

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA CORRESPONDENTE

THE IMPLICATIONS OF TIME ALLOCATION AND TEST-WISENESS
IN A TEST OF READING COMPREHENSION IN ENGLISH

MARCOS DA SILVA

Dissertação submetida à Universidade Federal de Santa Catarina em cumprimento
parcial dos requisitos para obtenção do grau de

MESTRE EM LETRAS

FLORIANÓPOLIS
Junho/2006

Esta dissertação de Marcos da Silva, intitulada *The Implications of Test-wiseness and Time Allocation in a test of Reading Comprehension in English*, foi julgada adequada e aprovada em sua forma final, pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção do grau de

MESTRE EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
Opção: Língua Inglesa e Lingüística Aplicada

José Luiz Meurer
Coordenador

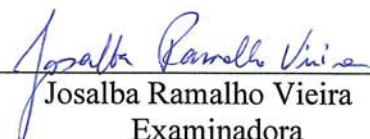
BANCA EXAMINADORA:



Lêda Maria Braga Tomitch
Orientadora e Presidente



Sara Rejane Freitas de Oliveira
Examinadora



Josalba Ramalho Vieira
Examinadora

Florianópolis, 30 de junho de 2006.

*To my parents, Nelson and Tereza,
with love and gratitude.*

ACKNOWLEDGEMENTS

I am indebted to many people who helped me on the way to writing this MA thesis. Perhaps I cannot remember all the names I should mention here, but the following ones will be unforgettable:

- Professor Dr Lêda Maria Braga Tomitch, for having advised me in this study, contributing with her invaluable suggestions, and certainly helping me to become a beginning researcher.
- CAPES, for the financial support granted to this study.
- The members of the examining committee, Professors Adriana Dellagnelo, Josalba Ramalho Vieira, and Sara Oliveira. Also, of course, Professor Lêda once again, for presiding over this committee.
- Denise, Liam and Sílvia, from *Phoenix*, in Joinville, Gary and Gareth from Cultura Inglesa in Florianópolis, Fábio from CCAA in Joinville, and Valéria Barreira, from UNIVILLE, for their helpful work as raters in this study.
- The professionals who helped me with statistics: Peters, from UDESC, in Joinville, and Professor Enori Carelli, from UDESC, Joinville; and very especially, Daru, from Faculdade Cenecista de Joinville, for his patience in making me understand some whys and hows concerning the statistical procedures used in this study.
- Cristala Athanázio Buschle and Paulo Leôncio, for having recommended me for the *Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente* at UFSC.
- Professor Josalba Ramalho Vieira, for having read the project of this MA thesis, for having taught me some academic writing skills, and for sharing materials on reading.
- Some friends from the MA course: Juliana Sartori, for the constant incentive, Miquéias Rodrigues for his precious readings of some parts of this thesis, and very especially, Juliane Massarollo, for always being so helpful in all possible ways, and for her intellectual contributions to this work.
- Some of my best friends in this world, Rose, Everton, Cúnico, Medeiros and Cátia (Vida), Raimundo and Crele, Eloízio and Leila, Ednilson and Salete, Fabiana and Fernando, Franciele and Márcio, Alcídio, Cleidi, Mariazinha, and Kátia Rocha, for I could always count on them at any time, and for their comprehension at the moments I had to choose not to be with them.
- My students Rodrigo, Emerson, Andrea, Diogo, Marineusa, Fabiana, Clio, Fraser, and Yelva, for their comprehension for all the times I had to be absent from work.
- Last, but surely not least, my family, for their invaluable and constant support, and because they always believed I would get here.

Thank you all for having helped me to achieve such a milestone as this MA thesis already means to me.

ABSTRACT

THE IMPLICATIONS OF TIME ALLOCATION AND TEST-WISENESS
IN A TEST OF READING COMPREHENSION IN ENGLISH

MARCOS DA SILVA

UNIVERSIDADE FEDERAL DE SANTA CATARINA

2006

Supervising Professor: Dr Lêda Maria Braga Tomitch

This study investigated the implications of time allocation and test-wiseness in an English reading test. Thirty EFL readers and 10 native English readers participated in the study. The issue of time allocation was investigated by having the participants take a reading test in two different conditions. In condition 1, participants had a set time limit to conclude their tests. In condition 2, they were allowed to spend as long as they needed to finish their tests. Although there was a tendency for the participants to have higher marks in the no time limit condition, the results showed no statistical significance for the intra-group comparisons. To investigate the implications of test-wiseness, a comparison was made between the mean scores of the only group of participants known to be highly familiar with the format of the test used in this study and the mean scores of each of the other groups in each condition. The results showed that the group which was under test-wiseness effects outperformed only one of the other groups, and only in terms of their mean time spent on the test, in one condition only. Further research is suggested, which may confirm the tendency found in this study in relation to time allocation. As to the issue of test-wiseness, the suggestion is that a treatment in test format familiarity is given to a group of participants, instead of the option to select a group expected to possess test-wiseness.

Number of pages: 85

Number of words: 26.064

RESUMO

IMPLICAÇÕES DE ALOCAÇÃO DE TEMPO E DE PERÍCIA EM TESTE
NUM TESTE DE COMPREENSÃO DE TEXTO EM INGLÊS

MARCOS DA SILVA

UNIVERSIDADE FEDERAL DE SANTA CATARINA

2006

Professora Orientadora: Dra. Lêda Maria Braga Tomitch

Este estudo investigou as implicações de alocação de tempo e de perícia em teste num teste de compreensão de texto em inglês. Trinta leitores de inglês como língua estrangeira e dez leitores nativos de inglês participaram deste estudo. Para investigar a questão da alocação de tempo, os participantes fizeram um teste de leitura em duas condições diferentes. Na condição 1, eles tiveram um tempo limite para realizar o teste. Na condição 2, os participantes dispuseram de quanto tempo precisaram para concluir seus testes. Apesar de ter havido uma tendência para os participantes obterem notas mais altas ao realizarem os testes sem limite de tempo, os resultados não apresentaram significância estatística nas comparações intra-grupo. Para investigar as implicações de perícia em teste, foram comparadas as médias do único grupo de participantes que era altamente familiarizado com o formato do teste utilizado às médias de cada um dos demais grupos em cada condição. Os resultados mostraram que o grupo que estava sob os efeitos de perícia em teste foi superior a somente um dos demais grupos, somente em termos da média de tempo utilizado para a realização do teste, em somente uma das condições. Sugere-se mais pesquisa para verificar se a tendência observada neste estudo em relação à alocação de tempo é confirmada estatisticamente. Com relação à questão de perícia em teste, a sugestão é que seja dado tratamento no sentido de familiarizar um grupo com o formato de um teste, ao invés de optar-se por selecionar um grupo que se espere possuir tal conhecimento.

Número de páginas: 85

Número de palavras: 26.064

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION.....	1
1.1- Purposes of the Study.....	5
1.2- Research Questions.....	5
1.3- Motivation for the Study.....	6
1.4- Significance of the Study.....	7
1.5- Organisation of the Thesis.....	8
CHAPTER 2 – REVIEW OF LITERATURE.....	10
2.1- Models and Component Processes in Reading Comprehension.....	10
2.1.1- Bottom-up Models.....	11
2.1.2- Top-down Models.....	12
2.1.3- Interactive Models.....	14
2.1.4- About the Component Reading Processes	15
2.1.5- A view of Reading.....	16
2.2- Individual Differences in Reading Comprehension.....	17
2.2.1- Individual Differences in Declarative Knowledge.....	18
2.2.2- Reading Speed.....	20
2.2.3- Reading in L1 versus L2/FL Reading.....	24
2.3- The Assessment of Reading Comprehension.....	25
2.3.1- Validity.....	26
2.3.2- Reliability.....	27
2.3.3- Test-wiseness.....	28
2.3.4- Time Allocation.....	29
2.4- Relating the Issues Reviewed to the Present Study.....	31

CHAPTER 3 – METHOD.....	32
3.1- Participants.....	32
3.1.1- Group 1- EFL Students Preparing to Take the CAE Examination.....	33
3.1.2- Group 2- Advanced EFL Readers Not Familiar with the CAE Examination...	34
3.1.3- Group 3- EFL Students Taking an MA Course in English Language and Literature	34
3.1.4- Group 4- Readers of English as their Native Language.....	35
3.2- Materials.....	35
3.2.1- The CAE Reading Paper	36
3.2.2- The Readiness Test.....	38
3.2.3- The Practice Material.....	41
3.2.4- Grading the Tests.....	41
3.3- Procedures.....	42
3.3.1- Taking the Readiness Test.....	43
3.3.2- Results of the Readiness Test.....	44
3.3.3- Taking the CAE Reading Paper.....	48
3.3.4- Framework for Data Analysis.....	49
 CHAPTER 4 – RESULTS AND DISCUSSION.....	 51
4.1- Summarising the Objectives and the Design of the Study.....	52
4.1.1- Objective and Design Concerning Research Question 1	52
4.1.2- Objective and Design Concerning Research Question 2.....	54
4.2- Results and Discussion for Research Question 1.....	55
4.2.1- Answering Research Question 1.....	56
4.2.2- Observed Tendencies in Relation to time Allocation.....	59

4.3- Results and Discussion for Research Question 2.....	63
4.3.1- Answering Research Question 2	63
4.3.1.1- Comparing the mean marks of group 1 to those of the other groups.....	64
4.3.1.2- Comparing the mean times spent by group 1 to those spent by the other groups.....	66
4.3.2- Issues Relating Generally to Research Question 2.....	70
4.4- Final Discussion.....	72
4.4.1- Implications of the Readiness Test.....	72
4.4.2- Final Remarks	74
CHAPTER 5 – CONCLUDING REMARKS, LIMITATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH.....	76
5.1- Concluding Remarks.....	76
5.2- Limitations of the Study.....	78
5.3- Future Research Directions.....	79
REFERENCES.....	81

LIST OF TABLES

Table 1 – <i>Results of the Readiness Test for Group 1</i>	44
Table 2 – <i>Results of the Readiness Test for Group 2</i>	45
Table 3 – <i>Results of the Readiness Test for Group 3</i>	46
Table 4 – <i>Results of the Readiness Test for Group 4</i>	47
Table 5 – <i>Mean Mark of Each Subgroup in the CAE Reading Paper</i>	57
Table 6 – <i>Mean Mark from All the Participants in the CAE Reading Paper in Each Condition</i>	60
Table 7 – <i>Mean Time of Each Subgroup in the CAE Reading Paper</i>	67
Table D1 – <i>Group 1.1 – Participants’ Marks and Times in the CAE Reading Paper</i> ..	98
Table D2 – <i>Group 1.2 – Participants’ Marks and Times in the CAE Reading Paper</i> ..	98
Table D3 – <i>Group 2.1 – Participants’ Marks and Times in the CAE Reading Paper</i> ...	99
Table D4 – <i>Group 2.2 – Participants’ Marks and Times in the CAE Reading Paper</i> ...	99
Table D5 – <i>Group 3.1 – Participants’ Marks and Times in the CAE Reading Paper</i> ...	99
Table D6 – <i>Group 3.2 – Participants’ Marks and Times in the CAE Reading Paper</i> ..	100
Table D7 – <i>Group 4.1 – Participants’ Marks and Times in the CAE Reading Paper</i> ..	100
Table D8 – <i>Group 4.2 – Participants’ Marks and Times in the CAE Reading Paper</i> ..	100

LIST OF APPENDICES

APPENDIX A – The CAE Reading Paper.....	86
APPENDIX B – CAE Reading Paper Answer Sheet.....	96
APPENDIX C – Answer Key to the CAE Reading Paper.....	97
APPENDIX D – Participants’ Marks and Times in the CAE Reading Paper.....	97
APPENDIX E – The Readiness Test.....	101
APPENDIX F – Answer Key to the Readiness Test.....	107
APPENDIX G – The Practice Material.....	108
APPENDIX H – Answer Key to the Practice Material.....	112

CHAPTER I

INTRODUCTION

From the history of research on reading comprehension, it now seems reasonable to contend that every attempt to investigate text comprehension should be concerned with what the implications are of asserting that a reader has or has not comprehended a text. Certainly, another issue that deserves attention is the question of what exactly accounts for how skilled a person can be at reading. According to Just and Carpenter (1987), skilled reading involves many component processes and extensive declarative knowledge, and there is now research evidence to show that individual differences in reading comprehension are not attributed to a single process. Skilled readers perform differently from less skilled readers in almost all processes through which they go in order to reach text comprehension (Gagné, Yekovich & Yekovich, 1993; Just & Carpenter, 1987).

One assumption that has been supported in reading comprehension research is that reading speed is strongly associated with reading comprehension. In this way, it is assumed that the faster a reader can process a text, the easier he or she can build a mental representation of the text content. Conversely, longer reading times have been associated with greater difficulty in building that representation (Jones, 1995). In other words, research has pointed out that fast reading indicates that the reading process has been automatised.

A conventional formal way to assess reading comprehension has been through the use of standardised reading tests. These tests are intended to measure subjects' performance according to a pre-established consensus as to what would indicate good reading comprehension of the texts used in a given test (Urquhart & Weir, 1998). However, one criticism about standardised reading tests, which may classify individuals

as good or poor readers, is that most of them take into consideration only the product of reading, not the process. These tests are called norm-referenced tests, and their use usually implies the comparison of a reader's score in a given test to a set of other individuals' scores in the same test. The result of such a comparison indicates how well test takers have performed in that specific test (Just & Carpenter, 1987). One example of such tests is the reading papers of the proficiency examinations in EFL¹ (English as a Foreign Language) designed by the University of Cambridge, the Cambridge ESOL (English for Speakers of Other Languages) examinations.

Moreover, concerning reading speed and comprehension accuracy, one problem with standardised reading tests is that they may place too much emphasis on speed. That being the case, such tests may favour readers who, for whatever reason, read fast, be it because they may know they are being tested on speed, or because they happen to have good background knowledge on the topic of the text being read, or just because they are fast test takers. Conversely, readers who tend to be slower, for example because they choose to reread the text or portions of the text in order to make sure they will do well in the test, or because they get nervous when taking tests and so forth, might end up not even having the necessary time to conclude their tests, which might contribute to their being categorised as poor readers.

Commenting on the trade-off between comprehension accuracy and reading speed, Just and Carpenter (1987) suggest that a valid assessment of reading skill should measure both variables – reading speed and reading comprehension. When only one of those is evaluated, the result might not be a good predictor of reading skill. For instance, in one experiment which compared the performance of good readers and less proficient readers, Block (1992) found that two subjects from the less proficient group (one native speaker of English and one native speaker of Chinese, both reading expository prose in

¹ In this thesis, the term EFL is used interchangeably with the term English as a Second Language (ESL).

