingly through a less conscious process. This suggests that increasing awareness and conscious development of selective reading skills, in conjunction with developing task representation skills, could help inexperienced students produce better written work earlier in their courses.

## References


Alamargot, D. and Chanquoy, L. (2001) *Through the Models of Writing: With Commentaries by Ronald T. Kellogg & John R. Hayes.* Springer Science & Business Media.

Alamargot, D., Chesnet, D., Dansac, C. and Ros, C. (2006) 'Eye and pen: A new device for studying reading during writing', *Behavior research methods,* 38(2), pp. 287-299.

Alamargot, D., Plane, S., Lambert, E. and Chesnet, D. (2010) 'Using eye and pen movements to trace the development of writing expertise: case studies of a 7th, 9th and 12th grader, graduate student, and professional writer', *Reading and Writing,* 23(7), pp. 853-888.

Allen, S. (2004) 'Task representation of a Japanese L2 writer and its impact on the usage of source text information', *Journal of Asian Pacific Communication,* 14(1), pp. 77-89.

Anderson, R.C., Pichert, J.W. and Shirey, L.L. (1983) 'Effects of the reader's schema at different points in time.', *Journal of educational psychology,* 75(2), pp. 271.

Asención Delaney, Y. (2008) 'Investigating the reading-to-write construct', *Journal of English for Academic Purposes,* 7(3), pp. 140-150. doi: 10.1016/j.jeap.2008.04.001.

Ashby, J., Rayner, K. and 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), pp. 1065-1086.

Bax, S. (2013) 'The cognitive processing of candidates during reading tests: Evidence from eye-tracking', *Language Testing,* 30(4), pp. 441-465. doi: 10.1177/0265532212473244.

Beers, S., Quinlan, T. and Harbaugh, A. (2010) 'Adolescent students' reading during writing behaviors and relationships with text quality: an eyetracking study', *Reading and Writing,* 23(7), pp. 743-775. doi: 10.1007/s11145-009-9193-7.

Bereiter, C. and Scardamalia, M. (1987) *The psychology of written composition.* Lawrence Erlbaum Associates Hillsdale, NJ.

Biedert, R., Hees, J., Dengel, A. and Buscher, G. (2012) *A robust realtime reading-skimming classifier.* In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 123-130). ACM..

Blanchard, H.E., Pollatsek, A. and Rayner, K. (1989) 'The acquisition of parafoveal word information in reading', *Perception & psychophysics,* 46(1), pp. 85-94.

Boland, J.E. (2004) 'Linking eye movements to sentence comprehension in reading and listening', *The on-line study of sentence comprehension: Eye-tracking, ERP, and beyond,* pp. 51-76.

Bridgeman, B. and Carlson, S. (1983) 'Survey of academic writing tasks required of graduate and undergraduate foreign students', *ETS Research Report Series,* 1983(1), pp. 247-280.

Britt, M.A. and Sommer, J. (2004) 'Facilitating textual integration with macro-structure focusing tasks', *Reading Psychology,* 25(4), pp. 313-339. doi: 10.1080/02702710490522658.

Brunfaut, T. and McCray, G. (2015) 'Looking into test-takers' cognitive processes whilst completing reading tasks: a mixed-method eye-tracking and stimulated recall study'. *ARAG Research Reports Online.* British Council.

Brysbaert, M. and Vitu, F. (1998) 'Word skipping: Implications for theories of eye movement control in reading '*Eye guidance in reading and scene perception* Elsevier, pp. 125-147.

Campbell, C.S. and Maglio, P.P. (2001) *A robust algorithm for reading detection.* In *Proceedings of the 2001 workshop on Perceptive user interfaces* (pp. 1-7). ACM, pp. 1.

Caporossi, G., Alamargot, D. and Chesnet, D. (2004) *Using the computer to study the dynamics of the handwriting processes.* In *International Conference on Discovery Science* (pp. 242-254). Springer, Berlin, Heidelberg.

Carson, J. (2001) 'A task analysis of reading and writing in academic contexts', *Linking literacies: Perspectives on L2 reading-writing connections,* pp. 48-83.

Carson, J. and Leki, I. (1993) *Reading in the composition classroom: Second language perspectives.* Heinle and Heinle.

Carver, R.P. (1992) 'Reading rate: Theory, research, and practical implications', *Journal of Reading,* 36(2), pp. 84-95.

Cerdán, R. and Vidal-Abarca, E. (2008) 'The Effects of Tasks on Integrating Information from Multiple Documents', *Journal of Educational Psychology,* 100(1), pp. 209-222. doi: 10.1037/0022-0663.100.1.209.

Chai, C. (2006) 'Writing plan quality: Relevance to writing scores', *Assessing Writing,* 11(3), pp. 198-223.

Chan, S.H.C. (2013) *Establishing the validity of reading-into-writing test tasks for the UK academic context.* Doctor of Philosophy. University of Bedfordshire.

Chenoweth, N.A. and Hayes, J.R. (2001) 'Fluency in writing. Generating Text in L1 and L2', *Written Communication,* 18(1), pp. 80-98.

Clifton, C., Staub, A. and Rayner, K. (2007) 'Eye movements in reading words and sentences' *Eye Movements* Elsevier, pp. 341-371.

Cohen, A.D. and Upton, T.A. (2007) 'I want to go back to the text': Response strategies on the reading subtest of the new TOEFL', *Language Testing,* 24(2), pp. 209-250. doi: 10.1177/0265532207076364.

Cohen, A.D. and Upton, T.A. (2006) 'Strategies in responding to the new TOEFL reading tasks', *ETS Research Report Series,* 2006(1), pp. 162.

Conlon, G., Litchfield, A. and Sadlier, G. (2011) 'Estimating the value to the UK of education exports', BIS Research Paper Number 46, Department for Business Innovation and Skills.

Cooper, A. and Bikowski, D. (2007) 'Writing at the graduate level: what tasks do professors actually require?', *Journal of English for Academic Purposes,* 6(3), pp. 206-221.

Cumming, A.H., Kantor, R., Powers, D., Santos, T. and Taylor, C. (2000) *TOEFL 2000 writing framework.* Educational Testing Service Princeton, NJ.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K. and James, M. (2005) 'Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL', *Assessing Writing,* 10(1), pp. 5-43. doi: 10.1016/j.asw.2005.02.001.

Davies, A. (2008) *Assessing Academic English: Testing English Proficiency, 1950-1989: the IELTS Solution.* Cambridge University Press.

Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C. and Bivens-Tatum, J. (2008) 'Cognitive models of writing: Writing proficiency as a complex integrated skill', *ETS Research Report Series,* 2008(2), pp. 36.

Department for Business Innovation and Skills (2010) *Policy paper 2010 to 2015 government policy: higher education participation* Available at: https://www.gov.uk/government/publications/2010-to-2015-government-policy-higher-education-participation/2010-to-2015-government-policy-higher-education-participation (Accessed: 06/09/2018).

Duggan, G.B. and Payne, S.J. (2011) *Skim reading by satisficing: evidence from eye tracking.* ACM, pp. 1141.

Engbert, R., Nuthmann, A., Richter, E.M. and Kliegl, R. (2005) 'SWIFT: a dynamical model of saccade generation during reading.', *Psychological review,* 112(4), pp. 777.

Eysenck, M.W. and Keane, M.T. (2005) *Cognitive Psychology.* Hove: Psychology Press.

Fayol, M. (1999) 'From on-line management problems to strategies in written composition', *The cognitive demands of writing,* pp. 13-23.

Field, J. (2004) *Psycholinguistics: The key concepts,* London: Routledge.

Field, J. (2003) *Psycholinguistics. A resource book for students.* Abingdon, Oxon: Routledge.

Fisher, D.G. 1983, *An experimental study of eye movements during reading* (unpublished).

Fitzgerald, J. and Shanahan, T. (2000) 'Reading and Writing Relations and Their Development', *Educational Psychologist,* 35(1), pp. 39-50.

Flower, L. and Hayes, J. (1981) 'A cognitive process theory of writing', *College Composition and Communication,* 32(4), pp. 365-387.

Flower, L. (1990) *Reading-to-write: exploring a cognitive and social process.* New York: Oxford University Press.

Flower, L. and Hayes, J.R. (1981) 'A Cognitive Process Theory of Writing', *College Composition and Communication,* 32(4), pp. 365-387.

Flower, L., Stein, V., Ackerman, J. and Peck, W.C. (1990) *Reading-to-write: Exploring a cognitive and social process.* Oxford University Press on Demand.

Frenck-Mestre, C. (2005) 'Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings', *Second Language Research,* 21(2), pp. 175-198.

Gass, S.M. and Mackey, A. (2000) *Stimulated recall methodology in second language research illustrated edition.* London: Routledge.

Gebril, A. (2010) 'Bringing reading-to-write and writing-only assessment tasks together: A generalizability analysis', *Assessing Writing,* 15(2), pp. 100-117. doi: 10.1016/j.asw.2010.05.002.

Geisler, C. (1994) *Academic Literacy and the Nature of Expertise.* New York: Routledge.

Gil, L., Bråten, I., Vidal-Abarca, E. and Strømsø, H.I. (2010) 'Understanding and Integrating Multiple Science Texts: Summary Tasks are Sometimes Better Than Argument Tasks', *Reading Psychology,* 31(1), pp. 30-68. doi: 10.1080/02702710902733600.

Goldman, S.R. and Bloome, D.M. (2004) 'Learning to Construct and Integrate', in Healy, A.F. (ed.) *Experimental cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer* Washington, DC: American Psychological Association., pp. 169-182.

Grabe, W. (2003) 'Reading and writing relations: Second language perspectives on research and practice', *Exploring the dynamics of second language writing,* pp. 242.

Grabe, W. (2001) 'Reading-writing relations: Theoretical perspectives and instructional practices', *Linking literacies: Perspectives on L2 reading-writing connections,* pp. 15-47.

Grabe, W. and Kaplan, R.B. (1996) *Theory and practice of writing: an applied linguistic perspective.* London: Longman.

Grabe, W. and Stoller, F. (2011) *Teaching and researching reading.* 2nd ed.

Hale, G., Taylor, C., Bridgeman, B., Carson, J., Kroll, B. and Kantor, R. (1995) 'A study of writing tasks assigned in academic degree programs', *ETS Research Report Series,* 1995(2), pp. 61.

Hannus, M. and Hyönä, J. (1999) 'Utilization of illustrations during learning of science textbook passages among low-and high-ability children', *Contemporary educational psychology,* 24(2), pp. 95-123.

Hayes, J.R. (1996) 'A New Framework for Understanding Cognition and Affect in Writing', in Levy, C.M. and Ransdell, S. (eds.) *The Science of Writing. Theories, Methods, Individual Differences and Applications.* Mahwah, New Jersey: Lawrence Erbaum, pp. 1-27.

Hayes, J.R. and Flower, L. (1980) 'Identifying the organization of writing processes', in Gregg, L.W. and Steinberg, E.R. (eds.) *Cognitive Processes in Writing* Hillsdale, New Jersey: Lawrence Erlbaum, pp. 3-30.

Hayes, J.R., Flower, L., Schriver, K.A., Stratman, J.F. and Carey, L. (1987) '5 Cognitive processes in revision', *Advances in Applied Psycholinguistics: Volume 2, Reading, Writing, and Language Learning,* 2, pp. 176.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. and Van de Weijer, J. (2011) *Eye tracking: A comprehensive guide to methods and measures.* OUP Oxford.

Hyönä, J. (1994) 'Processing of topic shifts by adults and children', *Reading Research Quarterly,* pp. 77-90.

Hyönä, J. (1985) 'The On-line Investigation of Reading a Text: Methods and a Model', *AFinLAn vuosikirja,* pp. 43-55.

Hyönä, J., Lorch Jr, R.F. and Kaakinen, J.K. (2002) 'Individual differences in reading to summarize expository text: Evidence from eye fixation patterns.', *Journal of educational psychology,* 94(1), pp. 44.

Hyönä, J., Lorch, R.F. and Rinck, M. (2003) 'Eye movement measures to study global text processing' *The Mind's Eye* Elsevier, pp. 313-334.

Inhoff, A.W. and Rayner, K. (1986) 'Parafoveal word processing during eye fixations in reading: Effects of word frequency', *Perception & psychophysics,* 40(6), pp. 431-439.

