Readers' cognitive processes during IELTS reading tests: evidence from eye tracking

Other

BRITISH COUNCIL

TeachingEnglish

ELT Research Papers 13–06

Readers' cognitive processes during IELTS reading tests: evidence from eye tracking

Stephen Bax, University of Bedfordshire

ELT Research Papers 13–06

Readers' cognitive processes during IELTS reading tests: evidence from eye tracking

Stephen Bax, University of Bedfordshire

ISBN 978-0-86355-720-0

10 Spring Gardens
London SW1A 2BN, UK

www.britishcouncil.org

Contents

Abstract

The research described in this report investigates readers' mental processes as they complete onscreen IELTS (International English Language Testing System) reading test items. It employs up-to-date eye tracking technology to research readers' eye movements and aims, among other things, to contribute to an understanding of the cognitive validity of reading test items (Glaser, 1991; Field forthcoming).

Participants were a group of Malaysian undergraduates (n=71) taking an onscreen test consisting of two IELTS reading passages with a total of 11 test items. The eye movements of a random sample of these participants (n=38) were tracked. Questionnaire and stimulated recall interview data were also collected, and were important in order to interpret and explain the eye tracking data.

Findings demonstrated significant differences between successful and unsuccessful test-takers on a number of dimensions, including their ability to read expeditiously (Khalifa and Weir, 2009), and their focus on particular aspects of the test items and the reading texts. This demonstrates the potential of eye tracking, in combination with post-hoc interview and questionnaire data, to offer new insights into the cognitive processes of successful and unsuccessful candidates in a reading test. It also gives unprecedented insights into the cognitive processing of successful and unsuccessful readers doing language tests.

As a consequence, the findings should be of value to teachers and learners, and also to examination boards seeking to validate and prepare reading tests, as well as psycholinguists and others interested in the cognitive processes of readers.

1

Introduction

When we prepare reading tests, it is important to ensure that our tests are valid, and part of a test's validity involves ensuring that the mental processes which test-takers use as they respond are similar to and representative of the mental processes they would use in the target situation in real life – what is known as cognitive validity (Glaser, 1991; Field, 2012). In the case of tests of academic reading ability, such as IELTS, this means that our tests should demonstrate that they assess candidates on the same range and types of cognitive operations as those required of students in programmes of higher academic study. If they do not, then the test cannot be considered cognitively valid.

This report describes a research project investigating onscreen IELTS reading test items in terms of candidates' cognitive processing, in ways that can potentially assist in understanding and improving the cognitive validity of these and similar reading test items. The research project reported here draws on established methods for investigating these areas, including interviews in which candidates report on the processes they have employed, but in addition it makes innovative use of eye tracking technology in order to seek new insights into candidates' cognitive operations.

Eye tracking technology has been used extensively for several decades to investigate various forms of reading. However, since its use to research the particular type of reading employed during language tests is rare, one contribution that this research seeks to make is a methodological one, investigating the extent to which eye tracking technology can offer new understanding of test-takers' reading behaviour. A second potential contribution, as suggested above, is to use eye tracking to help assess the cognitive validity of elements of the IELTS reading paper. In terms of application, the research reported here can also assist the process of test design in future, and can in addition inform language teachers and students seeking to prepare for reading tests such as the IELTS reading component.

2

Cognitive validity in reading tests

It has long been argued that language tests should exhibit what is known as cognitive validity (Glaser, 1991; Baxter and Glaser,1998), since cognitive interpretive claims are 'not foregone conclusions, [but] need to be warranted conceptually and empirically' (Ruiz-Primo et al. 2001: 100). As Alderson (2000: 97) argued '[t]he validity of a test relates to the interpretation of the correct responses to items, so what matters is not what the test constructors believe an item to be testing, but which responses are considered correct, and what process underlies them.' In short, understanding of the trait being measured requires an insight into the cognitive processing required for completion of the task. Field has recently summarised the situation as follows: 'we need to find out if the *mental processes* that a test elicits from a candidate resemble the processes that he/she would employ in non-test conditions' (Field, 2012:140, emphasis in original). Field (forthcoming) sets out three central questions that a language test must deal with in terms of its cognitive validity, namely:

a. Similarity of processing. Are the processes adopted during a test sufficiently similar to those that would be employed in the target context? Or do candidates adopt additional processes that are a by-product of facets of the test (procedure, test method, item) rather than part of the normal operations associated with the construct being tested?

b. Comprehensiveness. Do the items in the test elicit only a small subset of the cognitive processes that a language user would employ in a natural context? Or do they tap into a sufficiently broad range of such processes for the test to be deemed representative of real-world behaviour?

c. Calibration. Across a suite of tests graded by reference to a scale, are the cognitive demands imposed upon test takers at each level appropriately calibrated in relation to the performance features that might be expected of a listener at these levels?

(Field, forthcoming)

These considerations, which reinterpreted Messick's (1989) notions of construct-irrelevant variance and construct under-representation from a cognitive processing perspective, imply the need for better understanding of candidates' cognitive processing in language tests, and underpin the research to be described in this report in ways to be further detailed below.

3

Modelling cognitive processing in reading

A recent analysis of the reading process, which not only addresses reading under test conditions but also pays close attention to the role of readers' cognitive operations, is the model set out by Khalifa and Weir (2009). This model derives in turn from work by Urquhart and Weir, which characterises reading as taking place at the local or global level, and being in nature either careful or expeditious (i.e. involving 'quick, selective and efficient reading to access desired information in a text': Khalifa and Weir, 2009). Khalifa and Weir's model describes cognitive processing in reading in terms of different levels of complexity, with, for example, lexical processing as the least complex, and intertextual reading as the most. Khalifa and Weir's model is therefore particularly valuable in that it operationalises the concept of cognitive processing in reading, and proposes in a way amenable to empirical investigation a hierarchy of cognitive processing complexity in reading. This hierarchy is summarised in simple form in Table 1.