English) performed much like proficient readers, using “good reader strategies”. The explanation was that those subjects might have been classified as ‘less proficient readers’ as a result of *slow* reading rather than *unskilled* reading.

If a text is well written, in the sense that it presents all the necessary elements to allow readers to establish a coherent mental representation of it, comprehension of that text should not take too long, a fact which has to do with an expected reduction in reading time when texts are locally coherent (Murray, 1995). Conversely, with more complex texts, if readers are also expected to perform higher-level comprehension processes in order to reach comprehension of the texts being read, then longer reading times might be necessary. In the case of a coherence break, for instance, reinstatement processes are supposed to take place in order for the reading to proceed, for example, the reinstatement of goals might need to take place in order for the reader to reestablish coherence to the text. Such reinstatements require readers to rely on their background knowledge; for instance, readers may search for elements processed at an earlier time in the text being currently read, that is, elements that might no longer be available in working memory, and/or they may have to activate some world knowledge from long-term memory. Those processes are assumed to increase the degree of difficulty for the building of the mental representation of the text, thus making reading slow down (O’Brein, 1995; van den Broek, Risdén & Husebye-Hartmann, 1995).

Among all the strategies that skilled readers may use in their attempts to fully comprehend especially a difficult text, rereading seems to be one that is most frequently used, and it is certainly an alternative that slows down the reading process (Block, 1986; Paris, Lipson & Wixson, 1983; Paris, Wasik & Turner, 1991). The findings reviewed by Jones (1995) are in line with this perspective. As she has noted:

Previous research suggests that a mental model may not be completely formed after a single exposure to the information. Perrig and Kintsch (1985) found that when subjects were given a single exposure to a complex text, they were unable

to make accurate spatial inferences based on the presented information. However, Taylor and Tversky (1992) found that when subjects were given as many as four self-paced exposures to a text, they were quite adept at making inferences from a variety of perspectives (p.340).

Research on language testing still does not have much to say about the issue of time allocation in tests. As for tests of reading comprehension, it has been suggested that time allocation should be in accordance with task types. For instance, more time should be allowed when *careful reading* is expected and, conversely, less time should be given to test takers when they are expected to perform *expeditious reading*. Careful reading should be necessary for answering comprehension questions which require high-order processes such as inference making, and expeditious reading includes the use of strategies such as skimming or scanning (Urquart & Weir, 1998).

Concerning task types, another factor which may have implications for the outcomes of reading tests is task familiarity (Alderson, Clapham, & Wall, 1995; Urquart, & Weir, 1998). Ideally, tests should present task formats known to test takers, so that what is measured is the test takers' reading comprehension ability and not their expertise in taking tests of a given format (Alderson, Clapham, & Wall, 1995). On the other hand, if test takers happen to be too familiar with certain task types, they may develop what Allison (1999) calls *test-wiseness*. This term refers to the development of strategies to approach particular tasks due to task familiarity, which in turn may also affect test scores interpretations, hence lowering the validity of the test.

A distinction has also been made between *power tests* and *speeded tests*. While in the former test takers are allowed to spend as much time as they need to perform the test tasks, in the latter, not all test takers are expected to conclude their tests (Allison, 1999; Bachman, 1990). However, despite the existence of such a distinction, according to Bachman, few examinations consist of power tests. What happens is that tests are

usually timed, although their designers do not usually consider such tests to be speeded tests.

1.1 – Purposes of the study

Given the aforementioned discussion about the trade-off between reading speed and comprehension accuracy, and taking into account also the issue of test-wiseness, I conducted this study with two main objectives. First, I intended to verify whether individuals at a comparable level of reading proficiency in EFL, as well as proficient readers of English as their native language, would achieve higher marks when taking a reading test with unlimited time, that is, as opposed to their taking the same test within a set time limit. The second objective relates to the implications of test-wiseness. In this respect, I aimed at scrutinising the effects of test-wiseness by comparing the performance of advanced EFL readers who were very familiar with the tasks present in a reading test in English, with that of readers of English who were only slightly familiar with that test. The latter type of readers were either of comparable proficiency in EFL reading or proficient readers of English as their native language.

1.2 – Research Questions

In order to guide my attempt to reach the objectives of this study, as mentioned in the previous sub-section, I posed the following research questions:

- 1) Do advanced EFL readers and proficient readers of English as their native language have higher marks when taking a reading test within a given time limit or when they take the same test in a no time limit condition?

2) Do advanced EFL readers under the effects of test-wiseness have higher marks and finish their tests faster than advanced EFL readers and/or proficient readers of English as their native language who are both being exposed to the test for the first time?

1.3 – Motivation for the study

Two factors motivated me to conduct this study. Firstly, I shall not avoid mentioning that there was a somewhat personal reason for having chosen to study the assessment of reading comprehension through the use of proficiency tests, especially focusing on time allocation and test-wiseness. As a matter of fact, when I myself took a proficiency examination in EFL some years ago, I experienced the great pressure of time when doing the reading part of that examination. Also, having worked at a school in which EFL proficiency examinations were applied, several times I heard examinees leaving the exam rooms after having taken the reading papers of most proficiency examinations, commenting on the little time that they had to perform all the tasks. Those examinees included both candidates who had taken specific courses preparing them for the examinations that they were taking and candidates who had not. Since then, the issues of time allocation and test-wiseness have intrigued me.

The other factor that guided my choice for this research field was my taking two courses in reading comprehension at PPGI (Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente) at UFSC. The courses were *Constructing Meaning from Text* and *Individual Differences in Reading Comprehension*, both of each were given by my adviser in this study, professor Dr Lêda Maria Braga Tomitch. I then had the opportunity to become familiar with several pieces of research that somewhat approached the topics in which I had become interested. Therefore, I can say that the

motivation for this study is the result of interest that stemmed both from personal experience in taking proficiency examinations in EFL, and from the academic study of reading comprehension.

1.4 - Significance of the Study

This study has as its main intention to be an attempt to better understand whether and to what extent the allocation of time in tests of reading comprehension affects subjects' performance, as expressed by their test scores, and, in turn, test results interpretations. Therefore, this study may have pedagogical implications as to how the issue of time allocation is tackled in EFL classrooms, be it in regular classes or in testing situations.

In a broader sense, the results of this study might also bring about a discussion on the design of some EFL examinations, especially in what concerns the reading papers of such examinations, in terms of how they approach the trade-off between test takers' speed of performance and their comprehension achievements as expressed by their test scores. At this point, it may be enlightening to mention that, according to Hughes (1989), it was mainly due to feedback from teaching professionals that the Test of Written English (TWE) was designed, back in 1986, as a supplement to the Test of English as a Foreign Language (TOEFL). While the present study certainly does not intend to be so audacious as to want to promote any changes in the design of any examinations, it does intend to throw some light on the discussion of the issues to which it relates.

This study is also relevant for it discusses the implications of test-wiseness. As Bachman (1990) and Alderson (2000) point out, minimising as much as possible the effects of test method in language tests is of the essence. Hence, this issue of test-

wiseness should be taken into account by any researcher who intends to conduct research to better understand the outcomes of reading comprehension by using reading tests. For instance, if we are to compare the reading achievements of two different readers as measured by their scores in a given reading test, we must control for task familiarity effects. This means that every effort must be made to ensure that, all other factors being equal, neither of the two readers takes advantage over the other by being so familiar with the task types in a given test to the point of having developed specific strategies to perform such tasks.

1.5- Organisation of the thesis

This thesis is divided into five chapters, namely *introduction, review of literature, method, results and discussion, and concluding remarks, limitations, and suggestions for further research.*

In the present chapter, I tried to contextualise the study that I conducted. To achieve such a purpose, I presented some theoretical background, the objectives of my study, and the questions that I intended to answer along the study. I also mentioned the factors that motivated me to start this research, as well as the significance I attribute to it.

The review of literature, presented in chapter II, summarises the literature that I found most relevant to support this study. Sub-sections of this review include studies on the models and component processes of reading comprehension, on individual differences in reading comprehension, and on the assessment of reading comprehension.

Chapter III presents the method used in the data collection and analysis of the study in as detailed a manner as possible, describing the participants, the materials, the procedures for data collection, and a framework for data analysis.

In chapter IV, I present and discuss the results of the study. The discussion presented in chapter IV stems from the analysis of the collected data, in light of the research questions that were asked in this introductory chapter.

Finally, in the last chapter, I present some concluding remarks by summarising the main findings of this study. In this concluding chapter, I also discuss some limitations of the study and suggest possible directions for further research in the area.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter I present, in four different sections, some review of the literature on reading comprehension, which constitutes the theoretical basis of this study. In the first section, I will briefly discuss the most influential models of reading comprehension, as well as the component processes involved in text comprehension, which will help me to operationalise a view of reading comprehension to which I subscribe in this study. Given that the present study relates to the assessment of the reading performance of advanced EFL readers belonging to groups of different school background, and of native English speakers, I found pertinent to provide a section on individual differences in reading comprehension. This second section is divided into three sub-sections, to approach individual differences in declarative knowledge, the issue of how individuals differ in their reading speed, and the discussion of some research conclusions on how similar and how different it is to read in a first language from reading in a second or foreign language. The third section in this chapter is about the assessment of reading comprehension, in which I provide some chief concepts and definitions to research in this field. I do not intend to exhaustively review any of the aforementioned topics. What I will offer here, thus, is an overview of what I consider the most relevant references for the purposes of this study. Thus, I will also present a fourth section, in which I will attempt to interconnect the issues discussed in this chapter, pointing to their relevance to the present study.

2.1- Models and Component Processes in Reading Comprehension

Many researchers have reviewed the different models of reading comprehension (e.g., see reviews by Bernhardt, 1991; Carrell, Devine & Eskey, 1988; Clapham, 1996;

Davies, 1995; Nuttal, 2004; and Scaramucci, 1995) and many others have talked about the component reading processes (Daneman, 1991; Gagné, Yekovich & Yekovich, 1993; Gernsbacher, 1997; Just & Carpenter, 1987; O'Brien, 1995; to mention a few). Therefore, very little, if anything at all, that any other reviewer of those models and components say about them will be completely new. Yet it seems to me that reflecting on the most influential models of reading comprehension, as well as taking into account what is known about the component reading processes will help me to operationalise a definition of reading comprehension to which I may refer throughout this study. I shall now start talking about the models.

Davies (1995) suggests that the term 'model', in the context of reading comprehension refers to a well-structured theory of what takes place when someone processes a text. She also proposes that a first step in trying to evaluate the distinct models of comprehension is by paying attention to the factors in which the models are similar. The basic premise present in all models seems to be that for text comprehension to take place, the process starts with the reader having access to the visual information in the text. Then, in processing that information, the reader will contribute, according to some models to a larger, and to other models to a smaller extent, with his/her own (background) knowledge to finally achieve text comprehension (Davies, 1995; Kintsch, 1998). The three most influential reading comprehension models are those of Gough (1972), Goodman (1969), and Rumelhart (1977). I will next briefly describe and comment on each of these models.

2.1.1- Bottom-up Models

According to the model proposed by Gough (1972), reading comprehension starts from the lowest process of letter identification, followed by the identification of the sounds of the letters, and then by the recognition of words and their meanings. After

that, then, the reader will establish the relation between the different words identified in a sentence, until the meaning of a whole sentence finally occurs. This model is called ‘bottom-up’ because it proposes that reading comprehension is processed sequentially in that manner, from the lowest level to the highest level processes, with meaning achievement being only the final product in the process. It is also called ‘data-driven’, due to its emphasis on the words in the text and, conversely, its lack of emphasis on what the reader can contribute to the process (Carrell, Devine & Eskey, 1988; Clapham, 1996). This model was developed on the basis of research with fluent adult L1 readers.

The main criticisms to the model developed by Gough are the lack of emphasis on higher level processes such as inference making on the basis of the reader’s background knowledge (see review by Clapham, 1996), and the fact that, having been based on studies of fluent adult L1 readers, it may not apply to beginning readers, nor to L2 readers, as it may intend (Davies, 1995). Also, according to Urquart and Weir (1998), “Gough’s model of the reading process is a model of the **reading aloud** process” (p. 40), thus not necessarily applying to the more common practice of silent reading.

Bottom-up models of reading comprehension seem to give support to L1 literacy approaches called ‘phonic approaches’ (Davies, 1995). In Brazil, phonic approaches to literacy were common until the beginning of the 1980s; by then, authorities in education decided to subscribe to other, more ‘global’ approaches, which seem to be supported by top-down models like the one proposed by Goodman (1969). Top-down models are discussed in the next sub-section.

2.1.2- Top-down Models

As opposed to bottom-up models, the models called ‘top-down’ propose that readers have a central role in the process of reading, and that from the very beginning of

a reading process, a reader executes the higher level processes of predicting and making inferences on the basis of their background knowledge, paying minimal attention to the low-level processes involved in decoding (e.g., readers should not worry about establishing letter to sound correspondences) (Clapham, 1996; Davies, 1995). Although Goodman (1969) has usually been referred to as the main proponent of a top-down model of how reading occurs, Urquart and Wier (1998) claim that perhaps the credit should go to Frank Smith with his 1971 model, extended in 1973.

It is known that Goodman developed his model from a psycholinguistic perspective, based on research with beginning L1 readers. The suggestion was then that the lower level of decoding is rather unimportant, and that reading has to do with making hypotheses and going to the text in order to confirm the hypothesised predictions. Models like this are called top-down, or ‘conceptually-driven’, especially due to their emphasis on the active role of readers in executing higher-level processes such as inference making, attributing a relatively minor role to the decoding of printed symbols, which should only happen if necessary.

According to Capovilla and Capovilla (2004), in Brazil, the official method to teaching literacy since the beginning of the 1980s is based on a constructivist perspective, which seems to be supported by a top-down view of reading. Such an approach has proved not to be efficient and there is now a strong movement to take up a phonic method to literacy, which Capovilla and Capovilla claim to be the solution to several problems that emerged from the so-called ‘global’ approach to literacy that has been used for over 25 years.

It seems that either extreme, that is, subscribing to a purely bottom-up model, assuming that a text presupposes an only meaning in itself, or to a purely top-down model of reading comprehension, in which the text is at times ignored, would not be the ideal path to follow. A model that could account for the complexity of the reading

process, involving both bottom-up and top-down processing would seem more reasonable. That is what Rumelhart (1977) proposes, and what I will comment on in the next sub-section.