Johansson, R., Johansson, V., Wengelin, Å and Holmqvist, K. (2008) 'Reading during writing, four groups of writers', *Department of Linguistics and Phonetics, Centre for Languages and Literature, Lund University Working Papers,* 53.

Johns, A.M. (1993) 'Reading and writing tasks in English for academic purposes classes: Products, processes, and resources', *Reading in the composition classroom: Second language perspectives,* pp. 274-289.

Johnson, R.B. and Onwuegbuzie, A.J. (2004) 'Mixed methods research: A research paradigm whose time has come', *Educational researcher,* 33(7), pp. 14-26.

Just, M.A. and Carpenter, P.A. (1980) 'A theory of reading: from eye fixations to comprehension.', *Psychological review,* 87(4), pp. 329.

Kantz, M. (1990) 'Helping students use textual sources persuasively', *College English,* pp. 74-91.

Kellogg, R.T. (1996) 'A model of working memory in writing.', in Levy, C.M. and Ransdell, S. (eds.) *The Science of Writing* Mahwah, NJ: Lawrence Erlbaum Associates, Inc, pp. 57-72.

Kellogg, R.T. (1994) *The Psychology of Writing,* New York: Oxford University Press.

Kellogg, R.T. (2006) 'Professional Writing Expertise.' In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), The Cambridge handbook of expertise and expert performance (pp. 389-402). New York, NY, US: Cambridge University Press.

Kellogg, R.T. (1987) 'Writing performance: Effects of cognitive strategies', *Written communication,* 4(3), pp. 269-298.

Khalifa, H. and Weir, C.J. (2009) *Examining Reading: Research and practice in assessing second language reading.* Ernst Klett Sprachen.

Kintsch, W. and Van Dijk, T.A. (1978) 'Toward a model of text comprehension and production.', *Psychological review,* 85(5), pp. 363.

Kliegl, R., Nuthmann, A. and Engbert, R. (2006) 'Tracking the mind during reading: The influence of past, present, and future words on fixation durations.', *Journal of Experimental Psychology: General,* 135(1), pp. 12.

Kollmorgen, S. and Holmqvist, K. (2007) 'Automatically detecting reading in eye tracking data', *Lund University Cognitive Studies*.

Kuperman, V. and Van Dyke, J.A. (2011) 'Effects of individual differences in verbal skills on eye-movement patterns during sentence reading', *Journal of memory and language,* 65(1), pp. 42-73.

Kuzborska, I. (2018) 'Interactive Reading Strategies', *The TESOL Encyclopedia of English Language Teaching,* pp. 1-7.

Lea, M.R. and Street, B.V. (2006) 'The" academic literacies" model: Theory and applications', *Theory into practice,* 45(4), pp. 368-377.

Leki, I. (1993) 'Reciprocal Themes in ESL Reading and Writing', in Leki, I. and Carson, J. (eds.) *Reading in the Composition Classroom: Second Language Perspectives* Boston, Massachusetts, USA: Heinle & Heinle, pp. 9-32.

Leki, I. and Carson, J. (1997) '"Completely different worlds": EAP and the writing experiences of ESL students in university courses', *TESOL quarterly,* 31(1), pp. 39-69.

Leki, I. and Carson, J.G. (1994) 'Students' perceptions of EAP writing instruction and writing needs across the disciplines', *TESOL quarterly,* 28(1), pp. 81-101.

Lenski, S.D. (1998) 'Intertextual Intentions: Making Connections across Texts.', *Clearing House,* 72(2), pp. 74-84.

Lenski, S.D. and Johns, J.L. (1997) 'Patterns of reading-to-write', *Literacy Research and Instruction,* 37(1), pp. 15-38.

Lorch, R.F., Lorch, E.P. and Klusewitz, M.A. (1993) 'College students' conditional knowledge about reading.', *Journal of educational psychology,* 85(2), pp. 239.

Lorch, R.F., Lorch, E.P. and Matthews, P.D. (1985) 'On-line processing of the topic structure of a text', *Journal of memory and language,* 24(3), pp. 350.

McConkie, G.W. (1981) 'Evaluating and reporting data quality in eye movement research', *Behavior Research Methods & Instrumentation,* 13(2), pp. 97-106.

McCutchen, D. (1996) 'A capacity theory of writing: Working memory in composition', *Educational Psychology Review,* 8(3), pp. 299-325.

Mellors-Bourne, R., Humphrey, C., Kemp, N. and Woodfield, S. (2013) *BIS Research Paper 128: The wider benefits of international higher education in the UK;* Department for Business Innovation and Skills. Available at: https://dera.ioe.ac.uk/18261/1/13-1172-wider-benefits-of-international-higher-education-in-the-uk.pdf (Accessed: 25/09/2018).

National Association for Language Development in the Curriculum *EAL pupils in schools.* Available at: http://www.naldic.org.uk/research-and-information/eal-statistics/eal-pupils/ (Accessed: 25/09/2018).

Nesi, H. (2008) 'A multidimensional analysis of student writing across levels and disciplines', In *Taking the Measure of Applied Linguistics: Proceedings of the BAAL Annual Conference,* University of Swansea *(pp. 11-13).*

Newell, A. and Simon, H.A. (1972) *Human Problem Solving.* Englewood Cliffs, New Jersey, USA: Prentice Hall.

Newell, A., Shaw, J.C. and Simon, H.A. (1958) 'Elements of a theory of human problem solving.', *Psychological review,* 65(3), pp. 151.

Oakes, L.M. (2010) 'Infancy guidelines for publishing eye-tracking data', *Infancy,* 15(1), pp. 1-5.

Olive, T. and Kellogg, R.T. (2002) 'Concurrent activation of high-and low-level production processes in written composition', *Memory & cognition,* 30(4), pp. 594-600.

Olsen, A. and Matos, R. (2012) Identifying parameter values for an I-VT fixation filter suitable for handling data sampled with various sampling frequencies. In *proceedings of the symposium on eye tracking research and applications.* (pp.317-320). ACM.

Parodi, G. (2007) 'Reading–writing connections: Discourse-oriented research', *Reading and Writing,* 20, pp. 225-250.

Perfetti, C.A. and Adlof, S.M. (2012) 'Reading comprehension: A conceptual framework from word meaning to text meaning', *Measuring up: Advances in how to assess reading ability,* pp. 3-20.

Perfetti, C.A. (1999) 'Comprehending written language: A blueprint of the reader', *The neurocognition of language,* pp. 167-208.

Pichert, J.W. and Anderson, R.C. (1977) 'Taking different perspectives on a story.', *Journal of educational psychology,* 69(4), pp. 309.

Plakans, L. (2010) 'Independent vs. Integrated Writing Tasks: A Comparison of Task Representation', *TESOL Quarterly,* 44(1), pp. 185-194. doi: 10.5054/tq.2010.215251.

Plakans, L. (2009a) 'Discourse synthesis in integrated second language writing assessment', *Language Testing,* 26(4), pp. 561-587.

Plakans, L. (2009b) 'The role of reading strategies in integrated L2 writing tasks', *Journal of English for Academic Purposes,* 8(4), pp. 252-266. doi: 10.1016/j.jeap.2009.05.001.

Plakans, L. (2008) 'Comparing composing processes in writing-only and reading-to-write test tasks', *Assessing Writing,* 13(2), pp. 111-129. doi: 10.1016/j.asw.2008.07.001.

Plakans, L. and Gebril, A. (2012) 'A close investigation into source use in integrated second language writing tasks', *Assessing Writing,* 17, pp. 18-34.

Quinlan, T., Loncke, M., Leijten, M. and Van Waes, L. (2012) 'Coordinating the Cognitive Processes of Writing: The Role of the Monitor', *Written Communication,* 29, pp. 345.

Rayner, K. and Pollatsek, A. (1989) *The Psychology of Reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K. (2009) 'Eye movements and attention in reading, scene perception, and visual search', *The Quarterly journal of experimental psychology,* 62(8), pp. 1457-1506. doi: 10.1080/17470210902816461.

Rayner, K. and Duffy, S.A. (1986) 'Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity', *Memory & cognition,* 14(3), pp. 191-201.

Rayner, K., Foorman, B.R., Perfetti, C.A., Pesetsky, D. and Seidenberg, M.S. (2001) 'How psychological science informs the teaching of reading', *Psychological science in the public interest,* 2(2), pp. 31-74.

Rayner, K., Juhasz, B.J. and Pollatsek, A. (2005) 'Eye movements during reading', *The science of reading: A handbook,* pp. 79-97.

Rayner, K., Pollatsek, A., Ashby, J. and Clifton Jr, C. (2012) *Psychology of reading.* Psychology Press.

Reichle, E.D., Pollatsek, A., Fisher, D.L. and Rayner, K. (1998) 'Toward a model of eye movement control in reading.', *Psychological review,* 105(1), pp. 125.

Reichle, E.D., Rayner, K. and Pollatsek, A. (2003) 'The EZ Reader model of eye-movement control in reading: Comparisons to other models', *Behavioral and brain sciences,* 26(4), pp. 445-476.

Rinke, M., Gamez, E., Diaz, J. and De Vega, M. (2003) 'Processing of temporal information: Evidence from eye movements', *Memory and Cognition,* 31(1), pp. 77-86.

Rosenfeld, M., Leung, S. and Oltman, P.K. (2001) *The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels.* Educational Testing Service.

Rothkopf, E.Z. and Billington, M.J. (1979) 'Goal-guided learning from text: inferring a descriptive processing model from inspection times and eye movements.', *Journal of educational psychology,* 71(3), pp. 310.

Rouet, J.F. (2006) *The Skills of Document Use: from text comprehension to web-based learning.* New Jersey, USA: Lawrence Erlbaum.

Ruiz-Funes, M. (2001) 'Task representation in foreign language reading-to-write', *Foreign language annals,* 34(3), pp. 226-234.

Rumelhart, D.E. (1994) *Toward an interactive model of reading.* International Reading Association.

Scardamalia, M. and Bereiter, C. (1987) 'Knowledge telling and knowledge transforming in written composition', *Advances in applied psycholinguistics,* 2, pp. 142-175.

Scardamalia, M. and Paris, P. (1985) 'The function of explicit discourse knowledge in the development of text representations and composing strategies', *Cognition and instruction,* 2(1), pp. 1-39.

Segev-Miller, R. (2007) 'Cognitive processes in discourse synthesis: The case of intertextual processing strategies', *Studies in writing,* 20, pp. 231.

Shaw, S.D. and Weir, C.J. (2007) *Examining Writing - research and practice in assessing second language writing.* Cambridge University Press.

Simola, J., Salojärvi, J. and Kojo, I. (2008) 'Using hidden Markov model to uncover processing states from eye movements in information search tasks', *Cognitive systems research,* 9(4), pp. 237-251.

Smeets, W. and Solé, I. (2008) 'How adequate task representation can help students write a successful synthesis', *Zeitschrifts Schreiben*.

Spivey, N. (1990) 'Transforming texts: Constructive processes in reading and writing', *Written Communication,* 7(2), pp. 256-287.

Spivey, N. and King, J.R. (1989) 'Readers as Writers Composing from Sources', *Reading Research Quarterly,* 24(1 Winter), pp. 7-26.

Spivey, N.N. (1997) *The constructivist metaphor: Reading, writing, and the making of meaning.* Academic Press San Diego, CA.

Spivey, N.N. (1990) 'Transforming texts: Constructive processes in reading and writing', *Written communication,* 7(2), pp. 256-287.

Spivey, N.N. (1984) 'Discourse Synthesis: Constructing Texts in Reading and Writing.'

Stanovich, K.E. (1980) 'Toward an interactive-compensatory model of individual differences in the development of reading fluency', *Reading research quarterly,* pp. 32-71.

Stein, V. (1989) *Exploring the cognition of reading-to-write.* Center for the Study of Writing.

Stevenson, M., Schoonen, R. and de Glopper, K. (2006) 'Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL', *Journal of Second Language Writing,* 15(3), pp. 201-233.

Stotsky, S. (1983) 'Research on Reading/Writing Relationships: A Synthesis and Suggested Directions.', *Language arts,* 60(5), pp. 627-642.

Swales, J.M. and Lindemann, S. (2002) 'Teaching the literature review to international graduate students', *Genre in the classroom: Multiple perspectives,* pp. 105-119.

Swanson, H.L. and Berninger, V.W. (1996) 'Individual differences in children's working memory and writing skill', *Journal of experimental child psychology,* 63(2), pp. 358-385.