Table 1: Levels of cognitive processing in reading tests (adapted from Khalifa and Weir, 2009)

	Level of activity (ordered from more simple to more complex)	Readers' typical cognitive operations in language tests	Size of typical unit
1	Lexis: word matching	Reader identifies same word in question and text	Word
2	Lexis: synonym and word-class matching	Reader uses knowledge of word meaning or word class to identify synonym, antonym or other related word	Word
3	Grammar/syntax	Reader uses grammatical knowledge to disambiguate and identify answer	Clause/sentence
4	Propositional meaning	Reader uses knowledge of lexis and grammar to establish meaning of a sentence.	Sentence
5	Inference	Reader goes beyond literal meaning to infer a further significance	Sentence/ paragraph/text
6	Building a mental model	Reader uses several features of the text to build a larger mental model	Text
7	Understanding text function	Reader uses genre knowledge to identify text structure and purpose	Text

In terms of language testing, the authors then argue that for a high-level academic reading test to be cognitively valid it should test the full range of lower and higher cognitive processes. Research by Weir, Hawkey, Green and Devi (2009) then used this model to investigate the cognitive processes underlying the academic reading construct as measured by IELTS, using retrospective questionnaires and reports as the main research method.

Metacognitive strategies

Another related area of reading that has in recent years received research attention, and which will figure in the research reported here, is readers' metacognitive strategies. Carrell offers a straightforward definition:

> *The term* metacognition *refers to one's understanding of any cognitive process. Metacognition in the context of reading is usually understood to consist of two types of cognition: first, one's* knowledge *of strategies for learning from texts, and, second,* the control *readers have of their own actions while reading for different purposes* (Carrell, 1989: 650, emphasis in original).

The problem for researchers is where cognitive processes end and metacognitive strategies begin, since the difference in definition appears to reside in the reader's awareness and control, as Carrell emphasises, and these might well vary from person to person. Furthermore, metacognitive strategies might take place at various points in the reading process, which makes it difficult to separate them from other processes. Nonetheless, metacognition will be considered in the research in this paper because, as Carrell continues:

> *Successful readers monitor their reading and the state of their learning; they plan strategies, adjust effort appropriately, and evaluate the success of their ongoing efforts to understand.* (Carrell, 1989: 650).

The implication of this, which will be examined in this research project, is that good readers can be distinguished from weaker readers in terms of their planning, direction of effort and self-evaluation, a point which will be of relevance to the research described below.

4

Eye tracking in reading research

4.1 Eye tracking and 'default' reading

The research described in this report builds on the research by Weir, Hawkey, Green and Devi (2009) described above, but also uses eye tracking technology to gain further insight into readers' cognitive processes while they complete IELTS reading test items. The use of eye tracking to investigate reading is not new. Rayner (1998) reviews 100 years of research into reading using eye tracking of various sorts, divided into three periods before we reach what Duchowski (2002) has called the current 'fourth era', distinguished by the possibility of interactivity. Most of this research has been carried out on what has been termed the 'default' mode of reading, i.e. 'when comprehension is proceeding without difficulty and the eyes are continuing to move forward along a line of text' (Reichle et al., 2009: 9), that is, with relatively few regressions.

This mode of reading is certainly employed by some candidates in parts of reading tests, for example, when a candidate reads a text before reading a test item. However, much eye-movement behaviour during reading tests is significantly different from that in 'default' modes of reading, and is particular to language testing, for example, when a reader reads a test item then searches for relevant parts of a text, or skims for general meaning, or uses eye regression so as to check a suspected answer. These behaviours, although significantly different from those used in 'default' reading mode, have received almost no research attention. Furthermore most research with eye tracking has examined native-speaking English readers, not second-language readers of the kind discussed in this report, whose behaviour could be significantly different. For these reasons many of the findings derived from previous eye-movement research are not directly applicable to the research to be discussed here. Even so, it is useful briefly to review the main strands of previous eye-movement research in order to gather insights into methods and possible avenues that could usefully be explored in researching test-taker behaviour among second-language (L2) readers.

Rayner (1998) highlights some of the main insights that eye tracking has offered for our understanding of (mainly 'default') reading. First, when reading English, it is noted that eye fixations (when the eye dwells momentarily on a particular point) typically last about 200–250 milliseconds and the mean saccade size (when the eye moves from one point to another) is seven–nine letter spaces (Rayner, 1998: 375). This is of interest in the present study, particularly when identifying individual words in a text that constitutes the answer to a test item.

Research further demonstrates that saccades can be classified into different types according to their roles (Rayner, 1998; Dussias, 2010; Rayner, Pollatsek, Ashby and Clifton, 2012). In left-to-right languages rightward saccades typically drive onwards through the text, while four other types typically have the function (at least in 'normal' L1 reading) of correcting 'inefficient' text processing (Rayner, 1998). The first of these, 'regressions', are defined as backwards motions for a distance of a few letters so as to reprocess a word which may not have been analysed properly during a prior fixation. In Rayner's (1998) view regressions of more than a few letters are indicative of the reader's probable failure to understand the content (though see below for the case of L2 readers).

The second type, termed 'return sweeps', consist of the eye's return to a precise fixation point, probably recalled by the reader as a source of processing difficulty. Importantly for the research discussed in this report, it is usually surmised that higher-proficiency readers typically use return sweeps more efficiently, since they are able to determine and recall the position in the text that caused them difficulties.

By contrast, readers at lower-proficiency levels tend instead to adopt the third type, as they 'backtrack' through the text less efficiently until they (re)discover the source of difficulty (Rayner, 1998). The fourth type, termed 'corrective saccades', are eye-movements that tend successfully to re-identify text (Rayner, 1998), and again are considered a mark of higher-proficiency readers. These last two types are reminiscent of what Khalifa and Weir (2009) term 'expeditious reading', in which effective readers quickly find areas on which they need to focus, whereas poorer readers by contrast are less efficient at finding parts of a text which they need.

It is also apparent that eye movements are influenced by numerous textual and typographical variables, for example 'as text becomes conceptually more difficult, fixation duration increases, saccade length decreases, and the frequency of regressions [where the eye moves back rather than forwards] increases' (Rayner, 1998: 376). Again this is of significance in the present study, since one of the aims is to compare successful and unsuccessful readers in IELTS tasks. Indeed, it will be hypothesised below on the basis of research into fixation duration that those readers who find the text more difficult, i.e. those who are unsuccessful, will demonstrate greater fixation duration on areas of particular conceptual difficulty than those who are more proficient.