2.1.3- Interactive Models

Starting from the premise that bottom-up models fail to acknowledge that even the comprehension of small units, such as letters and words, as well as the comprehension of sentences, require higher level processes such as the use of background knowledge and inference making, Rumelhart (1977) proposes that both bottom-up and top-down processes are needed to account for a complete model of reading comprehension. According to his interactive model, reading is a cognitive as well as a perceptive process. Therefore, the idea is that readers simultaneously process information from different sources, what means that both bottom-up and top-down processing take part in reading (for detailed reviews see Carrell, Devine & Eskey, 1988; Davies, 1995; and Scaramucci, 1995).

Nowadays, most researchers agree that an interactive model gives a better account of the reading comprehension process (Carell, 1988; Clapham, 1996; Eskey & Grabe, 1988; Nuttall, 2004; Urquart, & Weir, 1998). In that respect, Block (1992) says that there is a tendency for theorists no longer to debate about whether a bottom-up or a top-down model of reading is the ideal one. Urquart and Weir (1998) seem to be in line with this perspective, adding the fact that, in psychology, it has been acknowledged that bottom-up processes do play a role in the low level processes of decoding and lexical access. That would justify the fact that the phonic method to teaching literacy, which is said to be supported by a more bottom-up view of reading, has become popular again in countries such as France, England and the USA, and in Brazil there has been a strong movement in favour of such a method (Capovilla & Capovilla, 2004).

As I proposed at the beginning of this section, I do not intend to deepen the discussion about the different models of reading comprehension; rather, I briefly presented the most influential models in the hope that such a reflection will be a first step to lead me towards a better understanding of how a reader might achieve comprehension of a text. For those who might be interested in reading more about reading models and text comprehension, I recommend, besides the references that I have already mentioned, the works of Alderson (2000), Anderson and Pearson (1988), Goodman (1970), Jones (1995), Kintsch and van Dijk (1978), O'Brien (1995), Retorta (2001), (Smith, 1978); Spiro and Myers (1984), and van Dijk and Kintsch (1983).

Following this brief discussion on the main models of reading comprehension, in the next sub-section, I tackle the component processes known to be involved in reading comprehension, after what I will try to present my own view of what it means to say that text comprehension has or has not been achieved by a given reader.

2.1.4- About the Component Reading Processes

Many authors have discussed the component processes of reading comprehension and, among those, the comments I will attempt to make here are especially related to my readings of Daneman (1991); Gagné, Yekovich and Yekovich (1993); Gernsbacher (1997); Just and Carpenter (1987); Lorch, Klusewitz and Lorch (1995); and O'Brien (1995). Instead of commenting on the perspective presented by each of those researchers, I will rather try to capture in this brief discussion what might be considered consensus among the issues they have discussed.

It is now known that a regular reading process starts with the readers' eyes fixating the words in a text, and that the outcome of a successful reading process is a coherent mental representation of the text read. Also, there is agreement among researchers that after having accessed the literal meaning of the words in a text,

inferential comprehension may take place, from the interaction between the readers' background knowledge (e.g., general world knowledge, what might include knowledge about the language of the text, knowledge of how texts are organised, and specific knowledge about the text content) and the information found in the text being read. The so-called high level process of comprehension monitoring is believed to be executed especially by proficient readers, who may have to re-start the reading process when finding inconsistencies in the texts they read, in order to try to establish coherence to such texts.

It seems to me that discussing the components involved in a reading process is an attempt to break down such a complex process as reading is, in order to better understand it. In this respect, I particularly like the words of Lorch, Klusewitz and Lorch (1995), when they say that:

Some reading situations entail the construction of a coherent text representation; some reading situations involve a search for specific information; other reading situations implicate study skills (e.g. memorization); and still other situations emphasize aesthetic purposes for reading. (p. 376)

At this point I will attempt to state how I myself view text comprehension. Of course my own view of reading is highly influenced by my readings of some of the aforementioned authors, and I now invite the reader to check that in the next subsection.

2.1.5- A View of Reading

In the first place, in line with Aebersold and Field (1997), Rumelhart (1977), and Tomitch (1991), among other authors, I think of reading comprehension as a cognitive activity which first of all involves a reader actively processing a text, what means that meaning is *assigned* to a text, as a result of the interaction between reader and text, rather than *extracted* only from the text itself. The interaction that takes place during

reading comprehension thus involves the processing of what the reader finds in the text, and the parallel and/or sequential processing of stored information the reader already has and with which the reader can establish relations in the attempt to understand, according to his or her purposes, the text being processed. The stored information, the reader's background knowledge, may refer to general world knowledge, including knowledge about the language of the text, knowledge of how texts are structured and more specific knowledge, for instance, knowledge of the text content.

In trying to define reading, one central aspect involved in all reading comprehension processes must be taken into account. It is the issue of reading purposes (see Aebersold & Field, 1997; Davies, 1995; Just & Carpenter, 1987; Nuttal, 2004; and Pressley & Afflerbach, 1995). Considering the several different types of texts and reading situations that may exist, it is reasonable to assume that a reader does not process all kinds of texts in the same way. For instance, readers do not normally read comic books with the same attentiveness that they are supposed to apply to the reading of academic texts for study purposes.

To conceive of reading as an interactive process between reader and text, and to assume that such a process is underlain by the reader's purposes for reading a given text, as in the perspective I have just presented, leads one to accept, as discussed by Gagné, Yekovich and Yekovich (1993), Just and Carpenter (1987), and Spiro and Myers (1984), that readers do differ in the way they process texts. Therefore, in the next section, I will approach this very issue of individual differences in reading comprehension.

2.2- Individual Differences in Reading Comprehension

There seems to be consensus in the literature that comes from research on reading comprehension, that skilled reading involves many component processes and

extensive declarative knowledge (Just & Carpenter, 1987). Researchers have also investigated the reading processes of “good” and “poor” readers, and there is evidence to claim that skilled readers perform better than less-skilled readers in almost all component reading processes (Clapham, 1996; Daneman, 1991; Gagné, Yekovich & Yekovich, 1993; Just & Carpenter, 1987). Another cognitive aspect in which individuals do differ and which is known to affect their reading performance is working memory capacity. The term working memory refers to memory which has storage as well as processing functions, a system of limited capacity (Daneman, 1991; Just & Carpenter, 1987; Tomitch, 1996, 1999-2000). Apart from the cognitive components involved in reading a text, there are other, non-cognitive factors that may influence the way a reader processes a text. For instance, issues related to social background, the text genre and wording, and the reading situation itself are known to affect reading processes (Aebersold & Field, 1997; Just & Carpenter, 1987; Zwaan & Brown, 1996).

To suit the specific purposes of this study, I divided this section into three sub-sections, in which I will approach only the issues of individual differences in reading comprehension that I understand as most closely related to the present study. Therefore, in the first sub-section, I will focus on what researchers have reported about individual differences in declarative knowledge. In the second sub-section, I will refer to issues of reading speed and reading time. Then, in the final sub-section, I will concentrate on differences in L1 (first language) reading as opposed to reading in an L2 (second language) or FL (foreign language).

2.2.1- Individual Differences in Declarative Knowledge

Declarative knowledge relates to conceptual understanding, that is, it relates to a reader’s knowledge about different elements involved in the reading of any text. It involves knowledge about small units of meaning such as letters and their

corresponding phonemes and morphemes, as well as knowledge about more complex units of meaning such as words, the content of a text, and possible schemas related to the reading of any particular text (Gagné, Yekovich & Yekovich, 1993). As to the issue of differences in declarative knowledge, Gagné, Yekovich and Yekovich discuss research evidence that shows that skilled reading is highly correlated to a reader's declarative knowledge. Therefore, declarative knowledge is considered a source of individual differences in reading, and it is believed to contribute to all the processes involved in reading a text. Presumably, the low-level processes of decoding and literal comprehension are fostered by a reader's declarative knowledge.

When discussing research on domain knowledge (i.e., knowledge about the content of a text, which usually refers to specialised knowledge in a certain field), Ferstl and Kintsch (1999) observed differences in a variety of reading tasks between readers with high domain knowledge and readers with low domain knowledge. According to the researchers, the effect of knowledge in the comprehension processes of different readers can be controlled either by choosing readers known to possess significantly different amounts of domain knowledge on a certain subject, or by providing readers with different amounts of prior knowledge on the same subject. Nutall (2004) and Alderson (2000) agree with the perspective that readers' background knowledge facilitates their comprehension. Thus, as Nuttall points out, any measure of comprehension which do not take that fact into account may be biased, favouring better comprehension for those readers who possess more of such knowledge.

Contributing to the discussion about the correlation between declarative knowledge and text comprehension, Clapham (1996) calls attention to the fact that, besides differences in declarative knowledge, the reading of a text is also dependent on the reader's purposes in reading the text. Adding to this discussion, Pritchard (1990) claims that "reading is a content-specific activity" (p. 291), what suggests that readers

do not process all texts in the same way. When the idea is to measure reading comprehension and not necessarily knowledge about any specific subject, Clapham suggests that research be conducted with groups of participants that are as homogeneous as possible in terms of background knowledge, so as the researcher would attempt to control for the effects of content familiarity. I would add to her suggestion that all efforts be made to ensure that those participants in the homogeneous groups also have as similar purposes as possible when performing the same reading tasks. Research that takes into account this discussion of individual differences in declarative knowledge and in reading purposes may help us to understand why a reader who is considered good in one context may prove not to be so skilled in another reading situation.

Finally, the work of Smith (1978) may be suitable to help me to conclude this sub-section on the implications of declarative knowledge in reading comprehension. According to Smith, the amount of time a reader may need to process the visual information found in a text depends on how much information the reader will need to process that information. It is reasonable to assume that the more information a reader has already activated at a certain point in the reading of a text, including knowledge brought by the reader to the reading process, the less information will be required to process the text in focus, and hence, less time will be needed. The issue of reading time will be discussed in the next sub-section.

2.2.2- Reading Speed

From a brief overview of several research reports that approach the issues of reading speed and time spent on reading, it seems that there is some controversy as to what it means to read fast or slowly. Several researchers in the field of reading comprehension have claimed that fast reading strongly correlates, positively, with ease of text comprehension, and conversely, longer reading times have been reported as

reflecting greater difficulty in constructing meaning from text (Gagné, Yekovich & Yekovich, 1993; Jones, 1995; Jordan, 1997; Just & Carpenter, 1987; Smith, 1978). Up to this point, there seems to be agreement among most researchers, and as a beginning researcher, I dare to include myself among them. What I see as conflicting, though, is to define good reading as fast reading, disregarding non-cognitive factors involved in a reading process, such as the reading situation, or the particular structure and/or wording of a text. For instance, some of those aforementioned authors (e.g., Gagné, Yekovich & Yekovich, 1993; Jones, 1995) who understand fast reading as an indicator of good flow of comprehension, and many others (e.g., Afflerbach, 1990; Block, 1986, 1992; Carver, 1997; Davies, 1995; Kintsch & van Dijk, 1978; Murray, 1995; Narvaez, van den Broek & Ruiz, 1999; Nuttal, 2004; O'Brien & Myers, 1999; Paris, Wasik & Turner, 1991; Pressley & Afflerbach, 1995; van den Broek, Ridsen & Husebye-Hartman, 1995; Williams, 1988) have identified reading situations in which a reader taking longer to read a text might have meant strategic reading, a behaviour expected to be exhibited by skilled readers.

According to O'Brien and Myers (1999), when a reader is able to identify a local coherence break in the text being processed, and if this reader attempts to re-establish coherence to the text, for example by rereading parts of it or even the whole text from the very beginning, the reading process of such a text certainly slows down. Going further into this perspective, Pressley and Afflerbach (1995) argue that when global coherence has not been achieved, the reader may need to reread the text in order to complete the mental model of that text, which in turn means that the processing of such a text will take longer. Still talking about rereading, it is important to reiterate, at this moment, that such a procedure of identifying that the comprehension of a text has not been achieved, and deciding to take some measure to solve the problem, for instance by rereading the text, has been characterised as a behaviour reflecting strategic reading.

Such a behaviour is expected to be performed by more skilled readers (Gagné, Yekovich & Yekovich, 1993; Paris, Wasik & Turner, 1991;).

Taken together, these findings seem to indicate that not always does fast reading mean skilled reading. For instance, if a text does not contain enough elements to allow the reader to establish a coherent mental representation of it, and if, even though, the reader quickly finishes reading it, this reader may end up miscomprehending such a text, or at least, parts of the text. On the other hand, when reading the same “difficult” text, a reader who is able to notice that it is hard to build a coherent mental representation of it, and who decides to spend more time on it, examining it more carefully, seems to be behaving as a skilled reader.

Reflections on a reader’s standards of coherence or on a reader’s decision to apply some strategies when noticing that a text is particularly difficult, bring back the discussion of the implications of reading purposes. Davies (1995), for example, argues that if the goal of reading is the learning of text content, for example for study purposes, a reader should read more slowly and take more time to reflect on the text content. Therefore, depending on their goals for reading, readers may take longer times to read, to guarantee that they will achieve better comprehension of a given text, what I believe may be the case in testing situations. Conversely, they may read faster if their goals are not related to study purposes or to any other purposes which require careful reading. This discussion converges to what has been usually referred to as a trade-off between reading speed and accuracy in reading comprehension (Gagné, Yekovich & Yekovich, 1993; Just & Carpenter, 1987). Such a trade-off implies that, at times, readers may choose to sacrifice comprehension in order to speed up their reading processes, especially when they believe they do not need to understand the text in depth; at other times, the same readers may decide to take longer and get the most of their reading. However, when such a trade-off is not the result of a conscious decision, it may mean

that readers will take longer to read their texts, in the hope they will better comprehend the texts, what may not happen for different reasons, for example due to a lack of relevant prior knowledge.

At this point, it may be suitable to mention what Nuttall (2004) says about the issue of reading time. According to Nuttall, good readers should ideally demonstrate flexibility and adjust the speed with which to read a text depending on their reading goals. A fact is that several researchers in the field of reading comprehension, when carrying out their experiments, ask readers to process texts one sentence at a time, and do not allow readers to have access to previously read portions of the text (e.g., Daneman & Carpenter, 1980). I certainly understand that the design of research must take into account, first of all, the purposes of the research, and the authors who design their experiments in the way that I have just mentioned, certainly have their reasons for doing so. I am, by no means, criticising any such studies. My wonder, however, comes from a perspective formed by the discussions posed by all the authors who claim that after a single exposure to a text, the reader may not have completed a mental model of the text content (e.g., Jones, 1995). Second, depending on factors such as the reading purpose (Nuttall, 2004), the level of difficulty of the text (Murray, 1995), and the reader's standards of coherence (O'Brien & Myers, 1999), a reader may decide and may in fact need to take longer to inspect a text, what may imply the use of strategies such as rereading.