The Complete University Guide *University of Bedfordshire* Available at: https://www.thecompleteuniversityguide.co.uk/bedfordshire/ (Accessed: 25/09/2018).

Tobii, A.B. (2017) *Tobii Pro Studio* [Computer program] Available at: https://www.tobiipro.com/product-listing/tobii-pro-studio/

Tobii and AB (2016) *Tobii Studio Users' Manual.* Tobii AB. Available at: https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf (Accessed: 01/02/2019).

Torrance, M. and Galbraith, D. (2006) 'The processing demands of writing', *Handbook of writing research,* , pp. 67-80.

Urquhart, A.H. and Weir, C.J. (1998) *Reading in a second language: Process, product and practice.* Harlow: Longman.

Vauras, M., Hyönä, J. and Niemi, P. (1992) 'Comprehending coherent and incoherent texts: Evidence from eye movement patterns and recall performance', *Journal of research in reading,* 15(1), pp. 39-54.

Weir, C.W. and Chan, S. (forthcoming) *Research and practice in assessing academic reading: the case of IELTS.* Studies in Language Testing, Volume 51. Cambridge University Press.

Wengelin, Å, Leijten, M. and Van Waes, L. (2010) 'Studying reading during writing: new perspectives in research', *Reading and Writing,* 23(7), pp. 735-742.

Wolverton, G.S. and Zola, D. (1983) 'The temporal characteristics of visual information extraction during reading', *Eye movements in reading: Perceptual and language processes,* pp. 41-51.

Wu, R.Y. (2011) 'Establishing the validity of the General English Proficiency Test Reading Component through a critical evaluation on alignment with the Common European Framework of Reference'.

Yu, G. (2008) 'Reading to summarize in English and Chinese: A tale of two languages?', *Language Testing,* 25(4), pp. 521-551.

# Appendix 1 Ethical approval

University of Bedfordshire

Centre for Research
in English Language
Learning and
Assessment
(CRELLA)
Putteridge Bury
Campus Luton, LU2
8LE
United Kingdom
t +44 (0)1582 489159
www.beds.ac.uk/crella

26 April 2016

To whom it may concern,

**Re: Ethical approval for Nicola Latimer's research project**

Please accept this letter as confirmation that Nicola Latimer's research project titled "Investigating the Cognitive Processes in Academic Reading into Writing" has been scrutinised by CRELLA's ethics committee and has been approved.

Yours faithfully,

Dr Chihiro Inoue

Lecturer in Language Assessment
CRELLA, University of Bedfordshire

## Appendix 2 Ethics statement

In accordance with guidance set out by BAAL (British Association of Applied Linguistics) the following statement outlines areas of ethical concern and summarises how this study proposes to deal with these issues.

**Ensuring there are no adverse effects on participants**.
In order to ensure that participants are treated fairly and sensitively with due regard to the avoidance of stress, intrusion and exploitation, this study will:

- Seek to pilot any materials with volunteers before extending them to a wider audience of students.
- exclude any students who are under the age of 18 or who have specific learning difficulties. This is because the writer believes the assessment / achievement scores of these groups may vary substantially from adults who do not experience any learning difficulties and will therefore not be valuable for the purposes of this research.

**Use of video / audio recording equipment**
Where interviews / computer sessions are recorded using audio / video the study will;

- Explain the purposes of the recordings to all participants to help them feel more at ease.
- Give careful consideration to the positioning of recording equipment to minimise its effect on participants whilst still being able to collect the required data,
- Make it clear to participants that the recording can be halted at any time if they decide to withdraw their consent.
- Allow a period of adjustment and any findings will take into consideration the influence on the participants of being recorded.
- Allow participants to refuse permission for certain parts of the recordings to be used if they so desire.

**Informed written consent.**
This study will

- Seek informed written consent from all parties / participants, involved either directly or indirectly, before any data is collected.
- Ensure that statements requesting informed written consent are written in language appropriate to the English levels of the participants, make it clear potential participants are under no obligation to participate, make explicitly clear participants right to withdraw at any time (without giving reasons), include unambiguous information about the objectives of the research and its possible consequences, give precise details of the data to be collected and the ways in which privacy and anonymity will be preserved.
- Make it clear to participants of the study that the relationship between the writer and the participant does not oblige them to participate in the study, any wish not to participate will be respected and will not affect the future relationship of the writer and respondent in any way.

**Safeguarding of data.**
This study will comply with the provisions of the 1998 Data Protections act by

- allocating all respondents' a reference number which will be used to identify their responses / materials

- not make reference to names in any published findings
- anonymise any responses or quotations reported in the research.
- password protecting any electronic document that contains any personal data
- shred any printed materials that contain any personal data.

**Debriefing of participants.**

This study will ensure participants are adequately debriefed by

- Making the findings of the study available in language appropriate to the English levels of the participants.
- Encouraging participants to comment on the fairness, relevance and accuracy of the findings and include these comments in the published report of the study.

**Maintenance of impartiality**.

In order to ensure that the findings of the research are represented fairly this study will:-

- Include a short biography of the writer, addressed to the reader, with the aim of exposing any unintended bias on the part of the writer due to unavoidable cultural influences / assumptions. This biography will make reference to the researcher's ethnicity, gender, educational background, age and religious beliefs.
- Include a statement, addressed to the reader, outlining the researcher's relationship to / with any of the participants, as the nature of the relationship may affect the likelihood of participants reporting/ behaving in a totally natural way.
- Include a statement indicating if the findings have been edited / changed in any way at the request of any of the participants.
- Attend to a wide variety of sources and perspectives to avoid presenting a biased view.

**Potential damage to the Reputation of the University.**

In order to ensure this study does not adversely affect the reputation of the university it will:

- Not intentionally misrepresent the work of others.
- Not present the work of others as its own.
- Acknowledge and give credit to all sources using the Harvard system of referencing.
- Not approach any people or organisations regarding participation in the study until ethical approval has been granted.
- Ensure that all material issued as a result of the study is well presented and free from errors.
- Ensure that any contact with people or organisations, as part of / as a result of the study, is conducted in a professional manner.

# Appendix 3 Research Information sheet

***RESEARCH INFORMATION SHEET***

<u>An investigation of the cognitive processing that tertiary level students engage in whilst completing academic reading into writing tasks from the UoB Reading into writing test.</u>

**Research carried out by:** Nicola Latimer, PhD student at University of Bedfordshire. CRELLA, Putteridge Bury Campus,Hitchin Road, Luton LU2 8LE

**Contact details: nicola.latimer@beds.ac.uk Tel:07718455270**

**<u>Background</u>**

Lots of research has been done in the past about reading and writing separately but not much has been done about how students coordinate their reading and writing skills whilst they are completing academic assignments.

There are several tests which indicate if a student has the right academic English skills to cope with assignments at university, however only a few require candidates to combine reading and writing in a similar way to that of writing an assignment based on reading source texts (books, journal articles etc.). One test which does demand integrated reading and writing is the University of Bedfordshire Reading into writing test. I want to explore how candidates coordinate their reading and writing skills whilst completing a task from the test.

By investigating how candidates coordinate their reading and writing behaviours on the test I hope to be able to understand how these skills are integrated and whether common patterns of reading into writing behaviour exist.

I am investigating the reading and writing processes used by candidates completing the test. I am interested in <u>how</u> you complete the task rather than making judgements about your ability to read and write in English. For the remainder of this document I will refer to the students who agree to help me as 'participants'.

**<u>Anonymity</u>**

All the data I gather will be reported anonymously - I will not publish participants' names or contact details. If I use samples of participants' writing or report participants' answers no one will know who wrote / said it. I will not publish any video images or voice recordings of participants.

Individual participants perceived level of English will not be reported to anyone within the university or to anyone, or any organisations, outside the university.

Any personal data gathered will be stored in accordance with data protection legislation.

## Extent of participation

Participation in the study will involve

1) An initial interview to gather information about participants linguistic and educational backgrounds. This will be audio recorded. (Approx. 20-30 mins)

2) Attending a session lasting approx. 1 hour 40 mins.

During the session participants do a reading & writing task on a computer. Participants will be asked to read some on-screen information and type a short written response. This session will be recorded using webcam, the participant's eye-movements and typing are recorded. Immediately after completing the computer session participants will be asked to watch the replay of their reading and writing activity with the researcher and answer a few questions about their reading and writing activities. Finally, participants will be asked to complete a short questionnaire.

**The right to withdraw**

Participants have the right to withdraw from the research at any time, without giving any reason.

**If you would like more information before deciding...**

Please feel free to ask any questions. I will be happy to explain more about my research and why I chose to do research in this area.

**Participants will receive £20 to cover any travel expenses incurred.**

# Appendix 4 Research consent form

## RESEARCH CONSENT FORM

<u>A comparison of the cognitive processing that tertiary level students engage in whilst completing academic reading into writing tasks and the UoB Reading into writing test tasks.</u>

**Research carried out by:** Nicola Latimer, PhD student at University of Bedfordshire. CRELLA, Putteridge Bury Campus,Hitchin Road, Luton LU2 8LE

**Contact details: nicola.latimer@beds.ac.uk Tel:07718455270**

Please Initial Box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason. ☐

3. I agree to take part in the above study. ☐

4. I agree to the initial interview being audio recorded. ☐

5. I agree to the computer sessions being recorded in the following ways: ☐
   a) video & audio recording of the session
   b) recording of my eye-movements using eye tracking equipment
   c) typing activity being logged by key-stroke logging software

6. I agree to the use of anonymised information submitted via a questionnaire. ☐

7. I agree to the use of anonymised quotes in publications ☐

____________________ ________________ ____________________

Name Date Signature

## Appendix 5 Assessment criteria and rating scales

# ASSESSMENT CRITERIA AND RATING SCALES
# FOR UoB READING INTO WRITING TEST

**Overall assessment approach**

A 3-band scale is proposed which reflects 3 levels of performance quality as follows:

U = unclassified performance the text is too short, completely off topic, illegible or plagiarised (No further marking is needed).

A = a quality of performance which meets or exceeds minimum requirements expected for students, and which therefore suggests that no additional systematic remedial intervention will be needed post entry

B = a quality of performance which falls below minimum requirements expected for students, and which therefore suggests that *some* additional systematic remedial intervention will be needed post entry

C = a quality of performance which falls significantly below minimum requirements expected for students, and which therefore suggests that a substantial level of systematic remedial intervention will be needed post entry

An analytical assessment approach is proposed, using 3 criteria, which will permit some a more fine-grained evaluation of reading–into-writing performance. This could have the potential for providing some meaningful diagnostic information to help inform decisions about the nature and extent of any systematic remedial intervention given to students post entry.

**Proposed analytical assessment criteria**

Plagiarism (copying chunks of more than 5 consecutive words) will result in overall failure.

1. **Relevance and adequacy of content** (coverage of key points)
2. **(Compositional) Organisation** (cohesion and coherence)
3. **Language** (choice and control of lexis, grammar)

Each of these 3 criteria is defined more fully below, immediately before the relevant scale and its draft descriptors.

**Score reporting procedures**

Scores could be reported both *overall* - for quick decision-making purposes (by summing the 3 subscores from the analytical criteria to produce a 'grade' of A, B or C), and also *at the subscore level* - in order to provide a profile of performance quality for diagnostic purposes and follow-up remedial intervention. The grade boundaries will need to be determined (presumably according to need and available resources). One scenario might be as follows:

| | | |
|---|---|---|
| A score of 8-9 | = | Grade A (no intervention required) |
| A score of 6/7 | = | Grade B ('low-level' intervention needed) |
| A score of 5 or < | = | Grace C ('high-level' intervention needed) |

### 1.Relevance and adequacy of content

This refers to the extent to which the writer has responded appropriately to the task and the specific instructions given about the relationship between the input reading material and the written output. It covers the inclusion of 4 essential key points, as well as communicative effect on the reader (i.e. awareness of writer-reader relationship and appropriate level of formality).

| Band | Descriptor of performance quality |
|---|---|
| **3** | • Relevant and fully adequate response to the task.<br>• All 4 key points required in the task included and expanded appropriately.<br>• Achieves desired communicative effect on target reader. |
| **2** | • Partially successful response to the task.<br>• One or two key points inadequately covered or omitted, and/or some irrelevant material included.<br>• May fail to communicate clearly to target reader and/or achieve the desired effect. |
| **1** | • Limited response to the task.<br>• More than 2 key points omitted and/or considerable irrelevance/repetition, possibly due to misinterpretation of the task.<br>• Fails to achieve the desired effect because considerable effort will be required of the reader. |

NOTE: The coloured shading in the tables below represents a 'traffic light' approach, i.e.