A question that should also be considered is whether or not eye tracking can assist at all in the identification and explanation of underlying cognitive processes. Importantly for the current project, Rayner affirms that the basic theme of his historical review, in particular of the third era from the 1970s onwards, is indeed that 'eye-movement data reflect moment-to-moment cognitive processes' (Rayner, 1998: 372). He expands the point as follows:

> *A crucial point that has emerged recently is that eye-movement measures can be used to infer moment-to-moment cognitive processes in reading [...] and that the variability in the measures reflects on-line processing. For example, there is now abundant evidence that the frequency of a fixated word influences how long readers look at the word* (Rayner, 1998: 376)

More recent studies concur with Rayner as to the value of eye tracking for researching cognitive processes in general (e.g. Bertram, 2011; Buscher, Biedert, Heinesch and Dengel, 2010; Eger, Ball, Stevens and Dodd, 2007). Spivey, Richardson and Dale offer a detailed discussion of how and why eye movements can be taken to be good indicators of cognitive processes, and term them 'a window into language and cognition' (2009: 225). The same metaphor is used by Salvucci and Goldberg who see eye tracking as 'a window into observers' visual and cognitive processes' (2000: 71; see also Anson, Rashid Horn and Schwegler, 2009). Some researchers, such as De Greef, Botzer and Van Maanen (2010) take this to extremes, suggesting – to quote the title of their article – that 'Eye Tracking = Reading the Mind', but this is arguably over-confident. A more cautious position would suggest that eye tracking data should be treated as indicative of cognitive processing, rather than a true and full reflection of it (Reichle, Warren and McConnell, 2009).

4.2 Eye tracking and post-lexical processing

Recent computational models of eye movements, such as the latest versions of the E-Z Reader (Reichle et al., 2009; Rayner, Pollatsek, Ashby and Clifton, 2012), are cautious about going beyond the lexical encoding level. The main reason for this concerns the limited evidence that eye tracking can provide for such higher-order processes when researching default-mode reading. Since such 'default' reading is restricted to reading 'when comprehension is proceeding without difficulty and the eyes are continuing to move forward along a line of text' (Reichle et al., 2009: 9), the reader by definition makes use of relatively few regressions, return sweeps, backtracking and corrective saccades, and therefore provides little evidence that could give insight into processing above the lexical/syntactic levels.

However, in the few studies that focus on more disrupted forms of reading, which, for example, cause readers to make more regressions, evidence does start to emerge which has started to shed light on post-lexical cognitive processes. For instance, studies of reading Finnish sentences containing long compound words (Hyönä and Pollatsek, 1998; Pollatsek, Hyönä and Bertram, 2000, cited in Reichle et al., 2009) were interpreted as giving evidence of post-lexical cognitive activity, and the E-Z Reader version 7 was adapted accordingly. The implication of this is that research into 'non-default' reading activity can potentially help to give insight into processes beyond the lexical/syntactic levels, because such reading provides different kinds of evidence of what readers are doing.

The research reported here concerns readers completing language-test items, which provides eye tracking evidence quite different from that offered in default reading research. As was illustrated in a previous study into cognitive processes in reading tests using eye tracking technology (Bax and Weir, 2012), the very nature of language tests means that readers frequently jump between the text and test item, and repeatedly regress and jump forward in various ways in their search for answers, in ways quite different from default reading patterns. It can therefore potentially offer evidence about higher-order processes of a kind not available to other reading research. For example, if a language test sets a question whose answer can only be found by reading through a whole paragraph using 'inferencing' to deduce the answer (see Table 1, above), and subsequently eye tracking data shows Reader X reading that question, then reading the correct paragraph, then answering the question correctly, it is legitimate to infer from these sources of evidence that this reader has used high-order inferencing strategies of the kind described by Khalifa and Weir (2009). If Reader X then confirms this behaviour through post-hoc questionnaire and stimulated interview, the researcher can surmise with even greater confidence that the reader has indeed made use of those higher-order post-lexical processes. If, by contrast, Reader Y reads the question but fails to read the correct paragraph fully, then fails to answer the test item correctly, and subsequently explains accurately what s/he did in post-hoc questionnaire and stimulated interview, then the researcher can be similarly confident that Reader Y did not use higher-order inferencing. In other words, carefully constructed research into language-test reading, making use of eye tracking technology in conjunction with other research tools, has the potential to offer interesting insights into readers' post-lexical processing behaviour.

5

Research methodology

5.1 Research questions

This report addresses three research questions, the first methodological:

1. To what extent and in what ways can current eye tracking technology shed light on the cognitive processing of participants completing IELTS Academic reading test items on screen?
2. To what extent and in what ways are successful readers differentiated from less successful readers in terms of their eye movements while completing IELTS Academic reading test items on screen?
3. To what extent and in what ways are successful readers differentiated from less successful readers in terms of their cognitive and metacognitive processing while completing on-screen reading test (IELTS) items, as evidenced from eye-movement data, and stimulated retrospective interview data?

5.2 Research approach and instruments

In order to investigate these questions, the following approaches were adopted.

5.3 Participants

A group of Malaysian undergraduates (n=71), with first languages including Bahasa Melayu, Tamil, Chinese and others, took an on-screen test consisting of two IELTS reading passages with a total of 11 test items that are considered to target the cognitive operations which this research seeks to investigate. The text items themselves are set out in Appendix 1 below. The students were first- and second-year undergraduates studying a BEd at a UK university, with an average of IELTS 6.5. The eye movements of a random sample of participants (n=38) were tracked in ways described below. Their activities were also recorded using screen-recording software, which captured every key press, mouse movement, eye movements and facial expressions.

5.4 Preliminaries

Participants signed appropriate ethics forms and personal-information forms. In addition, they were asked to rate their familiarity with computers in general and on-screen tests in particular. As was expected with this young and educated group, all reported extensive familiarity with computer technology and on-screen tests of various kinds.

In this study, it was decided to examine only Careful Local Reading and Expeditious Local Reading, in Khalifa and Weir's terms (2009). To this end, a sentence completion task was chosen of a type which, 'tests [the] ability to find detail/specific information in a text', in other words, testing careful local reading (Cambridge ESOL n.d.). The second was a matching task, which 'assesses [the] ability to scan a text in order to find specific information' (ibid), thereby testing expeditious local reading. The particular texts and task had previously been piloted by Devi (2010) and were selected from the IELTS Practice Papers series (Cambridge University Press), the Academic version, having been developed and trialled by Cambridge ESOL (English for Speakers of Other Languages), the partner responsible for IELTS test production. They were therefore considered representative of genuine IELTS reading tests and items. Devi's research identified some test items that were not functioning effectively, so these were dropped, with the 11 remaining items being those adopted for this research project.