Perhaps if replications are conducted, with two basic changes, of at least some of the experiments in which readers are instructed to read a text each sentence at a time, and in which readers are not allowed to reread previously read text, some light might be shed on this issue. The changes I recommend for those replications are then: (1) that more time and opportunity be given to the participants in those experiments, and (2) that the readers who may find it necessary, are allowed to use the extra time to reread any

portions of the text being processed. It may be the case that, having ample time to read a text, and being allowed to use this time for rereading when finding it necessary, the outcomes of the reading comprehension of the participants in such studies may turn out to be rather different from otherwise.

2.2.3- Reading in L1 versus L2/FL Reading

I shall now mention one more time that in the present study I was interested in analysing some aspects as to how the issues of test-wiseness and time allocation may influence the reading outcomes of both Brazilian EFL readers and readers of English as their native language. Therefore, in this sub-section, I will focus only on the similarities and differences in readers' performance due to their condition of being either readers who are native speakers of a given language or foreign readers of the same language.

For a brief introduction to the discussion of how similar or different the reading processes of native speakers and FL readers may be, I shall refer first to Clapham's (1996) review of related research. Clapham has reviewed some major studies on the comparison of L1 and L2 or FL reading comprehension processes. She observes that there are researchers who believe that L1 and L2 readers process texts in the same way, whereas other researchers claim that not everyone who is a good reader in their native language succeeds in reading in a foreign or second language.

According to Clapham, one point of agreement among researchers is that the reading processes shown by low proficient L2 readers differ from those of native speakers. However, although many researchers argue for the need of a certain threshold level of L2 language proficiency before it is possible to transfer L1 reading skills to the efficient reading of L2 texts, it is still not clear what determines such threshold level (Tomitch, 2002). For Alderson (2000), the linguistic threshold level will vary according

to task demands, which in turn depend on issues related to text content and wording, the reader's background knowledge, and the task type itself.

Another issue raised by Clapham is that there is also widespread agreement among L2 reading researchers that the reading comprehension process of L2 readers is affected by their (lack of) background knowledge, although when it comes to state *how* background knowledge affects SL reading performance, there seems to be less agreement.

From the evidence discussed by the aforementioned studies, and also in accordance with Zwaan and Brown (1996), it seems reasonable to expect that if the performance of FL readers can at all be comparable to that of native speakers, this would happen when those FL readers achieve a high level of proficiency in the target language. Moreover, Alderson (2000) claims that, despite the fact that reading knowledge in L1 and target language knowledge are both important, there is evidence to say that in L2 reading, knowledge of the target language contributes more to the reading process than reading proficiency in the L1.

2.3- The Assessment of Reading Comprehension

The present study can be viewed not only as one making use of assessment instruments to investigate reading comprehension, but also as a study *into* the assessment of reading comprehension. Therefore, I dedicate this section to discuss some issues related to research on the assessment of reading comprehension.

The terms *assessment* and *testing* have been used interchangeably by some researchers (Allison, 1999), and they will be used as such in this study unless indicated otherwise. In the case of reading comprehension, the reasoning is that reading tests are used to assess people's reading ability, which means, in this case, that reading

comprehension can be assessed through the use of reading tests. While some may claim that *assessment* does not only or always include *testing*, given that there are other, less formal ways of assessment (e.g., continuous and self-assessment), it seems that the term *testing* is always associated with the assessment of some ability.

The literature in language assessment, or language testing, is vast, and even in what concerns specifically the assessment of reading comprehension, it would be beyond the scope of this review to discuss all related issues. Therefore, in the subsections that follow, I will briefly discuss only the issues on the assessment of reading comprehension that are most closely related to the objectives of this study, namely *validity*, *reliability*, *test-wiseness*, and *time allocation*.

2.3.1- Validity

Validity has been often referred to as the most crucial feature in language testing in general (see Alderson, Clapham & Wall, 1995; Bachman, 1990; Bernardt, 1991; Henning, 1987). The first and foremost question that test developers or users must ask in relation to a test is, therefore: *Does the test test what it is supposed to test?* Extending on that, Douglas (2000) claims that “validity is about interpretations of test scores in the light of the purpose for which the test was developed” (p. 257). From that, it is reasonable to assume that a test may be valid for certain purposes, but not for others, or put it another way, that what is actually valid or not, is not the test itself, but the interpretations we attribute to the outcomes of a test, most commonly expressed by test scores.

According to Weir (1993), a test which is designed to measure a certain ability should measure that specific ability only, that is, as far as possible, a test should not measure anything else apart from what it was designed to measure. The implication of this is that a test may not be considered valid if it measures abilities other than that or

those for which it was intended. Once it is recognised that a test taker's performance may be affected by factors other than what is intended to be measured by a test, all efforts should be made to control for such factors. In a reading comprehension test, for instance, a test which brings one single task based on a single text may not be valid if the intention is to use such a test to measure the reading comprehension abilities of different readers, some of whom may have considerable larger amounts of background knowledge on the topic of the text.

The concept of validity is one that can be approached by different perspectives, that is, there are different types of validity. In the words of Allison, "validity is best considered as a unitary construct, though with many facets" (p. 14). Examples of the different types of validity include *construct validity*, which refers to the content of a test in relation to the concept of the ability to be measured; *concurrent validity*, which has to do with comparing one measure with other measures of a same ability; *face validity*, which usually involves the judgement of test users as to what a given test looks like (see Alderson, Clapham & Wall, 1995; Allison, 1999; and Henning, 1987).

One point of agreement among researchers discussing test validity is that a test which is not reliable cannot be considered valid. Reliability is then the next issue to be approached.

2.3.2- Reliability

Reliability seems to be best understood as the extent to which a test is consistent in measuring the ability it intends to measure (Allison, 1999; Bernhardt, 1991; Lado, 1961). This would imply the verification of whether a test measures a certain ability in the same way from time to time. For instance, would a test taker have the same results when taking the same test for the second time, given that such a test taker has received no instruction in the meantime? If so, then the test is said to be reliable.

Certainly, the question of reliability is not a simple one, especially because one cannot account for all variables that are known to exert effect on test performance. For instance, while we can try to minimise the effects of background knowledge and test-method effects (e.g., by elaborating tests with different texts and tasks), how can we possibly control for factors such as health conditions, tiredness, or lack of interest or motivation? As Bachman (1990) claims, the more we minimise the effects of any factor other than the ability of interest in a test, the more we are maximising the reliability of the test. For Henning (1987), reliability is associated with test length, with longer tests having the potential to be more reliable than shorter tests. Such reasoning would favour, hence, the inclusion of different tasks in a single test to measure the same general ability from different perspectives.

Finally, several researchers have ascertained that, as it is the case with validity, reliability is also obviously related to test scores. The distinction is, though, that while validity is associated with the interpretations and uses of test scores, reliability is an essential quality of test scores themselves (Bachman, 1990; Wier, 1993).

2.3.3- Test-wiseness

Among the factors that are known to have the potential to affect test performance (see Bachman, 1990), test-wiseness is one of the two with which I am particularly concerned in this study. The other is *time allocation*, to be discussed in the next sub-section.

Test-wiseness has been associated with task familiarity (Aebersold & Field, 1997; Allison, 1999) and it may be best understood as “awareness of how best to approach particular types of question” (Allison, 1999: 214). Concerning the issues of validity and reliability previously discussed, test-wiseness may in fact represent a threat to the interpretations made from test scores. For instance, in a study involving the use of

multiple-choice and cloze procedures in reading tests, Retorta (2001) observed a strong effect of test-wisness. The observed effect was in the sense that the participants in her study seem to have concentrated on strategies to help them to find the correct answers to the test items, rather than on other reading strategies which might normally have been used to help them to make sense of the text. Similarly, in a study by Phakiti (2003), some participants reported the use of test-taking strategies to the detriment of reading strategies, that is, they may have better demonstrated how they approach a particular type of test task than how they would normally read a text for the comprehension of it.

The question of how familiar a test taker should be with a given test in order to possess test-wisness is not a simple one. It is reasonable to assume, however, that having been exposed to a test once only is not likely to be enough to guarantee test-wisness to any test taker.

The discussion about the issue of test-wisness could be extended to include, for example, a description of different types of tasks and tests, and to what extent they may trigger test-wisness. Suffice it to say for the moment, however, that the knowledge test takers have of how to tackle specific test tasks may affect their performance in testing situations. In this respect, what was proposed in the present study was an investigation of whether individuals known to be under the effect of test-wisness outperform, by having higher marks and being faster in the same reading tests, individuals who do not possess test-wisness.

2.3.4- Time Allocation

As I mentioned in the introductory chapter, not much has been said about the allocation of time in language tests. In testing, time allotment refers to the amount of time that test takers have to accomplish the test tasks, information that must be explicitly given to test takers (Douglas, 2000). Since reading speed has been shown to

correlate with reading comprehension, as it was discussed in chapter 1 (see also Alderson, 2000), time allocation is a delicate issue in reading tests. I have already provided the example of Block (1992), who concluded that some of the participants in her study may have been characterised as less proficient readers due to their longer reading times, what, in that case, did not mean that they were unskilled readers. According to Rankin (1970), when the content of reading material is complex, and when readers have a more demanding reading purpose, for example, when their reading processes involve critical thinking in order to better understand a complex text, then the correlation between reading speed and text comprehension will be minimal.

Perhaps allotting test time in accordance with the task purposes is a reasonable alternative (Urquart & Weir, 1998). Thus, it would make sense to allow test takers more time for tasks which require high order processes such as inference making, which may in turn require rereading. Conversely, it might be pointless to have longer time allotments for tasks which require the scanning of specific pieces of information in a text. However, at least in the case of proficiency examinations, rarely does a test contain explicit information about how much time a test taker should spend on each task; what is provided, rather, is the total amount of time allowed to perform all the tasks in a given test. Knowing only the maximum amount of time allowed to be used in a test leaves test takers in charge of deciding how much time to allocate for each test task, a decision which in itself may be related to test-wiseness. For instance, test takers who are highly familiar with test tasks, may be more aware of which tasks require more time and which can be accomplished faster, an advantage that the ones not familiar with test tasks will not have.

Finally, as to what concerns reading in a second or foreign language, Alderson (2000) discusses research evidence for the fact that even readers at an advanced level of

proficiency in a foreign language may not read with the same ease with which they read in their native languages. Having this concern in mind, I decided to include a group of readers of English as their native language in this study. The idea was to verify whether any of the groups of participants, the EFL readers or the readers of English as their native language, would profit from having no time limit as opposed to taking their tests within a given time limit.

Having discussed the assessment of reading comprehension in this section, as well as individual differences, and models and components of the reading process in the previous sections, I shall now present a brief final section where I will try to relate the literature discussed in this chapter to the present study.

2.4- Relating the Issues Reviewed to the Present Study

As I have said at the beginning of this chapter, I proposed to review here some of the vast literature related to reading comprehension. Thus, according to the objectives stated in chapter 1, I concentrated here on the subjects that I believe may serve the purpose of constituting the theoretical basis of this study. I shall reiterate at this point that this study proposed to investigate two different issues that relate to the testing of reading comprehension. The first one relates to the implications of time allocation in a reading test, and the second is the issue of test-wiseness related to the same reading test. How I planned to achieve those objectives will be explained in detail in the next chapter. For the moment, I shall mention that the discussion I presented in the present chapter, stating how I view reading comprehension, and also approaching some issues on individual differences and on the assessment of reading comprehension, intended to provide the theoretical grounding for this study.

CHAPTER III

METHOD

In this chapter, I will explain in detail how this study was conducted, in terms of the materials, participants and procedures involved in the process of data collection and in the analysis of such data. As I do this, I will also attempt to make it clear what parameters guided and supported my choices as to the aforementioned elements. Therefore, I may, at times, refer back to some of the theoretical background that I presented in the previous chapter. In order to be as reader-friendly as possible, I will also use some of the sub-sections in this chapter to describe and explain materials and procedures that do not directly constitute part of the data collection itself, but that refer to details which underlie the procedures involved throughout all the process of data collection.

3.1- Participants

A total of 63 people participated in the first part of this study, that is, they participated in the phase in which I selected the participants who were at an advanced level of reading proficiency in English, and who were, therefore, apt to take part in the second phase of the study, when I actually collected the data to be analysed. The 63 participants were divided into four different groups. Three of these groups were composed of Brazilian EFL readers, while the participants in the other group were all readers of English as their native language. Most of the Brazilian EFL readers who participated in this study were enrolled in courses in which English was the target language; a requirement being that everyone taking such courses must comprehend as well as produce the target language (i.e., English) fluently. The EFL readers who were

not enrolled in courses like those, had already concluded EFL courses at advanced levels, and some of those participants were also EFL teachers. From the sixty-three participants who took the *readiness test* (see sub-section 3.2.2 in this chapter), only 40 were actually considered to be at an advanced level of reading proficiency in English, which was the required level of reading proficiency for the purposes of this study. The four groups of participants will be henceforth referred to as groups 1, 2, 3, and 4. I will next describe each group in as much detail as possible.

3.1.1- Group 1 – EFL Students Preparing to Take the CAE Examination

The participants in group 1, in the first phase of the study, that is, the phase in which I selected the actual participants from whom I later collected the data, were 13 adult EFL students preparing to take the CAE (Certificate in Advanced English) examination. At the time of the data collection, they all had been taking a regular CAE course for at least the second consecutive semester. Therefore, in this study, the participants in group 1 were the ones considered to be very familiar with the task types present in the reading paper of the CAE examination.

To say that someone is regularly enrolled in a course preparing to take a proficiency examination such as the CAE, usually means that they had either gone through a placement test previously to their entering the course, or that they have been promoted from one level lower than that, for instance, by having passed an examination at an immediately lower level, such as the FCE (First Certificate in English). The participants in group 1 of this study were all included in one of those two cases. Nine of these participants were then students at *Phoenix Centro Joinvilense de Cultura e Idiomas*, in Joinville – SC, henceforth referred to as *Phoenix*, and the other four were students at *Sociedade Brasileira de Cultura Inglesa de Florianópolis Ltda.*, henceforth

referred to as *Cultura Inglesa*. From the thirteen participants in this group, 9 were considered apt for the second part of the study.

3.1.2- Group 2 – Advanced EFL Readers Not Familiar with the CAE Examination

Group 2 was formerly composed of 19 Brazilian adults who were also fluent EFL readers but who were not preparing to take any proficiency examinations in EFL; neither were they acquainted with the format of the reading paper of the CAE examination. In this group, twelve of the participants were or had been EFL teachers in different language schools in Joinville. Four of the participants in this group were regularly taking an EFL course at an advanced level at a language school in Joinville. Finally, the other 3 participants had already concluded their EFL courses at advanced levels in different language schools. *Seven* out of the nineteen participants in this group were considered advanced EFL readers, according to the readiness test, and therefore were apt to participate in the second part of the study.