- green signifies an *adequate* performance, by a student who should not need additional EL/study skills support post entry
- yellow signifies a *below adequate* performance, by student who will benefit from some targeted EL/study skills intervention post entry;
- red signifies a *significantly weak* performance, by a student who will be a high-priority candidate for substantial EL/study skills intervention post entry.

This 3-category approach should help to filter out those who do not really need extra help in a systematic way, and to identify 2 groups for follow-up training- a high-need group and a lower-need group.

## 2. Organisation

This refers to the way in which the written production has been structured and organised in terms of the overall format, the grouping and sequencing of ideas in paragraphs, and the coherence of the argumentation throughout. It covers notions of cohesion and coherence, across both sentences and paragraphs.

| Band | Descriptor of performance quality |
|---|---|
| 3 | ▯ Overall shape and internal pattern clear.<br>▯ Information and ideas organised logically and coherently.<br>▯ Satisfactory use of cohesion resulting in effective communication.<br>▯ Format appropriate to the purpose of the task and audience. |
| 2 | ▯<br>Evidence of some underlying structure but not adequately controlled.<br>▯ Information and ideas partially organised but sometimes incoherent.<br>▯ Communication not always effective due to inadequate cohesive control.<br>▯ Format may be inappropriate to the purpose of the task and audience. |
| 1 | ▯ Little evidence of overall shape or underlying structure. Information and ideas<br>▯ presented largely incoherently.<br>▯ Linking devices rarely used and erratic use of punctuation<br>▯ Attempt at appropriate format unsuccessful or inconsistent. |

### 3. Language

This refers to the clarity of linguistic expression in English, including the selection and control of grammar and vocabulary items (syntactic accuracy, lexical precision, use of topic/discourse markers). It also includes stylistic choices relating to academic register.

| Band | Descriptor of performance quality |
|---|---|
| **3** | Wide range of structure and lexis with a variety of linking devices. Errors minimal but do not impede and may be due to ambitious attempts at more complex language.<br>Register appropriate to purpose of the task and audience.<br>Few punctuation and spelling errors. |
| **2** | Adequate range of structure and lexis though with little complex syntax and/or lexical precision.<br>Errors can distract the reader and may obscure communication at times.<br>Register sometimes inappropriate.<br>Some errors in punctuation and spelling. |
| **1** | Unexpectedly narrow range of structure and vocabulary - tends to be simplistic or repetitive.<br>Frequent and/or basic errors in lexis and grammar obscure communication.<br>Little or no awareness of appropriate register.<br>Poor control of punctuation and spelling. |

## Appendix 6 Eye tracking data quality: confidence report scorers

| Participant | Number of fixations given a validity rating of 4 | Number of fixations* | % of fixations that Tobii interpolated |
|---|---|---|---|
| P11 | 272 | 9558 | 3% |
| P12 | 1226 | 7348 | 17% |
| P14 | 744 | 5785 | 13% |
| P15 | 497 | 5168 | 10% |
| P23 | 534 | 3897 | 14% |
| P25 | 559 | 5623 | 10% |
| P26 | 268 | 4313 | 6% |
| P28 | 384 | 8237 | 5% |
| P29 | 479 | 9381 | 5% |
| P31 | 265 | 8453 | 3% |
| P32 | 545 | 6754 | 8% |
| P34 | 215 | 3655 | 6% |
| P35 | 508 | 8107 | 6% |
| P36 | 181 | 6631 | 3% |
| P37 | 433 | 7606 | 6% |
| P38 | 292 | 5779 | 5% |
| P40 | 530 | 6922 | 8% |
| P41 | 330 | 7937 | 4% |
| P42 | 386 | 10492 | 4% |
| P45 | 585 | 6708 | 9% |
| P46 | 328 | 7961 | 4% |
| P47 | 355 | 7983 | 4% |
| P48 | 382 | 8019 | 5% |
| P49 | 476 | 7499 | 6% |
| P50 | 203 | 5560 | 4% |
| P51 | 636 | 10010 | 6% |
| P52 | 539 | 8740 | 6% |
| P53 | 81 | 4842 | 2% |
| P54 | 426 | 9638 | 4% |
| P55 | 504 | 5729 | 9% |
| **Total** | **13163** | **214335** | |
| Av across all participants | | | 6% |

| |
|---|
| * number of fixations is slightly higher than fixations analysed for the research because this figure included the minutes before and after the task. |

## Appendix 7 Stimulated recall protocol

Gass and Mackey (2000) advise that a detailed research protocol is important to facilitate the successful execution of a stimulated recall session. They suggest creating a clear protocol for the researcher to ensure that nothing is forgotten and to ensure that the wording focuses the participant on reporting their thoughts at the time of the activity rather than their present reaction to the replay. The following protocol is based on examples presented in Gass and Mackey (2000).

**Pre-stimulated recall instructions**

Researcher:

*What we're going to do now is watch a replay of your reading and writing. I am interested in what you were thinking at the time you were reading and writing. I can see where you were looking and I can see what you were writing, but I don't know what you were thinking. So what I'd like you to do is tell me what you were thinking, what was in your mind at that time, while you were doing the task.*

*I'm going to put the mouse on the table here and you can pause the replay at any time if you want to tell me what you were thinking*.**(Demonstrate pausing the replay)** *If I have a question about what you were thinking then I will pause the replay and ask you to try and remember what you were thinking.*

The researcher will also pause the replay at various stages in a bid to elicit recall about specific cognitive processes.

| |
|---|
| **Stage of task / activity:** **Task Representation**<br><br>Pausing the replay just after participants have read the question / task rubric (or returned to re-read the question) and prompting using one of the following... |
| **It seems as though you had just read the question.**<br>**Can you tell me about that? / Can you recall what you were thinking at that time?/ Do you remember thinking anything after you had read the question? / Can you tell me what you thought when you had read the question?**<br><br>If the candidate is not forthcoming the researcher may further prompt...<br>'**At that point, what were you thinking about finding ideas to write about**?'<br>'**At that point did you think about what style to use?**'<br>'**At that point, did you think about how to organise your answer**?<br>'**At that point did you think about what to do next**'<br>**'At that point did you have any goals or targets?'** |

**<u>Stage of task / activity:</u> Macro-planning / Organising**

By pausing the replay and prompting just as participants <u>begin to write</u> their answer or during the writing of their plan and asking...

**It seems as though you are planning / just about to start writing your answer.**
**Can you tell me about that? / Can you recall what you were think at that time?/ Do you remember thinking anything about planning your answer? / Can you tell me what you thought when you began writing / planning your answer?**

If the candidate is not forthcoming the researcher may further prompt...

**'Did you make a plan for your answer before you started writing?'**
**'What decisions did you made when thinking about your plan?'**
**'Did your plan change while you were writing your answer? Why?'**
If no **'If you didn't make a plan how did you decide what to write?'**
**'When did you make decisions about what to write?'**
**'What helped you to make those decisions?'**

**<u>Stage of task / activity:</u> Reading / Connecting and Generating**

By pausing the replay when the <u>participant is engaged in reading and prompting</u>

**It seems as though you are reading there.**
**Can you tell me about that? / Can you recall what you were think at that time?/ Do you remember thinking anything as you read that sentence? / Can you tell me what you thought as you read?**

If the candidate is not forthcoming the researcher may further prompt...

**'You seem to be reading quickly/carefully/searching? Can you tell me about that'**
**'What decisions did you make while you were reading - What helped you decide?'**
**'How did you decide what to / how to read?'**

**<u>Stage of task / activity:</u> Micro-planning**

By pausing the replay when the <u>participant has paused writing (particularly mid-sentence pauses</u>) and prompting

| |
|---|
| **It seems as though you had paused when you were writing that sentence.**<br>**Can you tell me about that? / Can you recall what you were think at that time?/ Do you remember thinking anything when you paused? / Can you tell me what you thought when you paused?**<br><br>If the candidate is not forthcoming the researcher may further prompt...<br><br>**'Did you plan what to write next?'**<br>**'What decisions did you make when you paused? - What helped you decide?'**<br>**'How did you decide what to write next?'** |
| Stage of task / activity: **Monitoring and Revising**<br><br>By pausing the replay during a bout of editing and prompting |
| **It seems as though you are making some changes there.**<br>**Can you tell me about that? / Can you recall what you were think at that time?/ Do you remember thinking anything when you made those changes? / Can you tell me what you thought when you made those changes?**<br><br>If the candidate is not forthcoming the researcher may further prompt...<br><br>**'How did you decide what to change or when to make changes?'**<br>**'Can you tell me why you decided to make that change?'** |

The researcher will try to provide non-committal but encouraging responses such as

**Oh / mmm / great / good / I see / uh-huh / ok**

Gass and Mackey suggest delaying any follow-up questions about specific comments made by the participant until the end of the stimulated recall session, once the replay has finished.

## Appendix 8 Questionnaire

**Investigating cognitive-processing in academic reading and writing.**

The questionnaire has 43 questions. To answer the questions put a cross (X) in the box to the right of each question to select the appropriate answer. The list of possible answers includes : strongly agree, agree, neither agree nor disagree, disagree, strongly disagree. So for example, if you don't agree that the weather today is lovely you would put 'x' in the disagree box.

| | Strongly agree | Agree | Neither agree nor disagree | Disagree | Strongly disagree |
|---|---|---|---|---|---|
| Think about what the weather is like today and rate the following statement | | | | | |
| 1) I think the weather is really lovely | | | | ***X*** | |

There are no right or wrong answers. Each individual tackles their assignments in a different way and I am interested in how you approached your assignment so please try to be spontaneous and honest.

Your answers will be completely confidential and individual respondents' answers will not be published or shown to anyone else.

Thank you very much for taking the time to complete my questionnaire.

This questionnaire forms part of my PhD research into cognitive processing during academic reading into writing. Please contact me if you require any further information: Nicola Latimer, Centre for Research into English Language Learning and Assessment (CRELLA), University of Bedfordshire. Email: nicola.latimer@beds.ac.uk Tel:XXXXXXXXXX

| | | Strongly agree | Agree | Neither agree nor disagree | Disagree | Strongly disagree |
|---|---|---|---|---|---|---|
| Think about **what** the assignment question asked you to include and rate the following statements | | | | | | |
| 1) | I thought I needed to include a lot of my own opinions | | | | | |
| 2) | I thought I needed to include a lot of ideas from the source texts | | | | | |
| 3) | I thought I needed to include a lot of ideas / information from my memory | | | | | |
| Think about whether you used **any strategies** to help you decide what to include in your answer and rate the following statement. **Examples of strategies** : listing the ideas from the source texts one at a time, dividing the ideas from the source texts into groups (e.g. for / against), finding an organising idea to help arrange the information from the texts. | | | | | | |
| 4) | I used a strategy / strategies to help me decide what to include | | | | | |
| Think about whether you returned to **read the question more than once** and rate the following statements. | | | | | | |
| 5) | I returned to re-read the assignment question several times. | | | | | |
| 6) | My understanding of the assignment question changed whilst I was completing the assignment. | | | | | |
| Think about whether you had any goals or aims **in addition to those set by the assignment question.** You might have wanted to impress the reader by, for example, showing what you have learnt in lectures or using new words or phrases that you have read. Thinking about this, rate the following statement | | | | | | |
| 7) | I set myself additional goals, beyond what was required for the assignment | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Think about when you were **starting to plan** your writing. Rate the following statements about your planning | | | | | |
| 8) I formed a plan before starting to read or write | | | | | |
| 9) My plan was formed / changed as I was reading the source texts | | | | | |
| 10) My plan was formed / changed as I was writing my answer | | | | | |
| Think about **the style** (business-like / journalistic…) and **formality** of your writing and rate the following statements | | | | | |
| 11) I thought about how to adapt my writing to suit a particular style (business / academic / journalistic etc.) | | | | | |
| 12) I thought about how formal / informal to make my writing | | | | | |
| Think about whether you had a clear idea of **who would read** your writing and rate the following statements. | | | | | |
| 13) I had a clear idea of who I was writing to. | | | | | |
| 14) I adapted my writing to suit the reader. | | | | | |
| Think about how you decided **which** information to include in your answer and rate the following statements. | | | | | |
| 15) I asked myself whether the information was relevant to the assignment question. | | | | | |
| 16) I asked myself whether there was enough information to answer all parts of the assignment question. | | | | | |
| Think about how you **planned the overall organisation** of your answer and rate the following statements | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 17) | I thought I would follow the same organisation as one of the source texts | | | | | |
| 18) | I thought I would use the assignment question to help me organise my answer | | | | | |