Table 2: Characteristics of the selected reading test texts and items (Devi, 2010)

Text topic	Number of test items	Type of items	Cognitive process targeted cf. Khalifa and Weir (2009)
Human Genome Project	5	Sentence completion – select words from the passage – constructed response.	Careful local reading
Biometric Security Systems	6	Matching	Expeditious local reading

5.5 Test delivery

The texts and items were delivered on screen using Adobe Flash, and linked to a database. The original IELTS test was delivered in paper format, and in order to mitigate the possible effects of a transfer to an on-screen version, it was ascertained in advance that participants were fully conversant with computer use in testing and non-testing environments. In addition, careful training was provided on the format and management of the on-screen test.

5.6 Eye tracking technology specifications

The eye tracker used was a Tobii T60, which dispenses with chin rests, helmets and other distractions, as the tracking cameras are hidden in the monitor casing, thus ensuring that users' behaviour is as natural as possible. The T60 sample rate is 60 Hz per second, which allows detailed tracking of normal reading, and was set to a screen recording rate of ten frames per second. (Full technical specifications can be found at: www.tobii.com). In addition, the device was furnished with binocular tracking (rather than tracking on one eye only), a user camera and speakers for playing the tutorial soundtrack.

5.7 Procedure

After all personal information forms, consent forms and computer-familiarity forms had been completed, the project consisted of the following steps:

Stage 1: for those participants using the eye tracker, individual eye fixations and saccades were calibrated for each participant using the device's calibration tool, which identified each person's individual pattern of gaze and saccade behaviour and ensured the accuracy of the subsequent tracking of their reading during the test.

Stage 2: all participants watched a short video tutorial explaining each aspect of the process they were about to follow, including the on-screen questionnaire, which appeared between each test item.

Stage 3: participants then completed the IELTS reading items on-screen. All answers and responses to the items were saved to a database.

Stage 4: a sample of participants (n=20) then completed a stimulated recall interview procedure, observing video of their own performance with the eye tracking data, and describing and explaining their reading behaviour. This is further explained below (Section 5.8).

5.8 Stimulated recall interviews

After the test had been completed, a sample of the eye-tracked candidates (n=20) was then selected at random to undergo a stimulated recall interview in which they each viewed the video footage of their own test, with their eye movements represented on the screen. Each was asked to describe their reading behaviour as they observed the video, and their analysis was recorded, with the video slowed and stopped at their request to allow them to view and comment freely, with additional prompt questions posed at various points. In this way readers offered a moment-by-moment commentary and explanation of why they read as they did, providing important evidence to triangulate with the eye tracking data and the on-screen questionnaire responses completed after each item had been completed.

6

Analysis and findings

On completion of the data collection, the process of analysis including statistical item analysis of the test items, analysis of the eye tracking data both quantitatively and qualitatively, analysis of the onscreen questionnaire results and analysis of the stimulated recall interview data, all of which will now be described.

6.1 Item analysis

As summarised in Table 3 below, item analysis was carried out on the results of the whole cohort of 71 students to examine the functioning of the reading test items. It can be seen that the reliability coefficient for the 11 items was .722 (Cronbach's Alpha), which is acceptable considering the limited number of items analysed here. The items seem on the evidence to be relatively easy for the tested population, but were nonetheless still targeting the participating students' proficiency levels reasonably well, since the mean of the most difficult item (item 5) was .54 and that of the easiest items (item 9) was .87. However, item 1 was not functioning adequately, the item-total correlation value being lower than .25 (Henning, 1987), so this was excluded from further analysis.

Table 3: Item analysis of the 11 reading items (N=71)

	Mean	Std. deviation	Corrected item-total correlation	Cronbach's Alpha if item deleted
Item 1	.58	.497	.046	.752
Item 2	.79	.411	.265	.716
Item 3	.68	.471	.295	.713
Item 4	.66	.476	.464	.687
Item 5	.54	.502	.501	.680
Item 6	.72	.453	.352	.705
Item 7	.77	.421	.291	.713
Item 8	.86	.350	.433	.696
Item 9	.87	.335	.492	.691
Item 10	.65	.481	.494	.682
Item 11	.56	.499	.501	.680

6.2 Analysis of the eye tracking data

Quantitative analysis

The eye tracking data gathered consisted of full recordings of 38 participants' complete eye movements for all test items. The procedure for analysis was as follows:

Step 1: each item of the reading test was analysed so as to identify the cognitive processes that a reader would need to use in order to answer the item correctly. For example, item 2, on the Human Genome text, was as follows:

> *2. To write out the human genome on paper would require __________ books.*

This was analysed as requiring the reader to locate the following segment of the text (paragraph 3):

> *The human genome is the compendium of all these inherited genetic instructions. Written out along the double helix of DNA are the chemical letters of the genetic text. It is an extremely long text, for the human genome contains more than 3 billion letters. On the printed page it would fill about 7,000 volumes.*

Analysis of this test item suggested that the reader, in order to answer the item correctly, would as a minimum need to use lexical matching – synonymy (in Khalifa and Weir's terms, see Level 2 of Table 1), by identifying the elements 'on paper' and/or 'books' in the question, and lexically matching them with the synonymous or otherwise lexically related elements 'volumes' or 'printed page' and the number '7,000' in the text (the answer), within a single sentence. The reader might also use other clues and cues, but this was taken as the minimum, and the item was therefore analysed as requiring a relatively low (lexical) level of cognitive processing in Khalifa and Weir's terms. The four elements aforementioned ('on paper', 'books' in the question and 'volumes' and 'printed page' in the text) were therefore identified as targets for eye tracking analysis, as will be described below.

By way of further exemplification, item 5, statistically the most difficult item according to the item analysis above, was as follows:

> *5. Research into genetic defects had its second success in the discovery of the cause of one form of_____*

This was analysed as requiring a higher level of cognitive processing in Khalifa and Weir's terms than item 2, since the reader must of necessity read across sentences and use inference, in addition to lexical matching (synonymy) to obtain the answer. The target elements for eye tracking analysis were identified as 'second success' and 'the cause of' in the question and in the text, 'In 1989' (required to identify this as the second success), 'gives rise to' (the idiom which implies causation) and the answer, 'cystic fibrosis'. Other items were analysed in the same way, and the full list of analyses is set out in Appendix 2 below.