3.1.3- Group 3- EFL Students Taking an MA course in English Language and Literature

Eighteen Brazilian adults were included in group 3, as prospective participants in the data collection phase. The participants in this group were all regularly taking the MA course in English Language and Literature at the *Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente* (PPGI) at UFSC (Universidade Federal de Santa Catarina). All the subjects offered at PPGI/UFSC are given in English, that is, the language used in those courses is English, with all the required readings for such courses being in English. Therefore, the participants in group 3 were expected to have a high level of reading comprehension in EFL. At the time this selection of participants happened, most of them were or had been EFL teachers, although they were not acquainted with the tasks present in the reading paper of the CAE examination.

Fourteen participants from this group were apt to take part in the subsequent phase of the study.

3.1.4- Group 4 – Readers of English as their Native Language

Group 4 was first composed of 13 adult English native speakers from different nationalities. Eight of these participants lived in Joinville and five of them lived in Florianópolis at the time they participated in this selection phase. Although most of them were or had been teachers of EFL in different language schools in Joinville, Florianópolis and/or in different countries, none of them was familiar with the reading paper of the CAE examination. Three of these participants did not achieve a satisfactory score to be considered advanced English readers according to the pre-established parameters for this study (as explained in subsection 3.2.2). Thus, in the second part of the study, group 4 was composed of 10 participants.

3.2- Materials

In this section, I will describe the three different materials that I used in this study. The data which I am going to analyse and discuss in the next chapter refers to the outcomes of participants taking a complete reading paper of a past application of the CAE examination; I will henceforth refer to that material as *the CAE reading paper*. However, prior to the data collection itself, a *readiness test* was used with the purpose of identifying the participants who were at the required level of reading proficiency in English for this study, namely advanced level; this test will henceforth be referred to as *the readiness test*. Still another reading material was applied to some of the participants before they actually took the readiness test. This material will be henceforth referred to as *the practice material*. From these materials, only one refers to the actual data

collection instrument in this study, and it is also the most complete of the three materials in terms of the tasks of which it is composed. It is the CAE reading paper. For these reasons, I will describe it first, in sub-section 3.2.1. The other two materials, namely *the readiness test* and *the practice material*, will be described respectively in sub-sections 3.2.2 and 3.2.3. I decided not to describe the materials in a conventional chronological order because I believe this will facilitate the comprehension of the uses and purposes of these materials. For instance, I believe that having first read the description of the tasks that are present in the CAE reading paper, as well as having well understood the purpose of such a test in this study, the reader will more quickly and easily understand the components and the purposes of the other two materials, and not otherwise. To my own understanding, it seems that it only makes sense talking about the practice material, and its purpose is better justified, when the reader has a clear understanding of the readiness test and its purpose in this study. In the same way, what justifies the presence of the readiness test here is the fact that a test is being used to measure reading comprehension at a given level of proficiency, what is discussed in more detail in the first sub-section. I do hope that it makes as much sense to the readers of this thesis. In the final sub-section, 3.2.4, I will describe in detail how the readiness test and the CAE reading paper were graded.

3.2.1- The CAE Reading Paper

The questions I proposed to investigate in this study are related to the performance of EFL and native English readers at an advanced level of reading proficiency. Therefore, for a matter of validity, I had to either create or choose an EFL reading test at that level. For some reasons which I will attempt to explain next, I decided to use, as the data collection instrument in this study, the reading paper from a past application of the CAE examination. This examination is at level four of the

Cambridge ESOL (English for Speakers of Other Languages) examinations, what means, according to information found in a CAE handbook from 2003, that it "...is designed to offer a high-level qualification in the language to those wishing to use English for professional or study purposes" (p.6).

In line with the issue of validity discussed in the previous chapter, one of the main reasons why I decided not to design a specific test to suit the purposes of this study was due to practical aspects. For instance, given that I am not an expert EFL test writer nor can I easily contact any such professionals, and also because of time constraints, it would be certainly too hard for me to design and pilot a test until I could say it was reliable to a safe degree and I could then use it properly. Thus, I chose a reading paper from a past application of the CAE examination (see appendix A, on p. 86). This test had supposedly been piloted and had actually been applied with the very purpose of identifying whether test takers were at an advanced level of EFL reading proficiency according to the examination specifications.

The fact that the CAE reading paper is composed of three different task types, therefore attempting to test different reading skills, was also a contributing factor guiding my choice. I am in favour of this kind of test because research has indicated that the use of different task types is expected to minimise method effects, therefore maximising the validity of the test (see Aebersold & Field, 1997; Alderson, 2000; Alderson, Clapham, & Wall, 1995; Nutall, 2004; and Urquart, & Weir, 1998). The choice for such a test was then also an attempt to control for the effects of task familiarity as well as for test-wiseness, in the hope that all participants would be able to perform to the best of their English reading ability.

In terms of format, the CAE reading paper contains four parts: two *multiple matching* tasks (parts 1 and 4) aiming at specific information from short thematically related texts; one *gapped text* task (part 2) which focuses on text structure, and a

multiple choice task (part 3), whose main focus is “detail, gist, opinion/attitude” (CAE Handbook, 2003: 9). The questions in parts 1 and 4 are worth one mark each, and each question in parts 2 and 3 carries two marks.

In part 1 of the CAE reading paper used as the data collection instrument in this study, there were 6 short texts on the topic of *films on video*, and there were 15 multiple matching questions which could relate to any of the texts. The texts in part 1 were all *film reviews* of video releases. Part 4 presented a magazine article with five short texts about *a writer's day*, and 17 multiple matching questions addressing any of the texts. In each part, the task consisted in matching every question to the short text(s) to which the test taker thought the question related.

Part 2 was composed of a magazine article under the title *Where the landscape will do the walking*, from which six paragraphs had been removed. On the page following the text, there were seven paragraphs from which six were the ones that had been removed from the text, and one worked as a distracter. The task consisted in inserting the six removed paragraphs back to the text, therefore not including the distracter paragraph.

As to the multiple choice task, the participants had to read a magazine article about an artist, and then answer seven multiple choice questions. For each question there were four possible answers from which the test takers could choose only one.

3.2.2 – The Readiness Test

In order to investigate the performance of individuals at an advanced level of reading proficiency in English, a *sine qua non* condition was to have, as participants in this study, individuals at that specific level of reading proficiency. Therefore, a readiness test (Bachman, 1990) was applied to all the possible participants, prior to their taking the CAE reading paper. According to Bachman, when a given test has the

purpose of selection, for example when a test intends to verify whether or not students are at a required level of ability to enter a course for instructional purposes, this test may be called a “readiness test”. In this study, I propose to extend this meaning of readiness tests, to include the purpose of verifying whether or not individuals are ready to take a test at a given level of proficiency.

Again, instead of designing a brand-new test, I decided to count on previously tested tests, especially due to the same aspect of practicality involving reliability and validity issues as explained in the previous sub-section. This time, though, the choice was not for a complete test; I rather chose specific tasks from the reading papers of two different examinations.

With the purpose of both minimising method/practice effects and at the same time including task types which the participants would find in the CAE reading paper, three different tasks were chosen out of the reading papers of two different examinations (see appendix E, on p. 101). The reasoning for the latter explanation is that having worked with such task types in the readiness test would ensure that all the participants would have been somewhat familiar with the tasks they were going to find in the CAE reading paper. This is reasonable having in mind that the participants in groups 2 (the advanced EFL readers not familiar with the CAE reading paper), 3 (the EFL students taking an MA course in English language and literature), and 4 (the readers of English as their native language) had not had any previous contact with a CAE reading paper.

The tasks which composed the readiness test were then one *multiple matching* task, one *multiple choice* task and one *gapped text* task. The first and the third tasks were taken from the reading paper of a past application of the FCE (First Certificate in English, by Cambridge ESOL Examinations) examination. Each question in the first task was worth one mark, and the questions in the third task carried two marks each. The multiple choice task was taken from a reading paper of a past application of the

CAE examination, different from the CAE reading paper which was the data collection instrument in this study. Each question in this task carried two marks.

Following the advice given by Henning (1987) (see also Alderson, Clapham, & Wall, 1995), when deciding about the validity of the proposed readiness test, I asked for the help of test users as well as of experienced proficiency examiners who had been working with Cambridge ESOL examinations for over 10 years. In order to reach an agreement as to what score in the readiness test would indicate that the participants were at the advanced level of reading proficiency in English, we took into account two main factors. The first one was that the FCE examination is at one level lower than the CAE examination, that is, the FCE examination is addressed to individuals at an intermediate stage of proficiency. The other factor relates to our consideration of the fact that a score of around 60% gives candidates a passing grade in both examinations. We then agreed that the participants who achieved a score of 90% (which tends to be considered an “A” grade) or above in the tasks from the FCE reading paper, and 60% or above in the task at the CAE level, would be considered “ready” to take a whole CAE reading paper. Those participants might be characterised as being at an advanced level of reading proficiency in English.

However, to make matters a little more complex, when the tests were graded I noticed that, unexpectedly, some participants who did not achieve the minimum set score of 90% in the tasks from the FCE examination, got 100% in the task at the CAE level, which was supposed to be at a higher level of complexity. I then had some more discussions with the experienced teachers and examiners who were helping me, and I asked advice from a professional in statistics as well. In the end, we agreed that, since the tasks from the two different levels were now part of a same test, it would be fair to consider a balance between the outcomes of the performance of the participants in each part of the test. Therefore, still considering those high and low scores of 90% and 60%

as previously described, for the purposes of this study, every participant who got a mean score of 75% or above in the whole readiness test was considered to be at the advanced level of reading proficiency in English. That is, they would be ready to take the CAE reading paper.

3.2.3- The Practice Material

When talking to the participants previously to the data collection, and even previously to their taking the readiness test, I found out that some of them had never had any previous contact with multiple matching or gapped text tasks. Therefore, those participants were provided with some practice material in order for them to get acquainted with those tasks even before they took the readiness test. This procedure was in agreement with Urquart and Weir (1998) and Weir (1993) who argues that “every attempt should be made to ensure that candidates are familiar with the task type and other examination features before sitting a test” (p. 152).

The practice material consisted of two tasks from the reading paper of a past application of the FCE examination (different from the ones used in the readiness test): one multiple matching task and one gapped text task (see appendix G, on p. 108). Those participants who needed the practice material worked on their own with the two tasks, and I provided them with an answer key for each task.

3.2.4- Grading the Tests

As it was previously mentioned, all reading materials used in this study were taken from actual reading papers from past applications of Cambridge ESOL examinations. The papers were gently granted by *Cultura Inglesa*, which is an authorised Cambridge ESOL examination centre. Given that the authorised centres which apply the examinations overseas do not normally have access to the answer keys

to the papers, I myself did not have access to them, either. Therefore, in order to grade the participants' tests, I had to prepare an answer key for each of the two tests I used. To achieve such a purpose, I certainly took both tests, but this might not have been reliable enough for research purposes. Thus, in order to increase the reliability of the answer keys that I would have to prepare, I invited seven EFL teachers to collaborate as raters in this study. Their task was to take the tests and inform me of the answer they attributed to each question, as well as to discuss in order to reach agreement in case there were any discrepancies among the answers given by the different raters. All of the raters were experienced EFL teachers and five of them were actually oral examiners for the Cambridge ESOL examinations; three of these were native English speakers. These raters were the same professionals who helped me to work on the readiness test specifications, with the exception of the statistician, who is not in the field of EFL teaching.

I must report here that it took longer than I expected to have the raters take the tests and discuss until we could reach an agreement as to the correct answers. Thus, because of this time constraint, and having in mind that all the participants were expected to have a good level of reading comprehension in English, as described in the previous section, I asked all the participants who took the readiness test to take the CAE reading paper as well. Otherwise it could happen that some of the participants might not be available when I finally needed them for the second and most important part of the study, that is, the data collection.

3.3- Procedures

In this section, I will explain first the procedures used in the application of the readiness test. In the sequence, I will explicitly present the results of the readiness test.

Finally, I will describe each step in the data collection process, as well as I will give a brief account of how the data from this study were analysed.

3.3.1- Taking the Readiness Test

The objective of the readiness test was to verify whether test takers were at the required, advanced level of reading proficiency in English. I was not interested, at that moment, in how fast the participants could finish doing their tests. Therefore, I made it more of a power test (Alderson, 2000) and allowed participants to take as much time as they needed to accomplish the three tasks in the test.

In some instances, as it was the case with some participants from group 1, and some more from group 3, due to constraints concerning participants' availability, more than one participant took the readiness test simultaneously and, in those instances, in the same room. However, in most of the cases the readiness test was taken by one participant at a time. In all cases participants worked on their own and were not allowed to check any information with anyone else, nor were they allowed to use dictionaries or any material other than a pen or a pencil.

Because neither the content nor the format of the test tasks were modified in any way from what is presented in the original papers, some instructions were given orally before the participants started the readiness test. For instance, in the test sheets, after each instruction for a task completion there is the following statement: "Mark your answers on the separate answer sheet", which the participants were not supposed to do since no answer sheet was provided. They were then instructed to mark their answer for each question in the corresponding spaces provided in the test sheets. I also instructed the participants to ask me any questions whenever they were in doubt about how to approach any of the tasks. Moreover, the very first instruction the participants were given orally was to carefully read all the instructions before completing their tests.

3.3.2- Results of the Readiness Test

In this sub-section, I will present the results of the readiness test for each group of participants. To achieve such a purpose, four tables will be displayed which show the number of participants in a sequence (i.e., from 1 to 63), and the participants' mark in each part of the test (i.e., a mark related to the tasks at the FCE level, as well as a mark relating to the task at the CAE level). The table also shows the actual result of the test, that is, whether each participant was considered ready (R) or not (NR) to participate in the second part of this study.

The following table, Table 1, presents the results of the readiness test for group 1, that is, the EFL readers preparing to take the CAE examination.

Table 1: *Results of the Readiness Test for Group 1*

Participant	FCE Tasks	CAE Task	Mean	Result
1	48,81	33,33	41,07	NR
2	89,28	100,00	94,64	R
3	92,86	100,00	96,43	R
4	59,52	83,33	71,43	NR
5	92,86	66,66	79,76	R
6	96,43	83,33	89,88	R
7	85,71	66,66	76,19	R
8	72,62	83,33	77,98	R
9	77,38	66,66	72,02	NR
10	76,19	100,00	88,10	R
11	77,38	83,33	80,36	R
12	100,00	100,00	100,00	R
13	51,19	66,66	58,93	NR

As the data in Table shows, only nine out of the 13 prospective participants from group 1 were considered apt to take part in the data collection process.