Think about the **different types** of reading you used to read the source texts and rate the following statements

| | | | | | | |
|---|---|---|---|---|---|---|
| 19) | When I wanted to **completely understand** the meaning of an individual sentence I read it slowly and carefully. | | | | | |
| 20) | When I wanted a **deep understanding of the main ideas** in a source text I read the whole of that source text slowly and carefully. | | | | | |
| 21) | When I wanted **to find a particular word, date or specific detail** I quickly scanned through the text, skipping over some sections of the text until I found the sentence I needed. | | | | | |
| 22) | When I wanted to **get the gist (main idea)** of what the text was about I quickly sampled the text to find out what it was generally about. | | | | | |
| 23) | When I wanted to **locate the main ideas** I searched quickly and selectively through the text for them. | | | | | |

Think about when you were **reading the source texts** and rate the following statements

| | | | | | | |
|---|---|---|---|---|---|---|
| 24) | I decided **which ideas** were more important than others for this assignment. | | | | | |
| 25) | It was important **to link** new information to what I already knew. | | | | | |
| 26) | I worked out how **the main ideas** in each source text were linked together within that source text. | | | | | |
| 27) | I tried to understand **how** the main ideas in the different source texts related to each other. | | | | | |
| 8) | The source texts helped me decide what order to put the ideas in. | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 29) | I organised the ideas into an order I thought of myself. | | | | | |
| 30) | I had a main theme that helped me select ideas from the source texts and from my ideas. | | | | | |
| 31) | While writing I wrote about each text, one at a time. | | | | | |
| 32) | While writing I combined information from different texts to support the points I was making. | | | | | |
| 33) | I corrected spelling mistakes and typing errors. | | | | | |
| 34) | I corrected grammar mistakes. | | | | | |
| 35) | I made changes to the main ideas. | | | | | |
| 36) | I made changes to the order of the paragraphs / ideas. | | | | | |
| 37) | I made changes to ensure my ideas were clearly linked together. | | | | | |
| 38) | I made changes to the vocabulary I had used. | | | | | |
| 39) | I removed ideas that weren't relevant to the question. | | | | | |
| 40) | I checked the quotations were properly formatted and referenced. | | | | | |
| 41) | I checked that I had put the ideas from the source texts into my own words. | | | | | |
| 42) | I edited my work while I was writing. | | | | | |
| 43) | I started editing my work after I had finished most of the writing. | | | | | |

Thank you for your time : )

# Appendix 9 Transcript word counts

| Y1 | | Y3+ | |
|---|---|---|---|
| Participant | Word count | Participant | Word count |
| P11 | 1925 | P32 | 1536 |
| P12 | 1343 | P38 | 2140 |
| P14 | 1863 | P40 | 1999 |
| P15 | | P41 | 2378 |
| P23 | | P42 | 1866 |
| P25 | | P45 | 1698 |
| P26 | 1765 | P46 | 2654 |
| P28 | 1351 | P47 | 1843 |
| P29 | 1734 | P48 | 2041 |
| P31 | 2177 | P49 | 1748 |
| P34 | 1271 | P50 | 2137 |
| P35 | 1699 | P51 | 2016 |
| P36 | 160 | P52 | 1550 |
| P37 | 1758 | P53 | 1200 |
| P55 | 1006 | P54 | 1210 |
| Total | 18052 | Total | 28016 |
| Mean* | 1627 | Mean | 1868 |

*excluding P36 which ended prematurely

## Appendix 10 Fixation numbers and total duration by participant

| Participant | Count of fixations | Total duration (ms) | Mean fixation duration (ms) | Std Dev of fixation dur |
|---|---|---|---|---|
| P11 | 9,585 | 2,016,982 | 210 | 207 |
| P12 | 7,349 | 1,466,960 | 200 | 181 |
| P14 | 5,787 | 1,184,921 | 205 | 154 |
| P15 | 5,173 | 1,098,589 | 212 | 236 |
| P23 | 3,903 | 662,147 | 170 | 142 |
| P25 | 5,470 | 1,000,243 | 183 | 146 |
| P26 | 4,327 | 1,125,168 | 260 | 248 |
| P28 | 8,282 | 1,880,319 | 227 | 181 |
| P29 | 9,259 | 1,993,241 | 215 | 184 |
| P31 | 8,489 | 2,153,378 | 254 | 215 |
| P32 | 6,762 | 1,908,876 | 282 | 330 |
| P34 | 3,691 | 1,070,505 | 290 | 279 |
| P35 | 8,727 | 1,612,517 | 185 | 116 |
| P36 | 6,637 | 1,983,859 | 299 | 357 |
| P37 | 7,639 | 1,521,218 | 199 | 157 |
| P38 | 5,793 | 1,178,974 | 204 | 164 |
| P40 | 6,923 | 1,874,642 | 271 | 272 |
| P41 | 7,950 | 1,961,769 | 247 | 243 |
| P42 | 10,363 | 2,337,084 | 226 | 201 |
| P45 | 6,715 | 1,344,722 | 200 | 157 |
| P46 | 8,000 | 1,692,297 | 212 | 177 |
| P47 | 8,024 | 1,827,644 | 228 | 158 |
| P48 | 8,038 | 2,015,156 | 251 | 216 |
| P49 | 7,511 | 1,869,640 | 249 | 239 |
| P50 | 5,580 | 1,085,933 | 195 | 164 |
| P51 | 10,025 | 2,416,980 | 241 | 215 |
| P52 | 8,762 | 2,054,874 | 235 | 288 |
| P53 | 4,867 | 1,418,554 | 291 | 280 |
| P54 | 9,683 | 2,418,265 | 250 | 234 |
| P55 | 5,738 | 1,254,622 | 219 | 184 |
| Grand Total | 215,052 | 49,430,079 | | |
| Mean for all participants | 7,168 | 1,647,669 | | |
| Std dev for all participants | 1,808 | 459,659 | | |

# Appendix 11 Number of fixations per sentence of the written source text

| | Undergraduates | | | Postgraduates | | | All participants | | |
|---|---|---|---|---|---|---|---|---|---|
| **Number of fixations** | **Mean** | **SD** | **%** | **Mean** | **SD** | **%** | **Mean** | **SD** | **%** |
| Total Count | **40362** | **2,691** | 100% | **39,022** | **2,601** | 100% | 2,646 | 872 | 100% |
| P2TITLECount | 115 | 8 | 0% | 95 | 6 | 0% | 7 | 5 | 0% |
| P2S1Count | 993 | 66 | 2% | 904 | 60 | 2% | 63 | 35 | 2% |
| P2S2Count | 1857 | 124 | 5% | 1,658 | 111 | 4% | 117 | 50 | 4% |
| P2S3Count | 627 | 42 | 2% | 597 | 40 | 2% | 41 | 22 | 2% |
| P2S4Count | 2566 | 171 | 6% | 2,349 | 157 | 6% | 164 | 76 | 6% |
| P2S5Count | 2968 | 198 | 7% | 3,285 | 219 | 8% | 208 | 94 | 8% |
| P3TITLECount | 71 | 5 | 0% | 38 | 3 | 0% | 4 | 6 | 0% |
| P3S1Count | 1381 | 92 | 3% | 1,305 | 87 | 3% | 90 | 45 | 3% |
| P3S2Count | 1226 | 82 | 3% | 988 | 66 | 3% | 74 | 43 | 3% |
| P3S3Count | 2104 | 140 | 5% | 1,620 | 108 | 4% | 124 | 74 | 5% |
| P3S4Count | 577 | 38 | 1% | 469 | 31 | 1% | 35 | 21 | 1% |
| P3S5Count | 1984 | 132 | 5% | 1,746 | 116 | 4% | 124 | 66 | 5% |
| P3S6Count | 1499 | 100 | 4% | 1,511 | 101 | 4% | 100 | 54 | 4% |
| P3S7Count | 1237 | 82 | 3% | 1,361 | 91 | 3% | 87 | 61 | 3% |
| P3S8Count | 1605 | 107 | 4% | 1,724 | 115 | 4% | 111 | 78 | 4% |
| P3S9Count | 1081 | 72 | 3% | 1,049 | 70 | 3% | 71 | 49 | 3% |
| P4TITLECount | 182 | 12 | 0% | 157 | 10 | 0% | 11 | 8 | 0% |
| P4S1Count | 957 | 64 | 2% | 670 | 45 | 2% | 54 | 34 | 2% |
| P4S2Count | 1025 | 68 | 3% | 884 | 59 | 2% | 64 | 32 | 2% |
| P4S3Count | 1371 | 91 | 3% | 1,078 | 72 | 3% | 82 | 35 | 3% |
| P4S4Count | 2569 | 171 | 6% | 2,244 | 150 | 6% | 160 | 67 | 6% |
| P4S5Count | 2038 | 136 | 5% | 2,267 | 151 | 6% | 144 | 58 | 5% |
| P4S6Count | 1598 | 107 | 4% | 1,776 | 118 | 5% | 112 | 65 | 4% |
| P5TITLECount | 42 | 3 | 0% | 47 | 3 | 0% | 3 | 4 | 0% |
| P5S1Count | 1806 | 120 | 4% | 2,031 | 135 | 5% | 128 | 56 | 5% |
| P5S2Count | 2254 | 150 | 6% | 2,590 | 173 | 7% | 161 | 77 | 6% |
| P5S3Count | 717 | 48 | 2% | 681 | 45 | 2% | 47 | 24 | 2% |
| P5S4Count | 877 | 58 | 2% | 922 | 61 | 2% | 60 | 31 | 2% |
| P5S5Count | 286 | 19 | 1% | 351 | 23 | 1% | 21 | 10 | 1% |
| P5S6Count | 890 | 59 | 2% | 904 | 60 | 2% | 60 | 27 | 2% |
| P5S7Count | 999 | 67 | 2% | 958 | 64 | 2% | 65 | 30 | 2% |
| P5S8Count | 860 | 57 | 2% | 763 | 51 | 2% | 54 | 34 | 2% |