Step 2: once the items had been analysed the eye tracking targets identified for each item were selected and highlighted in the eye tracking software as Areas of Interest (AOIs), to allow for the subsequent analysis of each participant's eye-visit count and duration, plus eye-fixation duration and fixation count on each target. A suitable margin of error was allowed in the drawing of each AOI to allow for individual variation in fixation location. This then allowed the software to compare all participants' eye movements, for example to see if successful readers focused more on the target elements, or dwelled longer on each one.

Step 3: the next step was to run analysis on the eye tracking software, and then use statistical analysis so as to answer research questions 2 and 3 (Section 5.1 above). This meant comparing, for each test item, the eye movements of those students who were correct on that item with those who were incorrect on that item, the aim being to see whether differences in eye movement could in part explain readers' success or failure. Analysis examined the attention (in terms of visits and fixations) paid by each student:

1. to the text as a whole while seeking specifically to answer that test item (not including pre-reading or post-reading);
2. to key sections of the text at sentence level or beyond (e.g. on the whole text, on the correct page of the text, or on the paragraph containing the answer), while seeking specifically to answer that test item;
3. to those more specific sections (e.g. words or phrases) of the text and test question previously identified as targets for each test item.

To compare such reading behaviours of successful and unsuccessful test-takers on each item, it was decided to use non-parametric tests, as the datasets did not meet the normality and homogeneity of variances assumptions, and in each case the Independent samples Mann-Whitney U test was used.

It was anticipated that of the three areas of investigation cited above, the first (1) might give insight into readers' expeditious reading ability, since it was hypothesised that better readers, i.e. those who answered correctly in each item, would be able more quickly and efficiently (expeditiously) to find the correct parts of the text so as to answer the question.

The second dimension (2) was intended to give insight into readers' higher order processing of a kind that requires reading across larger segments of text. It was surmised that in cases where better readers paid more attention to these larger segments of text, the reason might be that this was necessary in order to carry out the required cognitive processes beyond the clause level. If weaker readers spent significantly longer, it could be assumed that it was because they either could not find the answer (again a failure of expeditious reading) or else had problems with comprehension. Both possibilities could then be checked through the onscreen questionnaire and stimulated recall interview data.

It was anticipated that the third dimension (3) could give insight into the extent to which better readers or weaker readers attended to key areas of lexis and syntax. If unsuccessful readers were shown to spend longer over a particular word or phrase, it could be surmised that they found it more taxing, and this could be confirmed or disconfirmed at interview; if, by contrast, it was successful candidates who attended significantly more to a target element, it could be supposed that that element was crucial in their successful answering of the question, a point which could also be confirmed or disconfirmed by the interview data.

Step 4: in all cases where the results of the analysis were found to be statistically significant, the findings were analysed closely with regard to individual eye movements (using GazePlot data, for example, as will be illustrated below) and also drawing on the stimulated recall interview data from successful and unsuccessful participants as a means of interpreting and explaining the findings. For example, as will be reported below, it was found that in item 2, unsuccessful students (those who answered the item incorrectly) spent significantly more time reading the page of the text where the answer was to be found than did successful students. Interview data was then used to explain this finding: unsuccessful students reported that they had had difficulty finding the correct part of the text, clearly a weakness in their expeditious reading skills. A similar approach was adopted for each test item, as can now be reported in detail. Only those issues that were found to be statistically significant need be discussed here, and since for items 4, 6, 8, 9 and 11 no significant differences between the eye movements of successful and unsuccessful readers were found, in the dimensions analysed attention will focus on items 2, 3, 5, 7 and 10. (It will be recalled that item 1 was excluded as not functioning well in discriminating between successful and unsuccessful test-takers). The findings will first be reported for each test item individually and then the overall picture discussed in greater detail below.

Item 2

Item 2 required readers, as a minimum, to identify lexical synonymy within a single sentence, a relatively low level of processing. Analysis of eye movements for item 2 showed a significant difference between successful and unsuccessful candidates on one measure, namely the amount of time spent (in seconds) on the correct page of the text. This suggests that although this was a relatively easy item (see Table 3), unsuccessful students were unable to read expeditiously so as to find the location of the answer, and so spent more time looking fruitlessly than the successful readers. Interview data corroborated this impression, with unsuccessful students stating for example 'I was trying here to find where the answer was' and 'Now I was looking for the answer up and down'.

Item 3

Item 3 was as follows:

> *3. A genetic problem cannot be treated with drugs because strictly speaking it is not a______.*

The answer can be found in this sentence:

> *None of the single-gene disorders is a disease in the conventional sense, for which it would be possible to administer a curative drug: the defect is pre-programmed into every cell of the sufferer's body.*

It can be seen that this required readers, as a minimum, to focus on specific terms in the question ('drugs', 'genetic problem') and lexically match them with related terms in the correct part of the text ('drug', 'single-gene disorders', 'diseases'). This is in itself a relatively low-level cognitive process in Khalifa and Weir's terms (see Table 1). Unlike item 2, however, the lexis here is notably more complex and technical, and the lexical relation between terms in the question and terms in the text is more vague. In addition, the reader first has to find the correct part of the text (using expeditious reading skills) and then to disambiguate the complex syntax, in order to identify the answer as 'disease'.

There were no significant differences between the successful students and unsuccessful in terms of expeditious reading, but numerous significant differences in the area of lexis. To summarise in brief, significant differences were found between the successful and unsuccessful groups in terms of their total fixation duration, fixation count, visit duration and visit count on one element in the question (the term 'genetic problem'), and on two elements of the text (the terms 'drug' and 'single-gene disorder', except that the fixation count for the last element was not significantly different). In all cases the unsuccessful students showed greater focus on these elements than successful students.

This is an interesting finding, pointing strongly in one direction. Interview data revealed that weaker students could not understand some of the terms, and – more importantly – could not with confidence match the elements of the answer with the appropriate elements in the text, nor disambiguate the syntax so as to arrive at the correct answer. Unsuccessful students said, for example 'I am here trying to understand these words, looking at them a lot'. This explains the significantly different dwell times and fixations on these particular elements of the question and text, as weaker students focused on them repeatedly in a vain attempt to answer. According to their retrospective reports, in short, the majority of them did not have the lexical knowledge, nor the syntactic knowledge, to disambiguate the target sentence appropriately, and the-eye tracking data appears to support that analysis.