As to the advanced EFL readers not familiar with the CAE examination, group 2, from the 19 prospective participants only seven were considered ready to take part in the data collection in the present study, as shown in Table 2 below.

Table 2: Results of the Readiness Test for Group 2

Participant	FCE Tasks	CAE Task	Mean	Result
14	71,43	83,33	77,38	R
15	88,10	66,66	77,38	R
16	41,66	16,66	29,16	NR
17	50,00	66,66	58,33	NR
18	52,38	50,00	51,19	NR
19	55,95	16,66	36,31	NR
20	82,14	33,33	57,74	NR
21	89,28	100,00	94,64	R
22	76,19	66,66	71,43	NR
23	64,28	66,66	65,47	NR
24	48,81	33,33	41,07	NR
25	51,19	66,66	58,93	NR
26	89,28	66,66	77,97	R
27	92,85	50,00	71,43	NR
28	96,43	83,33	89,88	R
29	71,43	83,33	77,38	R
30	89,28	83,33	86,31	R
31	52,38	83,33	67,86	NR
32	45,24	66,66	55,95	NR

The following table presents the results of the readiness test for group 3, which was composed of EFL students taking an MA course in English language and literature. From the data displayed in the table, it is possible to visualize what was briefly commented in sub-section 3.1.3, when I explain that fourteen out of the eighteen readers from group 3, who took the readiness test, were considered ready to participate in the data collection phase of this study.

Table 3: Results of the Readiness Test for Group 3

Participant	FCE Tasks	CAE Task	Mean	Result
33	59,52	66,66	63,09	NR
34	100,00	83,33	91,67	R
35	89,28	66,66	77,97	R
36	76,19	100,00	88,10	R
37	88,10	83,33	85,72	R
38	84,52	66,66	75,59	R
39	100,00	100,00	100,00	R
40	100,00	83,33	91,67	R
41	75,00	83,33	79,17	R
42	72,61	50,00	61,31	NR
43	47,62	33,33	40,48	NR
44	100,00	66,66	83,33	R
45	84,52	83,33	83,93	R
46	85,71	100,00	92,86	R
47	69,05	50,00	59,53	NR
48	96,43	83,33	89,88	R
49	92,85	83,33	88,09	R
50	92,85	66,66	79,76	R

Finally, the results of the readiness test for group 4, the readers of English as their native language, are presented in Table 4. It was already commented, in subsection 3.1.4, that 3 out of the 13 readers of English as their native language were not considered ready for the data collection according to the purposes of this study. Therefore, group four was composed of 10 participants in the data collection phase.

Table 4: *Results of the Readiness Test for Group 4*

Participant	FCE Tasks	CAE Task	Mean	Result
51	92,85	100,00	96,43	R
52	79,76	83,33	81,55	R
53	96,43	100,00	98,22	R
54	100,00	100,00	100,00	R
55	40,48	66,66	53,57	NR
56	89,28	100,00	94,64	R
57	96,43	100,00	98,22	R
58	71,43	66,66	69,05	NR
59	79,76	83,33	81,55	R
60	100,00	100,00	100,00	R
61	92,85	66,66	79,76	R
62	100,00	100,00	100,00	R
63	48,81	33,33	41,07	NR

Now that the purposes and the results of the readiness test used in this study were presented and discussed, it is time to explain in detail how the actual data for this study was collected. Therefore, in the next sub-section, I will present information about the two different conditions in which the readers who were considered ready to participate in the data collection process of this study took the CAE reading test.

3.3.3- Taking the CAE Reading Paper

In the second part of the study, participants took a complete reading paper from a past application of the CAE examination. Given that the CAE reading paper is the actual instrument of data collection in this study, each of the four groups of participants was divided into two sub-groups to suit the purposes of this investigation. In this way, the EFL students preparing to take the CAE examination, group 1 in this study, were divided into subgroups 1.1 and 1.2. The group of advanced EFL readers not familiar with the CAE examination was divided into subgroups 2.1 and 2.2. In the same way, group 3, the ELF students taking an MA course in English language and literature were divided into subgroups 3.1 and 3.2. Finally, the readers of English as their native language, group four in this study, were also divided into two subgroups, namely subgroups 4.1 and 4.2. At this point, it is worth reiterating that besides comparing the performance of the different groups of participants, I also intended to make some intra-group comparisons. I would like to observe how participants from the same group would perform in two different conditions, namely in a *time limit* condition, henceforth also referred to as *condition 1*, and in a *no time limit* condition, that is, *condition 2*. Therefore, in order to control for test method effects and maximise the internal validity of the study (Howell, 1992), the participants in each group (1, 2, 3, and 4) were randomly assigned to each of the conditions (e.g., participants from group 1 were randomly assigned to be part of either subgroup 1.1 or 1.2, and so forth).

Participants in *condition 1* worked in as similar a manner as possible to test takers in actual examination conditions. They worked individually and had a maximum time of 1hr and 15 min to conclude their tests, which also included the transferring of the answers to a separate answer sheet. This time of 1 hr and 15 min was chosen because it is the time limit test takers are allowed in the actual examination conditions.

A timer was used both to control for the maximum time limit and also to record the time that each participant actually spent to accomplish the whole test.

In *condition 2*, participants worked individually, too, but they could spend as long as they needed to conclude their tests. Oral instruction was given to ask the participants in this condition to disregard the instruction, in the test sheet, that told them to spend a maximum time of 1hr and 15 min. As in condition 1, again, a timer was used to register the total time spent by each participant who, as the ones in the other condition, also had to transfer their answers to a separate answer sheet. In this condition, the time for the transferring of the answers to the answer sheet was also counted in the total time spent by each participant on the test, as in condition 1.

3.3.4- Framework for Data Analysis

Thinking of the research questions that I asked in the introductory chapter of this thesis, I analysed the data from this research in three main ways. There were two types of inter-group comparisons and one of intra-group comparisons. The inter-group comparisons were made with the purpose of comparing the mean scores in the CAE reading paper of the participants in group 1, with the mean scores of each of the other groups in the same test. Therefore, the mean scores of subgroup 1.1 were compared to the mean scores of each subgroup in condition 1, and the mean scores of subgroup 1.2 were compared to the mean scores of each subgroup in condition 2. One of the mean scores compared refers to the participants' marks in the test, and the other score refers to the mean time spent by each subgroup to conclude their tests. The intra-group comparison aimed at observing whether it was in condition 1 or in condition 2 that each of the four main groups of participants had higher mean marks in the CAE reading paper. As to statistical methods, since the *t* test is considered appropriate for verifying

whether there is statistical significance between two means (see Howell, 1992; and Woods, Fletcher & Hughes, 1986) , I opted for this statistical test.

CHAPTER IV

RESULTS AND DISCUSSION

In this chapter, I will present and discuss the results of this study. The first and foremost objective of this chapter is, certainly, to answer the research questions that were posed in chapter 1. However, as each of the research questions is answered, the discussion will be also extended to include some comments about the general tendencies that were observed during the data analysis. Still while discussing the results of this study, I will try to relate them to those of some of the studies that were reviewed in previous chapters. After the research questions have been discussed and the observed tendencies have been commented, a section will be included with two purposes. The first one is to tackle the issue of the readiness test. Although it was not the main concern in this study, it was surely an important part prior to the data collection process. I will therefore discuss the implications of the readiness test used in this study, an issue that I understand as relevant for research in language testing, more precisely, in the testing of reading. The other purpose is to make a general concluding discussion trying to put together all the issues that will be discussed in this chapter. According to the aforementioned overview of this chapter, and in the hope to make it as clear as possible to the reader, it is divided into four sections. In an attempt to minimise any complexities that may be related to the design of this study, in the first section I will briefly restate the objectives of the study and how I actually collected and analysed the data to achieve each of them. In the second section, I will present and discuss the results of the study that are pertinent to research question 1. The third section is dedicated to answering research question 2. Finally, in the fourth section, I will briefly discuss the implications of the readiness test in this study in one sub-section, as well as I will try to make a final general discussion including all the issues tackled in this chapter in the final sub-section.

4.1- Summarising the Objectives and the Design of the Study

In this section, I shall elaborate more on each of the research questions that I posed in the first chapter before I actually answer them in the subsequent sections. Given that the research questions are obviously related to the objectives of the study, I will also reemphasise the objectives as well as I will briefly refer to some of the underlying theoretical background that supports this study. This section is divided into two sub-sections, each one approaching the objective related to one of the two research questions, and also clarifying how I collected and analysed the data to answer each of the two questions.

4.1.1- Objective and Design Concerning Research Question 1

The first objective of this study relates to the issue of time allocation in tests of reading comprehension. As I discussed in chapters 1 and 2, there are many researchers who found evidence to say that fast reading strongly correlates with ease of text comprehension (e.g., Gagné, Yekovich & Yekovich, 1993), and conversely, that slow reading would indicate unskilled reading, that is, more difficulty in the process of text comprehension (e.g., Jones, 1995). However, there are those who claim that not always does fast reading indicate good text comprehension (e.g., Rankin, 1970), or that slow reading cannot always be related to unskilled reading (e.g., Block, 1992).

Based on my readings of what some research outcomes have indicated concerning the related issues of reading speed and time allocation in reading tests, I found pertinent to ask the following research question: *Do advanced EFL readers and proficient readers of English as their native language have higher marks when taking a reading test within a given time limit or when they take the same test in a no time limit condition?* That was the first question, which I will approach more deeply in the next

section. For the moment, I shall just elaborate a little more on what exactly I meant with such a question, and on how I collected and analysed the data to answer it.

As to the issue of time allocation, as I mentioned in the first chapter, it is known that many test takers complain, at least informally, about the time they are allowed to spend on the reading papers of examinations such as the CAE. The usual complaint is that those test takers usually find that they have too strict a time limit to be able to perform all the test tasks and to fill in their respective answer sheets. Therefore, I thought it could be enlightening to conduct an experiment with the reading paper of one such examination in two different conditions. In the study I am reporting, the participants in condition 1 had exactly the same time limit as they would have in actual examination conditions, that is, 1 hr and 15 min. The participants in condition 2, on the other hand, did not have any time limit at all.

The procedure applied in condition 2, of having no time limit to conclude the test, is the proposal of *power tests*. Power tests intend to measure the ability for which they are designed, not taking into account the time test takers spend to demonstrate such ability, that is, allowing test takers to spend as long as they need to conclude their tests (see Allison, 1999; Bachman, 1990).

As I have informed in the method chapter, there were four different groups of participants in this study. Therefore, in order to compare the performance of the participants in each of the two conditions, I divided each of the four groups into two subgroups. That is why I had subgroups 1.1 and 1.2, 2.1 and 2.2, 3.1 and 3.2, and 4.1 and 4.2. I shall mention thus, that the comparisons that were made with the purpose of answering research question 1 were intra-group comparisons, that is, I compared the scores of the participants in a same group in the two different conditions. For instance, the mean score of subgroup 1.1, in terms of the participants' marks in the test, were compared to that of subgroup 1.2; subgroup 2.1 was compared to subgroup 2.2 and so

on. In this way, I could verify whether the participants in each of the four groups had higher marks when ample time was given or whether allowing them to spend as much time as they needed did not make any difference in the outcomes of their tests.

4.1.2- Objective and Design Concerning Research Question 2

As a second objective of this study, I intended to investigate the implications of test-wisness in a reading comprehension test. I am using the term test-wisness just as Allison (1999) presents it, that is, relating it to task familiarity effects, in the sense that being very familiar with task types may help test takers to approach such tasks more strategically. In order to achieve this second purpose, I could either give special treatment, in terms of task familiarity, to one of the groups of participants, or include a group of participants known to possess test-wisness for they were very familiar with the task types included in the test I would use. The latter option was chosen. At this point, it is worth reinforcing the assumption that one or two exposures to a given test task will not normally be enough to make a test taker become test-wise in the sense that test-wisness has been defined in this study.

Once test takers know how best to perform the tasks in a test, such test takers might be expected to finish their tests faster and perhaps to achieve better results than test takers who do not possess such test-wisness. Therefore, the implications I attempted to investigate with research question 2 relate to the mean marks and to the mean time spent by the participants in the CAE reading paper. Thus, the second research question in this study reads as follows: *Do advanced EFL readers under the effects of test-wisness have higher marks and finish their tests faster than advanced EFL readers and/or proficient readers of English as their native language who are both being exposed to the test for the first time?*

Given that I intended to compare the scores of the participants under the effect of test-wiseness to those of the participants in the other three groups, who were not so influenced, the comparisons established in order to answer research question 2 were inter-group comparisons. Therefore, the mean scores, both in terms of the participants' marks, and of the total time they spent on the CAE reading paper, of subgroup 1.1 were compared to those of subgroups 2.1, 3.1, and 4.1. In the same way, the mean scores of subgroup 1.2 were compared to those of each of the other subgroups in condition 2, that is, subgroups 2.2, 3.2, and 4.2.

I shall reiterate at this point, that every participant in this study was timed, even the ones in condition 1, in which the participants were only allowed a maximum time limit of 1 hr and 15 min. This timing procedure was a necessary one if the aforementioned comparisons were to be established. Besides that, there was also the intention to verify whether the participants in condition 1 in this study would actually be able to finish their tests within the established time limit. In addition, another intention was to observe whether any of the subgroups would spend a mean time which was shorter than the 75 min that they would be allowed in actual examination conditions. Since all the comparisons made in this study were comparisons between two means, a *t*-test was used as a statistical tool to verify whether any of the differences between the means that were compared was statistically significant.

4.2- Results and Discussion for Research Question 1

In this section, I will first explicitly answer research question 1. After that, I will extend the discussion to include comments on the general tendencies that can be observed from the analysed data. This section is thus divided into two sub-sections. In the first sub-section, which I will call *Answering research question 1*, as the title

suggests, I will answer that question based on the statistical analyses that were made with the collected data. In sub-section 4.2.2, which I will call *Observed tendencies in relation to time allocation*, I will still refer to research question 1, but this time with the purpose of focusing on the general tendencies observed in this study. The comments made in this sub-section may shed some light on the general discussion of the issue of time allocation in tests of reading comprehension.

4.2.1- Answering Research Question 1

It is now time to answer research question 1, which asked: *Do advanced EFL readers and proficient readers of English as their native language have higher marks when taking a reading test within a given time limit or when they take the same test in a no time limit condition?* As I stated in chapter 3, and reiterated in the previous section, all the comparisons that were made in this study refer to the mean scores of each subgroup of participants. Therefore, in order to present the relevant data to answer research question 1, I will display a table that presents the following information: a) the number of participants (N) in each subgroup; b) the mean score (M) of each subgroup in terms of the participants' marks in the CAE reading paper; the standard deviation (SD) of each mean; and the coefficient of variance (CV) for each subgroup. This information is presented in Table 5 on the next page.