# Appendix 12 Total fixation duration of fixations on each sentence of the written source text

| | Undergraduates | | | Postgraduates | | | All participants | | |
|---|---|---|---|---|---|---|---|---|---|
| **Duration of fixations** | **Mean** | **SD** | **%** | **Mean** | **SD** | **%** | **Mean** | **SD** | **%** |
| Total Duration | **486,492** | **169,837** | **100%** | **487,772** | **168,540** | **100%** | **487,132** | **166,249** | **100%** |
| P2TITLE | 1,340 | 973 | 0% | 1,057 | 882 | 0% | 1,198 | 924 | 0% |
| P2S1 | 12,594 | 7,524 | 3% | 11,444 | 6,673 | 2% | 12,019 | 7,012 | 2% |
| P2S2 | 23,696 | 11,498 | 5% | 20,956 | 7,522 | 4% | 22,326 | 9,647 | 5% |
| P2S3 | 7,988 | 3,978 | 2% | 7,764 | 5,331 | 2% | 7,876 | 4,623 | 2% |
| P2S4 | 29,493 | 12,189 | 6% | 27,753 | 15,908 | 6% | 28,623 | 13,953 | 6% |
| P2S5 | 35,380 | 17,579 | 7% | 38,908 | 17,091 | 8% | 37,144 | 17,129 | 8% |
| P3TITLE | 757 | 1,348 | 0% | 363 | 435 | 0% | 560 | 1,005 | 0% |
| P3S1 | 16,956 | 8,033 | 3% | 17,365 | 9,931 | 4% | 17,160 | 8,877 | 4% |
| P3S2 | 15,364 | 9,821 | 3% | 12,640 | 6,128 | 3% | 14,002 | 8,162 | 3% |
| P3S3 | 27,235 | 17,982 | 6% | 21,446 | 9,031 | 4% | 24,341 | 14,288 | 5% |
| P3S4 | 6,603 | 3,435 | 1% | 5,815 | 3,831 | 1% | 6,209 | 3,598 | 1% |
| P3S5 | 24,300 | 10,271 | 5% | 22,220 | 13,890 | 5% | 23,260 | 12,049 | 5% |
| P3S6 | 17,720 | 9,665 | 4% | 19,428 | 11,011 | 4% | 18,574 | 10,217 | 4% |
| P3S7 | 14,418 | 8,770 | 3% | 17,662 | 15,626 | 4% | 16,040 | 12,559 | 3% |
| P3S8 | 18,445 | 12,710 | 4% | 20,516 | 16,830 | 4% | 19,481 | 14,691 | 4% |
| P3S9 | 12,112 | 9,013 | 2% | 12,723 | 9,465 | 3% | 12,418 | 9,086 | 3% |
| P4TITLE | 1,943 | 1,454 | 0% | 1,877 | 1,384 | 0% | 1,910 | 1,395 | 0% |
| P4S1 | 11,809 | 8,488 | 2% | 8,388 | 3,751 | 2% | 10,099 | 6,678 | 2% |
| P4S2 | 12,280 | 5,393 | 3% | 10,815 | 6,394 | 2% | 11,548 | 5,860 | 2% |
| P4S3 | 15,982 | 5,275 | 3% | 12,989 | 7,215 | 3% | 14,485 | 6,394 | 3% |
| P4S4 | 30,151 | 10,994 | 6% | 27,299 | 13,938 | 6% | 28,725 | 12,419 | 6% |
| P4S5 | 23,277 | 9,937 | 5% | 27,613 | 12,002 | 6% | 25,445 | 11,049 | 5% |
| P4S6 | 18,546 | 11,738 | 4% | 21,584 | 12,449 | 4% | 20,065 | 11,988 | 4% |
| P5TITLE | 443 | 457 | 0% | 574 | 886 | 0% | 508 | 696 | 0% |
| P5S1 | 22,943 | 10,765 | 5% | 27,037 | 14,894 | 6% | 24,990 | 12,937 | 5% |
| P5S2 | 29,586 | 15,886 | 6% | 34,905 | 17,809 | 7% | 32,246 | 16,801 | 7% |
| P5S3 | 8,899 | 5,240 | 2% | 8,405 | 3,790 | 2% | 8,652 | 4,500 | 2% |
| P5S4 | 10,504 | 5,456 | 2% | 11,417 | 6,061 | 2% | 10,960 | 5,685 | 2% |
| P5S5 | 3,346 | 1,537 | 1% | 4,358 | 2,177 | 1% | 3,852 | 1,922 | 1% |
| P5S6 | 10,626 | 4,127 | 2% | 11,461 | 6,598 | 2% | 11,044 | 5,424 | 2% |
| P5S7 | 11,561 | 4,207 | 2% | 11,572 | 7,370 | 2% | 11,567 | 5,896 | 2% |
| P5S8 | 10,195 | 7,606 | 2% | 9,417 | 6,702 | 2% | 9,806 | 7,055 | 2% |

## Appendix 13 Mean and SD of fix dur per sentence

| Area of interest | Mean fixation duration (ms) | SD of dur (ms) | | Area of interest | Mean fixation duration (ms) | SD of dur (ms) |
|---|---|---|---|---|---|---|
| P2 TITLE | 171 | 84 | | P4 TITLE | 169 | 80 |
| P2S1 | 190 | 101 | | P4S1 | 186 | 99 |
| P2S2 | 191 | 110 | | P4S2 | 181 | 95 |
| P2S3 | 193 | 107 | | P4S3 | 177 | 85 |
| P2S4 | 175 | 89 | | P4S4 | 179 | 101 |
| P2S5 | 178 | 95 | | P4S5 | 177 | 92 |
| P3 TITLE | 154 | 75 | | P4S6 | 178 | 99 |
| P3S1 | 192 | 101 | | P5 TITLE | 171 | 100 |
| P3S2 | 190 | 100 | | P5S1 | 195 | 103 |
| P3S3 | 196 | 112 | | P5S2 | 200 | 108 |
| P3S4 | 178 | 96 | | P5S3 | 186 | 92 |
| P3S5 | 187 | 103 | | P5S4 | 183 | 95 |
| P3S6 | 185 | 101 | | P5S5 | 181 | 93 |
| P3S7 | 185 | 107 | | P5S6 | 185 | 95 |
| P3S8 | 176 | 96 | | P5S7 | 177 | 93 |
| P3S9 | 175 | 91 | | P5S8 | 181 | 106 |

# Appendix 14 Y1 / Y3+ fix by sentence of the written source texts

| | | Year one undergraduates | | | | Year three undergraduates / postgraduates | | | |
|---|---|---|---|---|---|---|---|---|---|
| | No. of words per sentence | No of fixations | Total fixation duration (hh:mm:ss) | Fixations per word | Total duration per word (secs) | No of fixations | Total fixation duration (hh:mm:ss) | Fixations per word | Total duration per word (secs) |
| P2 TITLE | 4 | 115 | 00:00:20 | 29 | 5 | 95 | 00:00:16 | 24 | 4 |
| P2S1 | 12 | 993 | 00:03:09 | 83 | 16 | 904 | 00:02:52 | 75 | 14 |
| P2S2 | 20 | 1,857 | 00:05:55 | 93 | 18 | 1,658 | 00:05:14 | 83 | 16 |
| P2S3 | 11 | 627 | 00:02:00 | 57 | 11 | 597 | 00:01:56 | 54 | 11 |
| P2S4 | 22 | 2,566 | 00:07:22 | 117 | 20 | 2,349 | 00:06:56 | 107 | 19 |
| P2S5 | 39 | 2,968 | 00:08:51 | 76 | 14 | 3,285 | 00:09:44 | 84 | 15 |
| P3 TITLE | 5 | 71 | 00:00:11 | 14 | 2 | 38 | 00:00:05 | 8 | 1 |
| P3S1 | 15 | 1,381 | 00:04:14 | 92 | 17 | 1,305 | 00:04:20 | 87 | 17 |
| P3S2 | 12 | 1,226 | 00:03:50 | 102 | 19 | 988 | 00:03:10 | 82 | 16 |
| P3S3 | 29 | 2,104 | 00:06:49 | 73 | 14 | 1,620 | 00:05:22 | 56 | 11 |
| P3S4 | 8 | 577 | 00:01:39 | 72 | 12 | 469 | 00:01:27 | 59 | 11 |
| P3S5 | 20 | 1,984 | 00:06:04 | 99 | 18 | 1,746 | 00:05:33 | 87 | 17 |
| P3S6 | 16 | 1,499 | 00:04:26 | 94 | 17 | 1,511 | 00:04:51 | 94 | 18 |
| P3S7 | 19 | 1,237 | 00:03:36 | 65 | 11 | 1,361 | 00:04:25 | 72 | 14 |
| P3S8 | 31 | 1,605 | 00:04:37 | 52 | 9 | 1,724 | 00:05:08 | 56 | 10 |
| P3S9 | 31 | 1,081 | 00:03:02 | 35 | 6 | 1,049 | 00:03:11 | 34 | 6 |
| P4 TITLE | 3 | 182 | 00:00:29 | 61 | 10 | 157 | 00:00:28 | 52 | 9 |
| P4S1 | 14 | 957 | 00:02:57 | 68 | 13 | 670 | 00:02:06 | 48 | 9 |
| P4S2 | 15 | 1,025 | 00:03:04 | 68 | 12 | 884 | 00:02:42 | 59 | 11 |
| P4S3 | 14 | 1,371 | 00:04:00 | 98 | 17 | 1,078 | 00:03:15 | 77 | 14 |
| P4S4 | 34 | 2,569 | 00:07:32 | 76 | 13 | 2,244 | 00:06:49 | 66 | 12 |
| P4S5 | 29 | 2,038 | 00:05:49 | 70 | 12 | 2,267 | 00:06:54 | 78 | 14 |
| P4S6 | 29 | 1,598 | 00:04:38 | 55 | 10 | 1,776 | 00:05:24 | 61 | 11 |
| P5 TITLE | 4 | 42 | 00:00:07 | 11 | 2 | 47 | 00:00:09 | 12 | 2 |
| P5S1 | 30 | 1,806 | 00:05:44 | 60 | 11 | 2,031 | 00:06:46 | 68 | 14 |
| P5S2 | 28 | 2,254 | 00:07:24 | 81 | 16 | 2,590 | 00:08:44 | 93 | 19 |
| P5S3 | 14 | 717 | 00:02:13 | 51 | 10 | 681 | 00:02:06 | 49 | 9 |
| P5S4 | 17 | 877 | 00:02:38 | 52 | 9 | 922 | 00:02:51 | 54 | 10 |
| P5S5 | 5 | 286 | 00:00:50 | 57 | 10 | 351 | 00:01:05 | 70 | 13 |
| P5S6 | 21 | 890 | 00:02:39 | 42 | 8 | 904 | 00:02:52 | 43 | 8 |
| P5S7 | 26 | 999 | 00:02:53 | 38 | 7 | 958 | 00:02:54 | 37 | 7 |
| P5S8 | 22 | 860 | 00:02:33 | 39 | 7 | 763 | 00:02:21 | 35 | 6 |
| **Total** | **599** | **40,362** | **02:01:37** | | | **39,022** | **02:01:57** | | |

## Appendix 15 Y1 / Y3+ RTA comments tally

| | | | | | |
|---|---|---|---|---|---|
| **Total number of comments for Task Representation** | | **31** | **30** | **21%** | **14%** |
| Sub themes | Genre | 6 | 6 | 4% | 3% |
| | Goal setting | 8 | | 5% | 0% |
| | Recall of task | 6 | 8 | 4% | 4% |
| | Rhetorical function | 11 | 15 | 7% | 7% |
| | Time constraints | | 1 | 0% | 0% |
| **Total number of comments for Macro-planning** | | **18** | **26** | **12%** | **12%** |
| Sub themes | Finding content | 1 | 7 | 1% | 3% |
| | Structure | 17 | 19 | 11% | 9% |
| **Total number of comments for High level reading processes** | | **41** | **58** | **27%** | **27%** |
| Sub themes | Deeper understanding | 11 | 6 | 7% | 3% |
| | Extracting info | 8 | 4 | 5% | 2% |
| | Failure to understand | 5 | 14 | 3% | 6% |
| | Gist | 4 | 15 | 3% | 7% |
| | Lack of engagement | 4 | 4 | 3% | 2% |
| | Scanning | | 6 | 0% | 3% |
| | Search reading | 6 | 1 | 4% | 0% |
| | Targeted | 3 | 8 | 2% | 4% |
| **Total number of comments for Organising** | | **3** | **6** | **2%** | **3%** |
| Sub theme | Categorise | 3 | 6 | 2% | 3% |
| **Total number of comments for Connecting and generating** | | **12** | **28** | **8%** | **13%** |
| Sub themes | Familiarisation | | 3 | 0% | 1% |
| | Forming opinions | | 6 | 0% | 3% |
| | Generating new ideas | | 4 | 0% | 2% |
| | Makes links between text ideas | 6 | 10 | 4% | 5% |
| | Own experience | 6 | 5 | 4% | 2% |
| **Total number of comments for Micro planning & translating** | | **10** | **9** | **7%** | **4%** |
| Sub themes | Prompting | 2 | 2 | 1% | 1% |
| | Reflecting | 3 | 2 | 2% | 1% |
| | Rewording | 5 | 5 | 3% | 2% |
| **Total number of comments for High level monitoring and revising** | | **20** | **39** | **13%** | **18%** |
| Sub themes | Amount of info | 4 | 2 | 3% | 1% |
| | Coherence / Cohesion | 4 | 19 | 3% | 9% |
| | Improving expression | 12 | 18 | 8% | 8% |
| **Total number of comments for Low level monitoring & revising** | | **15** | **21** | **10%** | **10%** |
| Sub themes | General | 3 | 3 | 2% | 1% |
| | Grammar | 3 | 6 | 2% | 3% |
| | Plagiarism | 3 | 3 | 2% | 1% |
| | Spelling | 5 | 3 | 3% | 1% |
| | Vocab | 1 | 6 | 1% | 3% |