Item 5
When item 5 was discussed above (Section 6.2.1) it was noted that it required a higher level of cognitive processing in term of inference, and reading across sentences, as well as lexical knowledge. It is worth noting that statistically it was the most difficult item (see Table 3).

With this item unsuccessful students looked significantly more at the whole text. For example, the total fixation duration on the whole text for the unsuccessful students was on average (by median) 122.5 seconds, while that for the successful students was 65.39 seconds. This suggested that, as in item 2, they were unable efficiently to find the correct location of the answer, apparently another failure of expeditious reading.

However, with the same item successful students focused significantly more on a key element of the text. The question for this item asked readers to consider 'the cause of' one form of a disorder ('cystic fibrosis', the answer to the item). In order to find the answer, the reader therefore had to find and disambiguate the piece of the text that matches this, in this case the phrase 'gives rise to'. Successful students did indeed focus more on this phrase in terms of fixation duration, fixation count and visit duration, indicating that they successfully identified it and worked on it as they proceeded towards the answer. For example, successful students spent an average (median) of 3.83 seconds fixating on this element while unsuccessful students spent only 0.9 seconds (median). This was supported by interview evidence, for example 'I was looking at that piece of the text a lot because I thought it fitted what the question wanted'. In short, the evidence from eye tracking and interviews seems to point to successful students' cognitive processing of the relevant syntax in the text so as to arrive at the correct answer.

Item 7
Unlike items 1–5, the later test items did not require the typing of a full word or number, but only the selection of a letter corresponding to a given correct answer. As reported above (Table 2), these items were designed to make students read *expeditiously* for *local* information, this being achieved by scattering the answers across a lengthy text and obliging readers to look for particular words or phrases in order to identify the correct match.

As was suggested by evidence from items 2 and 5, successful readers seem better at expeditiously locating the place of a correct answer in the text. In those two items it was the weaker students who spent longer on larger areas of text (a page in one case and the whole text in the other), presumably unable to locate the correct part quickly, and then read the target text carefully. However, with item 7 it was successful readers who focused significantly more on a smaller, correct part of the text (see Table 7), suggesting that they were better able to identify the correct segment so as then to work more intensively on it. For example, correct students had significantly more fixations on the correct paragraph (median = 15.19) than unsuccessful students (median = 2.51).

In addition, item 7 demonstrated a further significant difference between successful and unsuccessful students in terms of their focus on one element in the text which was key to the answer, namely the phrase 'Olympic athletes'. On four different measures, successful students gave significantly more attention to this element, while unsuccessful students virtually ignored it – to the extent that more than half did not fixate on it at all (visit count mean = 0.56, median = 0). The reason why this was significant to them is that elsewhere in the text other athletes were mentioned, and analysis of the video data shows that incorrect students were distracted by that. As the previous discussion shows, however, successful students demonstrated greater acuity in locating the correct paragraph, and in focusing on the key element, in this case the mention of Olympic athletes specifically, which gave them the correct answer.

Item 10

As with item 7 above, this item achieved its aim of testing expeditious reading skills, in that successful students attended significantly more to the paragraph with the correct answer than unsuccessful students, reinforcing the evidence from earlier items suggesting that unsuccessful students lacked the appropriate expeditious reading locating skills. For example, successful students spent 13.62 seconds on the paragraphs (visit duration median) in comparison with only 0.54 seconds for unsuccessful students.

Discussion

The eye tracking data, although it showed significant differences in only five of the ten items analysed, succeeded nonetheless in demonstrating clear differentiation between proficient and less less-proficient students at a range of levels of cognitive processing in different test items, as follows:

Table 4

Levels of processing (see Table 1 for details)	Gloss	
Lexis: word matching	Matching of identical word in question and in the text, as key to the answer	3 ('drug' in text) 7 ('Olympic athletes' in text)
Lexis: synonym and word-class matching	Matching of synonym or lexically related word in question and text	3 'genetic problem' in the question, and 'single-gene disorder' in the text
Grammar/syntax	Focus on significant syntactic structure or other grammatical element as part of working out the answer	5 'gives rise to' in the text
Propositional meaning	Focus on elements, in question and/or text, required for constructing a propositional meaning essential to answering the question correctly	
Inference	Focus on elements, in question and/or text, required for constructing inferences essential to answering the question correctly	
Building a mental model		None, as this was not the focus of these IELTS test items
Understanding text function		None, as this was not the focus of these IELTS test items

These test items were therefore successful in distinguishing between proficient and less proficient students in terms of their cognitive processing at the lexical (word matching), lexical (synonymy) and grammatical levels. There was no evident difference between successful and unsuccessful students at higher cognitive levels with this cohort, but this can be attributed to the fact that the test items were designed specifically to target lower-order specific information.

Expeditious reading

The results also showed significant differences in terms of expeditious reading, with weaker students apparently unable to locate the site of a correct answer as effectively as stronger students on some items and thus spent longer on larger chunks of text (items 2 and 5), while stronger students demonstrated that they could locate a smaller, particular part of a text and focus more expeditiously on it so as to extract the answer (items 7 and 10).

It should not be forgotten that in half of the test items (five of ten), no significant differences were identified in the eye-movement behaviour of those who were correct on each item and those who were not. The implication of this is that with these items the more successful students did not use any one cognitive process significantly more than the less successful students in arriving at the correct answer, but perhaps used a variety of processes and strategies, or other faculties related, for example, to memory or lexical knowledge, none of which was either predominant, or else was traceable in the eye tracking record. This in itself is of interest, suggesting that many test items – even items designed to target particular cognitive processes – might in practice be answered using a range of cognitive processes operating together.

However, since the research showed that in some test items there were clear and significant differences between the cognitive processes of proficient and less proficient test takers, it can be concluded that – in terms of the first research question set out in Section 5.1 above – the value of using eye tracking analysis to evaluate the cognitive processing of text takers in language tests has been established. In other words, it is indeed possible, through the use of eye tracking data, to gain insight into readers' cognitive processes during on-screen language tests.

At the same time, the research successfully offered an answer to the second and third research questions, in demonstrating that when completing some IELTS Academic test items, proficient readers use significantly different eye-movement behaviour in comparison with less proficient test takers. Their behaviour is presumably linked to different cognitive processing on a number of levels.