Table 5: Mean Mark of each Subgroup in the CAE Reading Paper

<i>Subgroup</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>CV (%)</i>
1.1	5	55,52	15,55	28,01
1.2	4	65,09	24,82	38,13
2.1	3	53,44	5,08	9,50
2.2	4	56,90	14,68	25,80
3.1	8	70,90	14,77	20,83
3.2	6	71,55	14,59	20,39
4.1	6	81,03	6,30	7,77
4.2	4	86,21	9,04	10,49

From the information presented in Table 5, one aspect which may call the reader's attention first, is the fact that all subgroups in condition 2 had higher marks than their counterparts, that is, the subgroups in condition 1. This might have indicated that, having had plenty of time to perform all the tasks in the CAE reading paper used in this study, may have really contributed for the participants with such an advantage to achieve higher marks than their counterparts. That would mean that the answer to research question 1 would be "yes". However, when it comes to analysing those results to verify whether the differences found are statistically significant, we have to accept that the answer may be just the opposite. I said "may be" because the precise answer to the question involves the establishment of a level of significance. In this study a significance level of .05 ($p < .05$) was set. I shall next present the precise results of the t -tests for the comparison between the mean marks of each subgroup in the CAE reading paper.

The first comparison verified whether there was any statistically significant difference between the mean marks of the participants in group 1, that is, the EFL students who were then preparing to take the CAE examination, in the two different

conditions. If we look at Table 5, we find the following numbers related to subgroup 1.1 (i.e., the participants in group 1 who took the test in condition 1, within the time limit) ($M = 55.52$, $SD = 15.55$), and ($M = 65.09$, $SD = 24.82$) are the figures for subgroup 1.2 (i.e., the participants in group 1 who took the test without a specified time limit, that is, in condition 2). Accordingly, the results of the t -test for this comparison is $t(7) = .62$, $p = .27$. Having found a p value higher than the .05 which was set indicates that the difference in the comparison of these scores did not reach statistical significance.

As to group 2, which included the advanced EFL readers not familiar with the CAE examination, the t -test showed that the difference between the mean marks of the participants in the two conditions was even less statistically significant than that found between subgroups 1.1 (related to condition 1) and 1.2 (participants in condition 2). The data displayed in Table 5 reveals that, again, it was in condition 1 that the participants, now from group 2, had lower marks ($M = 53.44$, $SD = 5.08$). Obviously then, subgroup 2.2 had a higher mean mark ($M = 56.90$, $SD = 14.68$), $t(5) = .33$, $p = .37$.

The difference that was least statistically significant was that found for group 3, the Brazilian students taking an MA course in English language and literature. The participants in this group achieved the following numbers when taking the CAE reading paper in condition 1 ($M = 70.90$, $SD = 14.77$). Their counterparts, that is, subgroup 3.2 (i.e., the ones in condition 2) obtained similar results ($M = 71.55$, $SD = 14.59$). The t -test comparing the performance of these two subgroups, as expressed by their mean marks in the CAE reading paper, presents the following result: $t(12) = .08$, $p = .47$.

Finally, a comparison was established with the mean marks of the participants in group 4, that is, the readers of English as their native language. We can clearly observe that the difference between the mean marks of subgroups 4.1 and 4.2 in the CAE reading paper was the least distant from .05. Although in both conditions (condition 1= time limit; condition 2= no time limit) the participants in group 4 had higher marks than

any of the other subgroups, again, in condition 1, their mean mark was lower ($M = 81.03$, $SD = 6.30$) than that of subgroup 4.2 ($M = 86.21$, $SD = 9.04$). Therefore, the result of the t -test for the comparison between the mean marks in the CAE reading paper of subgroups 4.1 and 4.2 are as follows: $t(8) = .95$, $p = .18$.

As can be observed from the results of the t -tests presented above, the better scores that we can observe for each of the subgroups in condition 2, the no-time limit condition, were not statistically different from the scores of their counterparts, that is, the subgroups in condition 1. This means that, even though we can observe in Table 5, that all subgroups in condition 2 apparently outperformed their counterparts in terms of their mean marks in the CAE reading paper, the precise answer to research question 1 in this study is *no*. That is, it is not statistically confirmed that either advanced EFL readers or proficient readers of English as their native language have higher marks on a reading test when they are allowed to spend as much time as they need on such a test, than when they are not so allowed.

This first look at the collected data was with the sole purpose of answering research question 1. However, I would like to extend the discussion on the issue of time allocation to include some informed considerations with reference to the collected data. This discussion is presented in the next sub-section.

4.2.2- Observed Tendencies in Relation to Time Allocation

In this sub-section, I will comment on the general tendencies that were observed in relation to the results of the data analysed in this study. I will also try to relate the results of this study, in what refers to time allocation, to some of the research that was reviewed in chapters 1 and 2. The discussion in this sub-section will refer to the data displayed in Table 5, and also to the data presented in Table 6. The data in Table 6 summarises the mean marks of all the participants in the CAE reading paper in

Condition 1 and in Condition 2, irrespectively of the subgroup to which each participant belongs.

Table 6: *Mean Mark from All the Participants in the CAE Reading Paper in Each Condition*

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>CV(%)</i>
Condition 1 (Time Limit)	22	67,79	16,19	23,89
Condition 2 (No Time Limit)	18	70,11	19,40	27,68

As it is already possible to observe in Table 5, but perhaps more clearly in Table 6, in general terms, the participants in this study achieved higher marks in the CAE reading paper when they took it in Condition 2. As it was discussed in the previous subsection, however, the difference in the participants' mean scores in terms of their marks in the test was not statistically significant for the purposes of this study.

Perhaps the factor which is responsible for the lack of statistical significance of the results obtained with the data collected in this study was a lack of statistical power due to the small number of participants in each group and, in turn, in each condition. However, if we leave that statistical significance ($p < .05$) aside for the moment, it seems that we are still left with a clear tendency in favour of having plenty of time to perform all the tasks in a reading test. I will next attempt to relate the results of this study, in terms of time allocation, to other studies that were previously reviewed in chapters 1 and 2.

I shall start by saying that the reading test used in the data collection of this study was not meant to distinguish between "good readers" and "less-skilled readers", as it seems to have been the case of the reading proficiency measure in the study by Block (1992). Nor did this study investigate how the different participants behaved as they approached each of the different tasks, so as the issue of task complexity could be

discussed in some detail, as it was discussed by Jones (1995). This study also did not control for how many instances of rereading of a single passage each participant may have attempted (see discussion by Jones, 1995). Notwithstanding, I shall reiterate that the precise reason why I decided to include the no time limit condition, condition 2 in this study, was because I did take into account the issues related to reading speed and time allocation discussed by authors such as Block (1992) and Jones (1995).

As I have already discussed, there seems to be in the literature in reading comprehension, and also in language testing, different views of what it means to read fast or slowly. Thus, as to the issue of time allocation, this study was actually an attempt to verify whether the results obtained from the data analyses would corroborate any of the current views in some way.

In general terms, the results of the data analysed in this study, in relation to time allocation, seem to demonstrate a tendency that would perhaps reinforce the claim of those researchers who argue in favour of ample time for text processing. Especially the ones who suggest that more time may be needed depending on the demands of the reading passages or tasks (e.g., Murray, 1995; Rankin, 1970), and on reading purposes (e.g., Davies, 1995; Nuttall, 2004). Thus, were these results statistically significant, it would be safe to say that they corroborate the results of previous research that show that when having ample time to read their texts, readers may achieve better results in terms of comprehension. However, the results in the present study only showed a “tendency”.

According to Davies (1995), and also to Nuttall (2004), a main factor that may influence the amount of time that any reader spends on a text is their reading purposes. For those authors, when a text is read for study purposes, readers may decide to spend more time in such a text than they would for other purposes, for example when they read for entertainment. To my understanding, in a reading test situation the purposes of readers may be similar to those of reading for study purposes. I believe the purposes are

similar to the extent that both situations, that is, reading for study purposes and reading in test situations, involve a reader reading at least one text for the purpose of demonstrating comprehension of that text in some way. Therefore, at least in the case of the participants involved in this study, we can notice this tendency for better results when readers have as much time as they need to process their texts, and in this case, to answer the test questions, as opposed to when they have only a set limited time. I shall again reiterate that it is a tendency only, thus not statistically confirmed.

In what concerns the complexity of the texts present in the CAE reading paper used in this study, no analysis was conducted to inspect whether one text was more demanding than any of the others in any way. Nor was this an objective in this study. As to task demands, however, although again, no analysis of task demands was carried out, it may be enlightening to mention one more time that only one group of participants was very familiar with all the tasks used in this study. To be more precise, some of the participants in the other three groups had their first contact with *gapped-text* and *multiple matching* tasks when they took part in this study. Therefore, at least for these participants, who may have considered such tasks more demanding than the *multiple-choice* task, with which they were all very familiar, having had plenty of time to perform all the tasks may have been of some help.

Given that the results reported so far did not reach statistical significance, all aspects related to time allocation that I have discussed in this study need more empirical investigation to be confirmed. As the title in this sub-section suggests, for the moment it is all about tendencies.

4.3- Results and Discussion for Research Question 2

In this section, I will discuss the results of this study in relation to the issue of test-wiseness. Therefore, I will explicitly answer research question 2 in this section. While discussing the data in order to answer research question 2, I will also refer to some of the theoretical background reviewed in chapter 2 in order to help me to explain the results. Also, as I did in the section related to research question 1, I will extend the discussion in the present section to make some comparisons that do not directly relate to research question 2, but that may be pertinent for discussions of test-wiseness related issues. As I make these comparisons, I will also comment on an observed tendency in relation to the groups which had the highest and lowest marks. Therefore, the present section is divided into two sub-sections. In the first one, I will explicitly answer research question 2 and discuss possible explanations for it, in light of the literature reviewed in chapter 2. In sub-section 4.3.2, I will go a step further and make a few more comparisons, as well as I will also comment on an observed tendency.

4.3.1- Answering Research Question 2

As I discussed again at the beginning of this chapter, research question 2 was asked with the intention to investigate whether test-wiseness would affect test takers' results in a given reading test. The idea was first to compare the mean marks of the subgroups of participants known to be under test-wiseness effects to the mean marks of their counterparts. This was with the intention to verify whether the participants under test-wiseness effects would have higher marks in the CAE reading paper than the other participants. Also, the mean times spent by subgroups 1.1 and 1.2 were compared to the mean time spent by each of the other subgroups, according to the time condition in which they took the test. The objective was then to verify whether the participants in group 1 (the ones who were preparing to take the CAE examination) would be faster

than the other participants. Therefore, the following research question was posed: *Do advanced EFL readers under the effects of test-wiseness have higher marks and finish their tests faster than advanced EFL readers and/or proficient readers of English as their native language who are being exposed to the test for the first time?*

Before I actually present and discuss the results for research question 2, I would like to reiterate two points once again. The first is that the participants considered to be under test-wiseness effects in this study are those belonging to group 1 (i.e., the EFL students preparing to take the CAE examination). The second point is that the mean scores of the participants in group 1 were compared to those of the participants in the other groups according to the condition in which they took the test. That is, the mean scores of the participants in subgroup 1.1 (the participants in group 1 who took the test in the time limit condition) were compared to those of all other subgroups in condition 1. In the same way, the mean scores of subgroup 1.2 (the participants in group 1 who took the test in condition 2, the no time limit condition) were compared to those of all the other subgroups in condition 2.

Since research question 2 relates to two different issues, I will approach each of them separately. I will first approach the inter-group comparisons in relation to the mean marks that each subgroup got in the CAE reading paper. Then, I will tackle the inter-group comparisons in relation to the mean time spent by each subgroup on the test.

4.3.1.1- Comparing the mean marks of group 1 to those of the other groups

To answer research question 2, in what concerns the comparisons of the mean marks of the participants in group 1 to those of the other groups of participants, we shall inspect Table 5 once again. We can clearly see that subgroup 1.1 only outperformed subgroup 2.1, in terms of their mean marks in the CAE reading paper. Group 2 included the advanced EFL readers not familiar with the CAE examination.

Given that subgroup 1.1 did not outperform any of the other subgroups in condition 1, in terms of their mean marks, the difference between the mean marks of subgroups 1.1 and 2.1 was the only one which required statistical analysis to verify its significance at the .05 level. In other words, subgroup 1.1 was actually outperformed by the other two subgroups in condition 1, namely subgroups 3.1 and 4.1, which included the Brazilian students taking an MA course in English language and literature, and the readers of English as their native language, respectively.

A *t*-test was performed computing the values for subgroup 1.1 ($M = 55.52$, $SD = 15.55$) and those for subgroup 2.1 ($M = 53.44$, $SD = 5.08$), and thus presenting the following result: $t(6) = .19$, $p = .42$. Therefore the difference between the mean mark of subgroup 1.1 and that of subgroup 2.1 was not statistically significant for the purposes of this study.

As to subgroup 1.2, it only outperformed one of the other subgroups in condition 2, namely, subgroup 2.2. In other words, subgroup 1.2 was then outperformed by the other two subgroups, that is, subgroups 3.2 and 4.2. Thus, again, the only inter-group comparison which required statistical analysis to verify its significance was that involving groups 1 and 2, more precisely, between subgroups 1.2 and 2.2.

The numbers related to the mean mark of subgroup 1.2 ($M = 65.09$, $SD = 24.82$) and those related to subgroup 2.2 ($M = 56.90$, $SD = 14.68$) were computed by the statistical program which performed the *t*-test. This is the result of the *t*-test: $t(6) = .49$, $p = .32$. Therefore, the difference observed between the mean mark of subgroup 1.2 and that of subgroup 2.2 is not statistically significant for the purposes of this study.

At this point, after having discussed the data presented in Table 5, and also in light of the results calculated by the aforementioned *t*-tests, it is possible to give a partial answer to research question 2. I say partial because, up to this moment, I have been discussing the results for research question 2 only in what refers to the comparison

between the mean marks of group 1 in the CAE reading paper to the mean marks of the other groups in the same test. The discussion related to the comparison between the time spent by these groups in the test will be presented in the subsequent sub-section.

For the purpose of answering research question 2 only partially at the moment, I will break that question so as the answer given matches only the proper part of research question 2, that is, the part to which the answer is actually addressed. Thus, the question being answered here should read only: *Do advanced EFL readers under the effects of test-wiseness have higher marks in a reading test than advanced EFL readers and/or proficient readers of English as their native language who are both being exposed to the test for the first time?* Not surprisingly at this point, the answer is “no”. In this study, the participants known to possess test-wiseness due to task familiarity effects did not achieve statistically significant higher scores in the reading test used, than the other participants, who were not familiar with the test tasks.