# Appendix 16 Y1 / Y3+ questionnaire scores

| | | Y1 | | | | | | Y3+ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Question** | **Task Representation** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** |
| **1** | I thought I needed to include a lot of my own opinions | 0 | 3 | 3 | 8 | 1 | **2.53** | 1 | 1 | 6 | 6 | 1 | **2.67** |
| **2** | I thought I needed to include a lot of ideas from the source texts | 5 | 10 | 0 | 0 | 0 | **4.33** | 8 | 7 | 0 | 0 | 0 | **4.53** |
| **3** | I thought I needed to include a lot of ideas / information from my memory | 1 | 1 | 3 | 7 | 3 | **2.33** | 0 | 2 | 2 | 6 | 5 | **2.07** |
| **4** | I used a strategy / strategies to help me decide what to include | 7 | 6 | 0 | 1 | 1 | **4.13** | 5 | 10 | 0 | 0 | 0 | **4.33** |
| **5** | I returned to re-read the assignment question several times. | 10 | 4 | 0 | 0 | 1 | **4.47** | 10 | 4 | 0 | 1 | 0 | **4.53** |
| **6** | My understanding of the assignment question changed whilst I was completing the assignment. | 2 | 4 | 2 | 5 | 2 | **2.93** | 3 | 3 | 2 | 7 | 0 | **3.13** |
| **7** | I set myself additional goals, beyond what was required for the assignment | 1 | 4 | 7 | 2 | 1 | **3.13** | 1 | 2 | 6 | 6 | 0 | **2.87** |
| | **Macro-planning** | | | | | | | | | | | | |
| **8** | I formed a plan before starting to read or write | 2 | 6 | 3 | 3 | 1 | **3.33** | 3 | 7 | 1 | 4 | 0 | **3.60** |
| **9** | My plan was formed / changed as I was reading the source texts | 3 | 6 | 2 | 4 | 0 | **3.53** | 3 | 6 | 2 | 4 | 0 | **3.53** |
| **10** | My plan was formed / changed as I was writing my answer | 1 | 7 | 3 | 2 | 2 | **3.20** | 3 | 5 | 2 | 5 | 0 | **3.40** |
| **11** | I thought about how to adapt my writing to suit a particular style (business / academic / journalistic etc.) | 3 | 7 | 3 | 2 | 0 | **3.73** | 3 | 8 | 2 | 2 | 0 | **3.80** |
| **12** | I thought about how formal / informal to make my writing | 4 | 10 | 1 | 0 | 0 | **4.20** | 5 | 8 | 1 | 1 | 0 | **4.13** |
| **13** | I had a clear idea of who I was writing to. | 2 | 5 | 4 | 3 | 1 | **3.27** | 4 | 6 | 4 | 1 | 0 | **3.87** |
| **14** | I adapted my writing to suit the reader. | 1 | 8 | 3 | 2 | 1 | **3.40** | 1 | 4 | 5 | 4 | 1 | **3.00** |
| **15** | I asked myself whether the information was relevant to the assignment question. | 6 | 8 | 0 | 1 | 0 | **4.27** | 9 | 6 | 0 | 0 | 0 | **4.60** |
| **16** | I asked myself whether there was enough information to answer all parts of the assignment question. | 5 | 4 | 3 | 3 | 0 | **3.73** | 5 | 6 | 1 | 3 | 0 | **3.87** |
| **17** | I thought I would follow the same organisation as one of the source texts | 1 | 3 | 2 | 8 | 1 | **2.67** | 1 | 1 | 2 | 9 | 2 | **2.33** |
| **18** | I thought I would use the assignment question to help me organise my answer | 7 | 6 | 0 | 1 | 1 | **4.13** | 10 | 3 | 1 | 1 | 0 | **4.47** |

| | High-level reading including connecting and generating | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | Mean | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **19** | When I wanted to **completely understand** the meaning of an individual sentence I read it slowly and carefully. | 8 | 7 | 0 | 0 | 0 | **4.53** | 6 | 9 | 0 | 0 | 0 | **4.40** |
| **20** | When I wanted a **deep understanding of the main ideas** in a source text I read the whole of that source text slowly and carefully. | 5 | 4 | 4 | 2 | 0 | **3.80** | 4 | 10 | 0 | 1 | 0 | **4.13** |
| **21** | When I wanted **to find a particular word, date or specific detail** I quickly scanned through the text, skipping over some sections of the text until I found the sentence I needed. | 10 | 5 | 0 | 0 | 0 | **4.67** | 13 | 2 | 0 | 0 | 0 | **4.87** |
| **22** | When I wanted to **get the gist (main idea)** of what the text was about I quickly sampled the text to find out what it was generally about. | 8 | 7 | 0 | 0 | 0 | **4.53** | 8 | 3 | 2 | 0 | 2 | **4.00** |
| **23** | When I wanted to **locate the main ideas** I searched quickly and selectively through the text for them. | 6 | 6 | 3 | 0 | 0 | **4.20** | 7 | 6 | 0 | 2 | 0 | **4.20** |
| **24** | I decided **which ideas** were more important than others for this assignment. | 4 | 10 | 1 | 0 | 0 | **4.20** | 7 | 6 | 1 | 1 | 0 | **4.27** |
| **25** | It was important **to link** new information to what I already knew. | 1 | 4 | 6 | 3 | 1 | **3.07** | 1 | 5 | 5 | 3 | 1 | **3.13** |
| **26** | I worked out how **the main ideas** in each source text were linked together within that source text. | 4 | 9 | 1 | 1 | 0 | **4.07** | 2 | 9 | 4 | 0 | 0 | **3.87** |
| **27** | I tried to understand **how** the main ideas in the different source texts related to each other. | 7 | 7 | 1 | 0 | 0 | **4.40** | 6 | 4 | 5 | 0 | 0 | **4.07** |
| | **Organising** | | | | | | | | | | | | |
| **28** | The source texts helped me decide what order to put the ideas in. | 3 | 5 | 2 | 5 | 0 | **3.40** | 6 | 4 | 2 | 3 | 0 | **3.87** |
| **29** | I organised the ideas into an order I thought of myself. | 3 | 6 | 3 | 1 | 2 | **3.47** | 1 | 5 | 3 | 4 | 2 | **2.93** |
| **30** | I had a main theme that helped me select ideas from the source texts and from my ideas. | 3 | 9 | 3 | 0 | 0 | **4.00** | 2 | 11 | 0 | 2 | 0 | **3.87** |
| **31** | While writing I wrote about each text, one at a time. | 2 | 5 | 1 | 6 | 1 | **3.07** | 3 | 5 | 3 | 4 | 0 | **3.47** |
| **32** | While writing I combined information from different texts to support the points I was making. | 2 | 7 | 2 | 3 | 1 | **3.40** | 3 | 8 | 2 | 2 | 0 | **3.80** |

| | **Monitoring and revising** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** | Strongly agree | Agree | Ambivalent | Disagree | Strongly disagree | **Mean** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **33** | I corrected spelling mistakes and typing errors. | 12 | 3 | 0 | 0 | 0 | **4.80** | 8 | 6 | 1 | 0 | 0 | **4.47** |
| **34** | I corrected grammar mistakes. | 9 | 5 | 1 | 0 | 0 | **4.53** | 8 | 7 | 0 | 0 | 0 | **4.53** |
| **35** | I made changes to the main ideas. | 4 | 2 | 3 | 4 | 2 | **3.13** | 1 | 3 | 2 | 8 | 1 | **2.67** |
| **36** | I made changes to the order of the paragraphs / ideas. | 4 | 3 | 1 | 3 | 4 | **3.00** | 3 | 3 | 0 | 6 | 3 | **2.80** |
| **37** | I made changes to ensure my ideas were clearly linked together. | 6 | 5 | 2 | 2 | 0 | **4.00** | 4 | 7 | 3 | 1 | 0 | **3.93** |
| **38** | I made changes to the vocabulary I had used. | 7 | 6 | 1 | 1 | 0 | **4.27** | 7 | 6 | 1 | 1 | 0 | **4.27** |
| **39** | I removed ideas that weren't relevant to the question. | 4 | 3 | 3 | 3 | 2 | **3.27** | 2 | 4 | 4 | 5 | 0 | **3.20** |
| **40** | I checked the quotations were properly formatted and referenced. | 2 | 2 | 4 | 1 | 6 | **2.53** | 0 | 4 | 5 | 2 | 4 | **2.60** |
| **41** | I checked that I had put the ideas from the source texts into my own words. | 9 | 5 | 0 | 0 | 1 | **4.40** | 3 | 8 | 2 | 1 | 1 | **3.73** |
| **42** | I edited my work while I was writing. | 5 | 4 | 3 | 2 | 1 | **3.67** | 6 | 5 | 0 | 3 | 1 | **3.80** |
| **43** | I started editing my work after I had finished most of the writing. | 5 | 6 | 0 | 3 | 1 | **3.73** | 5 | 3 | 2 | 4 | 1 | **3.47** |

# Appendix 17 Y1, Y3+, low scorers and high scorers: comparison of fixation duration across AOIs

| Percentage of fixation duration | | | | |
|---|---|---|---|---|
| | Y1 | Y3+ | Low scorers | High scorers |
| Total | 100% | 100% | 100% | 100% |
| Fixations on Own Work | 53% | 59% | 59% | 56% |
| Fixations on Move Page buttons | 2% | 2% | 2% | 3% |
| Fixation on Screen Instructions | 0% | 0% | 0% | 0% |
| Fixations on task instructions | 5% | 5% | 4% | 4% |
| Fixations on written source texts | 33% | 27% | 29% | 30% |
| Fixations on diagrams | 3% | 3% | 2% | 4% |
| Fix outside any area of interest | 4% | 4% | 4% | 4% |

## Appendix 18 Low / high scorers' number of fixations per sentence

| | % of source text length | % of low scorers fixations | % of high scorers fixations | Low scorers no. fixations | High scorers no. fixations |
|---|---|---|---|---|---|
| P2 TITLE | 0.67% | 0.26% | 0.25% | 27 | 34 |
| P2S1 | 2.00% | 2.58% | 2.30% | 272 | 314 |
| P2S2 | 3.34% | 4.32% | 4.06% | 456 | 553 |
| P2S3 | 1.84% | 1.52% | 1.12% | 160 | 152 |
| P2S4 | 3.67% | 3.88% | 5.75% | 409 | 783 |
| P2S5 | 6.51% | 6.17% | 7.93% | 651 | 1,081 |
| P3 TITLE | 0.83% | 0.13% | 0.04% | 14 | 6 |
| P3S1 | 2.50% | 3.71% | 3.47% | 391 | 473 |
| P3S2 | 2.00% | 2.42% | 3.04% | 255 | 414 |
| P3S3 | 4.84% | 4.30% | 4.79% | 454 | 653 |
| P3S4 | 1.34% | 0.92% | 0.84% | 97 | 115 |
| P3S5 | 3.34% | 3.87% | 4.29% | 408 | 584 |
| P3S6 | 2.67% | 2.52% | 4.15% | 266 | 565 |
| P3S7 | 3.17% | 2.95% | 3.56% | 311 | 485 |
| P3S8 | 5.18% | 5.25% | 3.78% | 554 | 515 |
| P3S9 | 5.18% | 2.84% | 3.18% | 300 | 434 |
| P4 TITLE | 0.50% | 0.45% | 0.34% | 48 | 46 |
| P4S1 | 2.34% | 2.29% | 2.67% | 242 | 364 |
| P4S2 | 2.50% | 2.52% | 2.91% | 266 | 397 |
| P4S3 | 2.34% | 3.26% | 2.81% | 344 | 383 |
| P4S4 | 5.68% | 5.49% | 6.51% | 579 | 887 |
| P4S5 | 4.84% | 6.56% | 5.86% | 692 | 799 |
| P4S6 | 4.84% | 4.82% | 4.65% | 508 | 634 |
| P5 TITLE | 0.67% | 0.21% | 0.07% | 22 | 10 |
| P5S1 | 5.01% | 4.54% | 5.61% | 479 | 764 |
| P5S2 | 4.67% | 6.67% | 5.91% | 704 | 805 |
| P5S3 | 2.34% | 2.06% | 1.89% | 217 | 257 |
| P5S4 | 2.84% | 3.58% | 1.80% | 378 | 245 |
| P5S5 | 0.83% | 0.96% | 0.70% | 101 | 95 |
| P5S6 | 3.51% | 2.87% | 1.56% | 303 | 212 |
| P5S7 | 4.34% | 3.05% | 2.07% | 322 | 282 |
| P5S8 | 3.67% | 3.03% | 2.10% | 320 | 286 |
| | 100% | 100% | 100% | 10,550 | 13,627 |