At the same time, the research also demonstrated limitations in eye tracking data, since with several items there was no significant difference between the eye movements of successful and unsuccessful test takers, indicating that the former must have been using cognitive processes and other resources (such as memory), which did not show up in the eye tracking data, and which need therefore to be investigated by other means.

Implications for teachers, learners and test design

The research has interesting implications for teachers and learners, since it is clear that those candidates were more successful who made use of expeditious reading strategies, particularly to locate in the text the possible site of the correct answer as speedily as possible. Successful candidates also showed better abilities at the lexical level (in matching words in the question and the text, and in doing the same with synonyms), which suggests that it is valuable to work extensively on their lexical knowledge and their ability to identify lexical matches of various kinds. They also showed better ability in terms of dealing with syntactic ambiguities, so it is useful for learners to expect in reading test items, and then to deal with, grammatical ambiguity, which might obscure the correct answer.

In terms of designing reading tests, the findings from this study suggest that to some extent it is effective for test designers to target specific cognitive processes, since some items in this research proved able to distinguish between successful and unsuccessful candidates in terms of their cognitive processing operations. Test-item writers can therefore usefully draw on Khalifa and Weir (2009), for example, to plan the kinds of items they design so as to test different levels of cognitive processing, with a view to achieving greater cognitive validity in their reading tests.

In conclusion, then, the project reported here is among the first substantial research reports into reading tests using eye tracking technologies. It has successfully illuminated some of the darker corners of readers' mental processing while they try to complete such tests, demonstrating the potential of this research tool and also showing some of its limitations. It has offered insights into successful and unsuccessful readers' behaviour in ways that can be of value to test designers, to students, to teachers and to researchers interested in the wider cognitive dimensions of reading.

References

Afflerbach, P and Johnston, P (1984) On the use of verbal reports in reading research. *Journal of Reading Behaviour* 16/4: 307–321.

Alderson, JC (2000) *Assessing Reading*. Cambridge University Press, Cambridge.

Anson, C, Rashid Horn, S and Schwegler, R (2009) The Promise of Eye Tracking Methodology for Research on Writing and Reading. *Open Words: Access and English Studies* 3.1, 2009: 5–28.

Bax, S and Weir, C (2012) *Investigating learners' cognitive processes during a computer-based CAE reading test*. Cambridge Research Notes, Cambridge ESOL, 47/February 2012: 3–14. Available online at: www.cambridgeesol.org/rs_notes/rs_nts47.pdf

Baxter, G and Glaser, R (1998) Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practices* 17/3: 37–45.

Bertram, R (2011) Eye movements and morphological processing in reading. *The Mental Lexicon* 6/1: 83–109.

Birch, M (2007) *English L2 reading: getting to the bottom*. Mahwah: Lawrence Erlbaum Associates.

Britt, MA and Sommers, J (2004) Facilitating textual integration with macro-structure focusing tasks. *Reading Psychology* 25/4: 313–339.

Buscher, G, Biedert, R, Heinesch, D and Dengel, A (2010) 'Eye tracking analysis of preferred reading regions on the screen', in *Proceedings of CHI Extended Abstracts* 2010: 3307–3312

Carrell, PL (1983) Some issues in studying the role of schemata or background knowledge in second language comprehension. *Reading in a Foreign Language* 1/2: 81–92.

Carrell, PL (1984) Evidence of formal schema in second language comprehension. *Language Learning* 34/1: 87–112.

Carrell, P (1989) Metacognitive Strategy Training for ESL Reading. *TESOL Quarterly* 23/4: 647–678.

Carson, JG (2001) 'A task analysis of reading and writing in academic contexts', in Belcher, D and Hirvela, A (eds) *Linking literacies: Perspectives on L2 reading-writing connections* (pp. 48–83). Ann Arbor: The Michigan University Press.

Carver, R (1992) Reading Rate: theory, research and practical implications. *Journal of Reading* 36/2: 84–95.

Cerdan, R (2006) *Integration information processes from multiple documents*. Departemento de Psicologia Evolutiva y Educacion Valencia: Universitat de Valencia.

Cordon, L and Day, J (1996) Strategy use on standardized reading comprehension tests. *Journal of Educational Psychology* 88: 288–95.

Devi, S (2010) *Investigating the academic reading of first-year undergraduates at a British university*. PhD thesis, University of Bedfordshire.

Duchowski, A (2002) A Breadth-First Survey of Eye tracking Applications. *Behavior Research Methods, Instruments, & Computers (BRMIC)* 34/4: 455–470.

Eger, N, Ball, L, Stevens, R and Dodd, J (2007) *Cueing Retrospective Verbal Reports in Usability Testing Through Eye-Movement Replay*. Proceedings of HCI 2007, The 21st British HCI Group Annual Conference, University of Lancaster, UK. Available online at: www.bcs.org/server.php?show=ConWebDoc.13300

Enright, MK, Grabe, W, Koda, K, Mosenthal, P, Mulcahy-Ernt, P and Schedl, M (2000) *TOEFL 2000 Reading Framework: A Working Paper*. TOEFL Monograph Series MS-17. ETS.

Field, J (2012) *Cognitive Validity* seminar talk at the CRELLA Spring Seminar, University of Bedfordshire, 15 March 2012.

Field, J (forthcoming) 'Cognitive Validity' in Geranpayeh, A and Taylor, L (eds) *Examining Listening*. Cambridge: Cambridge University Press.

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: 175–198.

Gass, S and Mackey, A (2000) *Stimulated Recall Methodology in Second Language Research.* Mahwah, N.J., Lawrence Erlbaum.

Gernsbacher, MA (1990) *Language comprehension as structure building*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Glaser, R (1991) 'Expertise and assessment', in *Testing and Cognition*, Wittrock, MC and Baker, EL (eds). Prentice Hall, Englewood Cliffs, pp. 17–30.

de Greef, T, Botzer, A and Van Maanen, P-P (2010) *Eye tracking = Reading the mind* ECCE 2010 – European Conference on Cognitive Ergonomics 2010: The 28th Annual Conference of the European Association of Cognitive Ergonomics.

Henning, G (1987) *A Guide to Language Testing: Development, Evaluation, Research*. Cambridge, MA: Newbury House.

Hoover, WA, and Tunmer, WE (1993) 'The components of reading' in Thompson, GB, Tunmer, WE and Nicholson, T (eds) *Reading acquisition processes*. WBC Print Ltd, Great Britain, pp. 1–19.

Jordan, RR (1998) *English for Academic Purposes: A guide and resource book for teachers*. Cambridge: Cambridge University Press.

Just, MA and Carpenter, PA (1980) A theory of reading: from eye fixations to comprehension. *Psychological Review* 87/4: 329–354.

Just, MA and Carpenter, PA (1987) *The psychology of reading and language comprehension*. Newton, Massachusetts: Allyn and Bacon, Inc.

Just, MA and Carpenter, PA (1992) A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99/1: 122–149.

Khalifa, H and Weir, C (2009) *Examining Reading*. Cambridge: UCLES/Cambridge University Press.

Kintsch, W and van Dijk, TA (1978) Toward a model of text comprehension and production. *Psychological Review* 85/5: 363–394.

Kobayashi, M (2009) *Hitting the mark: how can text organisation and response format affect reading test performance?* New York: Peter Lang.

Koda, K (1987) 'Cognitive strategy transfer in second language reading', In Devine, J, Carrell, PL and Eskey, D (eds) *Research in reading in a foreign language*. Washington DC: TESOL (pp. 125–144).

Kroll, B (1990) *Second language writing: research insights for classroom*. Cambridge: Cambridge University Press.

Meyer, BJF (1975) *The organisation of prose and its effect on memory*. Amsterdam: North Holland Publishing Company.

Perfetti, CA (1985) *Reading ability*. New York: Oxford Press.

Perfetti, CA (1997) *Sentences, individual differences and multiple texts: three issues in text comprehension*. Discourse Processes 23, pp. 337–355.

Pressley, M and Afflerbach, P (1995) *Verbal protocols of reading: the nature of constructively responsive reading*. Hillsdale, N. J., Lawrence Erlbaum.

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

Rayner, K and Pollatsek, A (1989) *The psychology of reading*. Prentice Hall, Englewood Cliffs, NJ.

Reichle, E, Warren, T and McConnell, K (2009) Using E-Z Reader to Model the Effects of Higher-Level Language Processing on Eye Movements During Reading *Psychonomic Bulletin & Review* 16/1: 1–20.

Rosenfeld, M, Leung, S and Oltman, PK (2001) The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels. *TOEFL Monograph Series* MS – 21.

Ruiz-Primo, M, Shavelson, R, Li, M and Schultz, S (2001) On the Validity of Cognitive Interpretations of Scores From Alternative Concept-Mapping Techniques, *Educational Assessment* 7/2: 99–141.

Salvucci, D and Goldberg, J (2000) *Identifying fixations and saccades in eye tracking protocols*. In Proceedings of the Eye tracking Research and Applications Symposium (pp. 71–78). New York: ACM Press.

Snow, CE (2002) *Reading for understanding: toward a research and development program in reading comprehension*. Santa Monica: RAND Education.

Spivey, NN and King, JR (1989) Readers as writers composing from sources. *Reading Research Quarterly* 24/1: 7–26.

Spivey, M, Richardson, D, and Dale, R (2009) 'The movement of eye and hand as a window into language and cognition', in Morsella, E and Bargh, J (eds) *Oxford Handbook of Human Action*. New York: Oxford University Press, pp. 225–248.

Stanovich, KE (1980) Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly* 16: 32–71.

Stromso, HI and Braten, I (2002) Norwegian law students' use of multiple sources while reading expository texts. *Reading Research Quarterly* 37/2: 208–227.

Urquhart, A and Weir, C (1998) *Reading in a second language: process, product and practice*. Longman, London.

van Dijk, TA and Kintsch, W (1983) 'The notion of macrostructure' in van Dijk, TA and Kintsch, W, (eds), *Strategies of discourse comprehension*. New York: Academic Press , pp. 189–223.

Vipond, D (1980) Micro- and macro-processes in text comprehension. *Journal of Verbal Learning and Verbal Behaviour* 19: 276–296.

Weir, C (2005) *Language Testing and Validation: an evidence based approach*. Palgrave Macmillan.

Weir, C, Hawkey, R, Green, T and Devi, S (2009) *The cognitive processes underlying the academic reading construct as measured by IELTS*. British Council/IDP Australia IELTS Research Reports vol, 9, 4, pp. 157–189.

Appendix 1

Test items

a) Part 1: gap-fill questions

1. The passage compares the size of the project to the ______
2. To write out the human genome on paper would require _____books.
3. A genetic problem cannot be treated with drugs because strictly speaking it is not a______
4. Research into genetic defects had its first success in the discovery of the cause of one form of ______
5. The second success of research into genetic defects was to find the cause of ______

b) Part 2: matching task

List of biometric systems	List of users to be matched by the test taker while reading
A. fingerprint scanner	6. sports students
B. hand scanner	7. olympic athletes
C. body odour	8. airline passengers
D. voiceprint	9. welfare claimants
E. face scanner	10. home owners
F. typing pattern	11. bank customers

Appendix 2

Analysis of each test item in terms of anticipated cognitive processing

Item	Level of text	Type of processing required	Target in question item – AOI	Target in the reading texts – AOI
2	Within one sentence	Lexical, synonymy	Books On paper	Printed page 7,000 volumes
3	Within one sentence	Lexical matching Syntactic parsing	Drugs Genetic defects	Drug (only occurrence in text) Single-gene disorders disease
4	Across sentences	Inference (first) Lexical, synonymy	First success One form of	In 1986 (compared to 1989 later) One type of Muscular dystrophy
5	Across sentences	Inference (dates) Lexical, synonymy	Second success The cause of	In 1989 (compared to 1986 earlier) Gives rise to cystic fibrosis
6	Within one sentence	Lexical matching Synonymy	Sports students	Students, athletic (para A)
7	Within one sentence	Lexical matching	Olympic athletes	Olympic, athletes (para E)
8	Within one sentence	Lexical matching Synonymy	Airline passengers	Passengers, airport (para F)
9	Within one sentence	Lexical matching	Welfare claimants	Welfare, welfare payments (para D)
10	Within one sentence	Synonymy	Homeowners	Housing (para A)
11	Within one sentence	Lexical matching	Bank customers	Customers, bank (para A)

ISBN 978-0-86355-720-0