# Appendix 19 Low / high scorers' total duration per sentence

| | % of source text length | % of low scorers cum. duration | % of high scorers cum. duration | Low scorers duration | High scorers duration |
|---|---|---|---|---|---|
| P2 TITLE | 0.67% | 0.25% | 0.21% | 5,399 | 5,288 |
| P2S1 | 2.00% | 2.60% | 2.31% | 56,390 | 56,774 |
| P2S2 | 3.34% | 4.26% | 4.00% | 92,383 | 98,587 |
| P2S3 | 1.84% | 1.49% | 1.10% | 32,347 | 26,987 |
| P2S4 | 3.67% | 3.64% | 5.33% | 79,014 | 131,244 |
| P2S5 | 6.51% | 6.21% | 7.73% | 134,705 | 190,427 |
| P3 TITLE | 0.83% | 0.08% | 0.04% | 1,810 | 901 |
| P3S1 | 2.50% | 4.10% | 3.81% | 88,843 | 93,864 |
| P3S2 | 2.00% | 2.57% | 3.12% | 55,759 | 76,846 |
| P3S3 | 4.84% | 4.73% | 5.21% | 102,679 | 128,312 |
| P3S4 | 1.34% | 0.95% | 0.82% | 20,583 | 20,299 |
| P3S5 | 3.34% | 3.97% | 4.55% | 86,087 | 112,034 |
| P3S6 | 2.67% | 2.58% | 4.36% | 55,974 | 107,498 |
| P3S7 | 3.17% | 2.75% | 3.57% | 59,561 | 87,856 |
| P3S8 | 5.18% | 4.62% | 3.67% | 100,109 | 90,324 |
| P3S9 | 5.18% | 2.53% | 3.27% | 54,957 | 80,544 |
| P4 TITLE | 0.50% | 0.41% | 0.31% | 8,992 | 7,668 |
| P4S1 | 2.34% | 2.36% | 2.93% | 51,270 | 72,244 |
| P4S2 | 2.50% | 2.42% | 2.85% | 52,510 | 70,258 |
| P4S3 | 2.34% | 3.16% | 2.66% | 68,540 | 65,507 |
| P4S4 | 5.68% | 5.30% | 6.17% | 114,891 | 151,969 |
| P4S5 | 4.84% | 6.17% | 5.56% | 133,703 | 137,036 |
| P4S6 | 4.84% | 4.85% | 4.21% | 105,099 | 103,744 |
| P5 TITLE | 0.67% | 0.20% | 0.06% | 4,238 | 1,436 |
| P5S1 | 5.01% | 5.21% | 6.08% | 113,044 | 149,783 |
| P5S2 | 4.67% | 7.68% | 6.08% | 166,459 | 149,682 |
| P5S3 | 2.34% | 1.90% | 1.87% | 41,136 | 46,000 |
| P5S4 | 2.84% | 3.35% | 1.76% | 72,617 | 43,249 |
| P5S5 | 0.83% | 0.95% | 0.71% | 20,607 | 17,449 |
| P5S6 | 3.51% | 2.85% | 1.61% | 61,774 | 39,561 |
| P5S7 | 4.34% | 2.98% | 1.97% | 64,732 | 48,603 |
| P5S8 | 3.67% | 2.88% | 2.07% | 62,506 | 51,029 |
| | 100% | 100% | 100% | 2,168,718 | 2,463,003 |

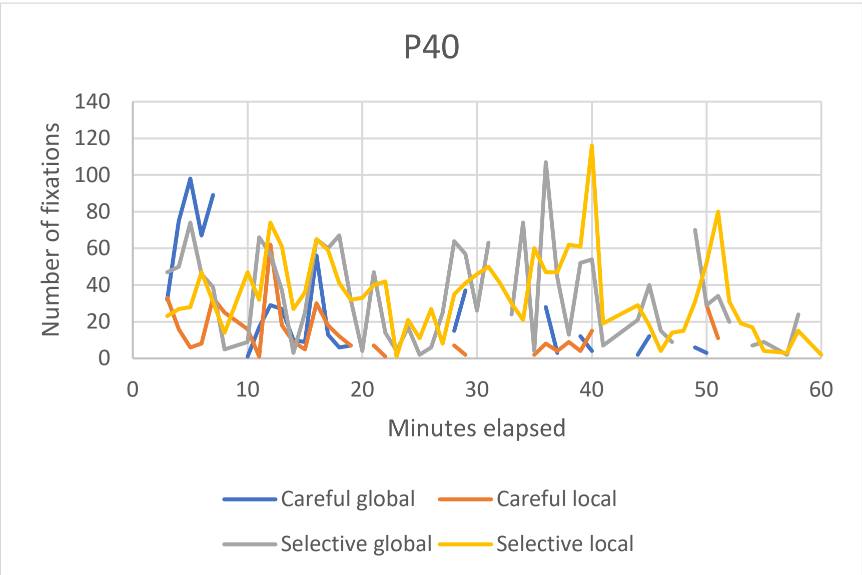
# Appendix 20 RTA tally of comments for low / high scorers

| | | | |
|---|---|---|---|
| **Total number of comments for Task Representation** | | **11** | **8** |
| Sub themes | Genre | 2 | 1 |
| | Goal setting | 1 | 1 |
| | Recall of task | 2 | 1 |
| | Rhetorical function | 6 | 5 |
| | Time constraints | | |
| **Total number of comments for Macro-planning** | | **5** | **7** |
| Sub themes | Finding content | 1 | 2 |
| | Structure | 4 | 5 |
| **Total number of comments for High level reading processes** | | 11 | **13** |
| Sub themes | Deeper understanding | 1 | 2 |
| | Extracting info | 1 | 2 |
| | Failure to understand | 1 | 4 |
| | Gist | 2 | 2 |
| | Lack of engagement | 1 | 0 |
| | Scanning | 0 | 2 |
| | Search reading | 3 | 0 |
| | Targeted | 2 | 1 |
| **Total number of comments for Organising** | | 0 | 2 |
| Sub theme | Categorise | 0 | 2 |
| **Total number of comments for Connecting and generating** | | **9** | **7** |
| Sub themes | Familiarisation | 1 | 0 |
| | Forming opinions | 1 | 1 |
| | Generating new ideas | 1 | 0 |
| | Makes links between text ideas | 3 | 6 |
| | Own experience | 3 | 0 |
| **Total number of comments for Micro planning & translating** | | **6** | **2** |
| Sub themes | Prompting | 1 | 1 |
| | Reflecting | 1 | 1 |
| | Rewording | 4 | |
| **Total number of comments for High level monitoring and revising** | | **8** | **13** |
| Sub themes | Amount of info | 2 | |
| | Coherence / Cohesion | 2 | 4 |
| | Improving expression | 4 | 9 |
| **Total number of comments for Low level monitoring & revising** | | **6** | **7** |
| Sub themes | General | 1 | 0 |
| | Grammar | 2 | 3 |
| | Plagiarism | 1 | 2 |
| | Spelling | 2 | 1 |
| | Vocab | 0 | 1 |

# Appendix 21 Questionnaire mean scores for low / high scorers

| Question | Mean | | Question | Mean | |
|---|---|---|---|---|---|
| | Low scorers | High scorers | | Low scorers | High scorers |
| Q1TR | 3.40 | 3.48 | Q23REA | 5.20 | 4.84 |
| Q2TR | 5.20 | 5.24 | Q24C&G | 5.20 | 5.64 |
| Q3TR | 3.40 | 1.88 | Q25C&G | 3.80 | 3.56 |
| Q4TR | 4.60 | 5.32 | Q26C&G | 4.60 | 4.52 |
| Q5TR | 5.20 | 5.84 | Q27C&G | 4.80 | 5.16 |
| Q6TR | 3.60 | 4.32 | Q28ORG | 3.60 | 4.92 |
| Q7TR | 4.40 | 3.88 | Q29ORG | 4.80 | 3.76 |
| Q8MP | 3.60 | 4.32 | Q30ORG | 4.80 | 4.76 |
| Q9MP | 3.80 | 4.56 | Q31ORG | 4.20 | 4.24 |
| Q10MP | 4.00 | 4.00 | Q32ORG | 4.40 | 5.08 |
| Q11MP | 4.40 | 4.08 | Q33LMR | 5.60 | 5.52 |
| Q12MP | 5.20 | 5.24 | Q34LMR | 5.00 | 5.60 |
| Q13MP | 4.80 | 4.56 | Q35HMR | 3.40 | 3.68 |
| Q14MP | 4.00 | 4.20 | Q36HMR | 4.00 | 3.40 |
| Q15MP | 5.20 | 5.84 | Q37HMR | 5.00 | 5.00 |
| Q16MP | 4.60 | 4.92 | Q38LMR | 5.20 | 5.64 |
| Q17MP | 2.80 | 3.16 | Q39HMR | 4.00 | 4.00 |
| Q18MP | 4.40 | 5.08 | Q40LMR | 3.20 | 3.04 |
| Q19REA | 5.00 | 5.80 | Q41LMR | 5.00 | 5.20 |
| Q20REA | 4.80 | 5.16 | Q42LMR | 4.60 | 4.52 |
| Q21REA | 5.60 | 5.72 | Q43HMR | 3.80 | 4.96 |
| Q22REA | 5.40 | 4.08 | | | |

# Appendix 22 Example of reading types as P40 progressed through task



# Appendix 23 Written source texts relevance scores

| TEXT 1 | Sentence reference | Total Relevance score | Relevance category |
|---|---|---|---|
| [redacted] | P2S1 | 11 | Highly relevant |
| [redacted] | P2S2 | 8 | Relevant |
| [redacted] | P2S3 | 7 | Less relevant |
| [redacted] | P2S4 | 12 | Highly relevant |
| [redacted] | P2S5 | 9 | Relevant |
| [redacted] | P3S1 | 8 | Relevant |
| [redacted] | P3S2 | 8 | Relevant |
| [redacted] | P3S3 | 7 | Less relevant |
| [redacted] | P3S4 | 11 | Highly relevant |
| [redacted] | P3S5 | 12 | Highly relevant |
| [redacted] | P3S6 | 12 | Highly relevant |
| [redacted] | P3S7 | 7 | Less relevant |
| [redacted] | P3S8 | 9 | Relevant |
| [redacted] | P3S9 | 9 | Relevant |

| TEXT 2 | Area of interest | Total Relevance score | Relevance category |
|---|---|---|---|
| [redacted] | P4S1 | 7 | Less relevant |
| [redacted] | P4S2 | 9 | Relevant |
| [redacted] | P4S3 | 7 | Less relevant |
| [redacted] | P4S4 | 12 | Highly relevant |
| [redacted] | P4S5 | 8 | Relevant |
| [redacted] | P4S6 | 8 | Relevant |
| [redacted] | P5S1 | 7 | Less relevant |
| [redacted] | P5S2 | 12 | Highly relevant |
| [redacted] | P5S3 | 9 | Relevant |
| [redacted] | P5S4 | 9 | Relevant |
| [redacted] | P5S5 | 6 | Less relevant |
| [redacted] | P5S6 | 6 | Less relevant |
| [redacted] | P5S7 | 6 | Less relevant |
| [redacted] | P5S8 | 7 | Less relevant |

## Appendix 24 Participant scores on the task according to level and place of study

| | Y1 undergraduates | | Year three undergrads / postgrads | |
|---|---|---|---|---|
| **Place of study** | **No. of participants** | **Average score* on task** | **No. of participants** | **Average score* on task** |
| Uni of Bedfordshire | 7 | 7.36 | 2 | 8.50 |
| Other universities | 8 | 7.44 | 13 | 8.12 |

* full score for the task = 9

## Appendix 25 Weighted gaze sample figures

| Participant | Weighted gaze samples | Participant | Weighted gaze samples |
|---|---|---|---|
| P11 | 79% | P38 | 48% |
| P12 | 47% | P40 | 67% |
| P14 | 40% | P41 | 72% |
| P15 | 44% | P42 | 71% |
| P23 | 37% | P45 | 45% |
| P25 | 45% | P46 | 68% |
| P26 | 69% | P47 | 54% |
| P28 | 61% | P48 | 67% |
| P29 | 69% | P49 | 70% |
| P31 | 66% | P50 | 49% |
| P32 | 63% | P51 | 65% |
| P34 | 55% | P52 | 67% |
| P35 | 52% | P53 | 74% |
| P36 | 80% | P54 | 68% |
| P37 | 64% | P55 | 55% |