Instead of looking for possible explanations for the fact that test-wiseness did not seem to be a factor affecting test takers performance, at least in what relates to their marks in the test, I prefer to discuss it the other way around. To my understanding, the point here is what may have helped the participants in groups 3 and 4 to compensate for not possessing test-wiseness. However, since this discussion will also involve the other half of the question, I will wait to discuss it together with the issue of “reading speed” at the end of the next sub-section.

4.3.1.2- Comparing the mean times spent by group 1 to those spent by the other groups

Having answered the first part of research question 2, it is now time to answer the other half of the question: *Do advanced EFL readers under the effects of test-wiseness finish a reading test faster than advanced EFL readers and/or proficient*

readers of English as their native language who are both being exposed to the test for the first time?

While to answer the first part of the question we referred back to the data displayed in Table 5, the other part of the answer requires information other than what is presented in Table 5. The missing information refers to the mean time spent in the CAE reading paper by each subgroup of participants, and it is presented in Table 7 below.

Table 7: Mean Time of Each Subgroup in the CAE Reading Paper

<i>Subgroup</i>	<i>N</i>	<i>M (min)</i>	<i>SD</i>	<i>CV (%)</i>
1.1	5	71,80	6,40	8,91
1.2	4	65,00	7,95	7,61
2.1	3	75,00	0,00	0,00
2.2	4	94,25	13,05	13,84
3.1	8	69,50	5,61	8,08
3.2	6	82,83	18,84	22,74
4.1	6	64,83	6,52	10,05
4.2	4	74,50	16,29	21,86

The comparisons performed to verify whether the participants under test-wiseness conditions were faster than the participants in the other groups were, obviously, inter-group comparisons. However, I shall mention again that some of the participants were taking their tests in condition 1, that is, under time pressure, and some were allowed to spend as much time as they needed, namely the ones in condition 2. Therefore, the mean time spent by subgroup 1.1 will be compared only to the mean time spent by each of the subgroups in condition 1. In a similar way, the mean time of subgroup 1.2 will be compared only to the mean time of each subgroup in condition 2. It

might not be fair to compare the mean scores of subgroup 1.2 to the mean scores of any of those subgroups in condition 1, given that the latter performed their tests within a time limit, while the former had no time constraints.

Similarly to what happened in relation to the comparisons between the mean marks, in terms of reading speed, we can clearly observe that the participants in subgroup 1.1 had a lower mean time only than that of subgroup 2.1. The students taking the MA course in English, and the native readers were not at all slower than the participants under test-wiseness effects. Therefore, up to this point, the comparison between the mean times of subgroups 1.1 and 2.1 was the only one which required statistical analysis to prove or not the significance of the difference between the two means.

The numbers related do subgroup 1.1 (the EFL students preparing to take the reading test, in condition 1) ($M = 71.80$, $SD = 6.40$) were computed in a t -test, together with the numbers related to subgroup 2.1 (the advanced EFL readers not familiar with the CAE examination, in condition 1) ($M = 75.00$, $SD = .00$). The result of the test for this comparison was $t(6) = .75$, $p > .05$. Therefore, we cannot say that the participants in subgroup 1.1 were faster than those in subgroup 2.1 because the difference between their mean times was not statistically significant, according to the level of significance established for this study.

As to the subgroups in condition 2, the no time limit condition, looking at the data in Table 7 we can notice that, this time, group 1 apparently outperformed all the other groups in this condition. That is, the raw data show that the mean time spent by the participants in subgroup 1.2 was lower than that of each of the other subgroups in condition 2. We shall next see if any of these differences are statistically significant for the purposes of this study.

Computing the numbers for subgroup 1.2 ($M = 65.00$, $SD = 7.95$) and those for subgroup 2.2 ($M = 94.24$, $SD = 13.05$), we have a t -test result indicating $t(6) = 2.10$, $p < .05$. Therefore, in this case, the difference between the mean time spent by subgroup 1.2 (i.e., the participants in group 1 who took the test in condition 2, the no time limit condition) and that spent by subgroup 2.2 (the participants in group 2 who took their tests in the no time limit condition, too) was statistically significant. We can say, thus, with 95% of confidence, that in this study, the participants under test-wiseness effects were faster than those of group 2, when both groups took their tests in condition 2, that is, with no time limit.

Comparing the mean time of subgroup 1.2 with that of subgroup 3.2 ($M = 82.83$, $SD = 18.84$), we have $t(8) = 1.66$, and $p > .05$. The result of the t -test for this comparison showed that the difference between the two means is not statistically significant. This means that the participants under test-wiseness effects were not faster than the students taking the MA course in English and literature, when both groups were in the no time limit condition. That is, subgroup 1.2 was not faster than subgroup 3.2 in the sense that the difference between the mean time spent by each subgroup was not statistically significant.

Finally, when the mean time of subgroup 4.2 ($M = 74.50$, $SD = 16.29$) was computed, the t -test showed the following result: $t(6) = .56$, $p > .05$. Again, although we see different numbers expressing the mean time spent by subgroups 1.2 and 4.2, the difference between these two means did not reach statistical significance. Therefore, we cannot say, at least not with 95% of confidence, that the participants under test-wiseness influence were faster than the readers of English as their native language, when both groups took a reading test in a no time limit condition (i.e., condition 2 in this study).

4.3.2- Issues Relating Generally to Research Question 2

In this sub-section, I will discuss a few points that may relate to research question 2 as a whole, thus involving the issue of test-wiseness and its implications in a reading test.

First of all, according to what was already reviewed about test-wiseness (see Aebersold & Field, 1997; Allison, 1999), to my understanding, the biggest threat of test-wiseness can be best understood as twofold. One main concern is that, in testing situations, test takers under the effect of test-wiseness may be better able to demonstrate their ability in taking tests of a given format than the ability that the test was designed to measure. However, a second threat is that knowing how to best tackle the tasks in a test, test takers might be able to better allocate the time to perform the different test tasks. This in turn might contribute for the test takers with test-wiseness advantages to finish their tests faster than the test takers who are not so advantaged. Overall, it seems that in this study, the fact that one group of participants possessed test-wiseness did not affect the results of this group, neither in terms of reading speed, nor in terms of the marks the participants in this group got in the test they took.

I shall mention at this point that the idea to include groups 3 and 4 in this study, the MA students and the native speakers, respectively, had the very purpose of verifying to what extent test-wiseness would be an advantage to the test takers in condition 1. As it was reviewed in the section about individual differences in reading comprehension, if L2 readers are to be compared to proficient readers of a given language, we should include L2 readers at a very high level of reading proficiency in the L2.

Despite the fact that all participants in this study took the readiness test, whose results indicated whether the participants were ready or not to take the CAE reading paper, I now believe that some of the participants in this study might be beyond CAE level. For instance, most of the participants in group 3 were or had been EFL teachers

by the time the data was collected. Moreover, they were all regularly taking an MA course in English language and literature, reading a great deal of academic articles written either in EFL or in English as an L1. As to the natives, most of them were or had been teachers of EFL, too. Most of those who were teachers had already taught English at advanced levels in different language schools.

As it was reviewed in chapter 2, even readers at an advanced level of proficiency in a foreign language may not read as fast and as accurate as they do in their mother tongues (see Alderson, 2001). That perhaps explains why the participants in group four (the natives) achieved the highest marks in the test in both conditions, though they were only the fastest in condition 1.

In summary, it might be the case that, despite the readiness test, the participants in this study were not actually at the same level of reading proficiency in English. If on the one hand all participants were ready to take the CAE reading paper, on the other hand, I believe that only the participants in groups 1 and 2 were at the advanced level of proficiency required for the CAE examination. The other two groups were probably beyond CAE level. The fact that none of the groups achieved a mean mark of 95% or above in the CAE reading paper, what might be an indication that they were beyond CAE level, according to the parameters discussed for the readiness test in this study, is another issue. For the moment, the point is that ability in reading in the target language seems to have compensated for lack of familiarity with task types, at least in the case of the participants in groups 3 and 4.

Having answered the research questions that were asked in this study, there are still some comments that I would like to make before I can close this chapter. I will discuss the final issues in the next section.

4.4- Final Discussion

In this section, I will make the final comments in order to close this chapter. However, before I do this, I shall tackle, in the next sub-section, the implications of the readiness test used in this study. Thus, this section is divided into two sub-sections. In the first one, I will discuss some issues concerning the readiness test, and in the second one, I will make some final remarks.

4.4.1- Implications of the Readiness Test

As it was discussed before, the purpose of the readiness test used in this study was to verify whether the prospective participants, the ones who participated in the first part of the study, prior to the data collection, were “ready” to take the CAE reading paper. The test that was selected, as it was explained in chapter 3, served the purpose to which it was intended, and can therefore be considered valid. However, now that the data for this study was collected and analysed, I see that things could have been different.

I borrowed the concept of a readiness test from Bachman (1990), according to whom, when a test intends to verify whether students are ready for instruction at a certain level, this test can be called a ‘readiness test’. I proposed to extend the definition so as I could say that, if a test has the purpose of verifying whether individuals are ‘ready’ to take a given test, this test can be called a ‘readiness test’, too.

I still think that a readiness test is a valid type of test for the purpose of this study. However, perhaps the way I selected the test should have been different. At the time I selected the tasks to include in the readiness test used in this study, I asked the help of experienced EFL teachers who worked with proficiency examinations. We all agreed, then, that the test as it is (see appendix E, p. 101) would be a valid test for the purposes of this study. A closer look at the test, in light of the results discussed in this

chapter, make me think that the way I decided to assemble the tasks to form this readiness test would better serve the original purpose discussed by Bachman (1990). That is, I now believe it would be more appropriate to identify individuals who are ready to enter a preparatory course to take the CAE reading paper, than to identify test takers who are ready to take the CAE reading paper itself. This is mainly because in the readiness test I 'selected' there was only one task at the CAE level, the other two were one level lower than that, as explained in chapter 3.

The readiness test in this study was applied as a power test, therefore allowing test takers to spend as long as they needed to perform all the tasks in it. Thus, perhaps a better option of a readiness test for this study would have been the use of a complete reading paper from a past application of the CAE examination, different, of course, from the CAE reading paper used to collect the data. The reasoning would be that, any participant who got a mark equal or above 60 in the CAE reading paper used as a readiness test, would be considered apt to take a CAE reading paper, since 60 is a passing grade, according to the examination specifications.

We would still be left with the problem that some of the participants might be beyond CAE level. A way out could be, perhaps, to work with an average mark, between the lowest passing grade and the highest one, with anyone getting a mark lower than 60, or above the average between 60 and the highest mark, not being included as participants in the study. Then, perhaps one single readiness test would serve two different purposes, the one which was already discussed and a second one, aiming also at identifying test takers who might be beyond the desired level of proficiency.

I surely understand that having this kind of test used for research purposes is a delicate issue, as are many issues in language testing. For instance, in sub-section 3.2.2, I reported the fact some of the prospective participants who took the readiness test in this study got higher marks in the task that was meant to be testing comprehension at a

higher level than their marks in the tasks at a supposedly lower level. Would this indicate that the candidates were ready to take a reading test at an advanced level of proficiency but were not ready for a test at intermediate level? Probably not. Perhaps such a fact has to do with the level of complexity of the tasks included in the readiness test used in this study. What I can suggest at this point is that more research on the topic is needed before we can have a better position in relation to the application of readiness tests.

The discussion is not closed, but I will stop right here, since this chapter needs to be concluded, what I will do next.

4.4.2- Final Remarks

Investigating the assessment of reading comprehension is not at all a simple endeavour. Perhaps one of the most difficult tasks related to the testing of reading is to isolate the construct to be measured. The difficulty is in ascertaining that reading ability only is being measured and not other types of ability that may help a reader to process a text but which do not by themselves constitute reading ability. In this study, the implications of test-wiseness and time allocation were investigated in a test of reading comprehension at an advanced level of proficiency.

According to the results discussed in this chapter, test-wiseness did not prove to be a factor affecting testing performance. That is, the participants in this study who possessed test-wiseness (i.e., group 1) did not achieve higher marks in the test used in this study, in comparison to the participants who were not under test-wiseness effects. In one instance, though, the participants in group 1 were faster than the participants in group 2, but that does not seem to be a tendency, it was rather an exception.

As to the issue of time allocation, again, not only did the fact that not having a set time limit to complete their tests not help test takers, but also, some test takers

finished their tests before the time limit given in the actual application of the test used in this study.

More empirical investigation is certainly needed to find out more about the issues of test-wiseness and time allocation in reading tests. I hope the discussion carried out in this chapter may have contributed in some way to a better understanding of some aspects related to those issues, namely the establishment of a time limit to do a reading test as opposed to doing the same test in a no time limit condition, and the implications of being highly familiar with the tasks present in a given reading test.

CHAPTER V

CONCLUDING REMARKS, LIMITATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH

In this chapter, I will present some concluding remarks in order to highlight the main findings of the present study. I will also discuss some limitations of the study and, based on the main findings and on the limitations presented, I will as well suggest future research directions.

5.1- Concluding Remarks

As I have mentioned in chapter 1, from my personal experience in taking proficiency examinations in EFL, and from having been in contact with professionals who are involved in such examinations, I have noticed that examinees usually complain about the time limits that they are given to take the reading papers. Moreover, research has suggested that some people may have been labelled *poor readers* due to their longer reading times (Block, 1992).

The knowledge of those facts motivated me to carry out research in this field, and the present study is the outcome of my endeavours in the area. However, as I have discussed in chapter 4, a fact is that in this study, the proposed time limit of 75 min for the taking of the CAE reading paper seems to have been enough for all the groups of participants. That is, the only two groups that used more time, namely subgroups 2.2 and 3.2, did not have higher marks than their counterparts (i.e., 2.1 and 3.1), I mean, there was no statistical significance between the mean marks of subgroup 2.1 and 2.2, nor between the mean marks of subgroups 2.1 and 3.2.

Thinking of possible explanations for the results obtained in this study, one fact is that although the procedures used in the study were similar to the ones used in the actual examination conditions, the two situations certainly cannot be equated. Therefore, despite all the instructions and the formal atmosphere that I attempted to create, the participants surely knew that they were taking part in an experiment only.

In real-life situations, test takers may feel pressured not only by the time limit itself, but by other factors. For instance, they may feel pressured due to financial reasons, since they themselves or someone else has to pay in order for them to take the examinations, and therefore, failure might be faced as wasting of money. Another factor that is true for actual examination conditions, but was not the case of this study is that actual test takers usually have to handle self-expectations as well as others' expectations. For instance, in the case of EFL teachers or prospective EFL teachers taking a proficiency examination, the outcome might have implications as to those test takers getting or keeping a job.

Bearing in mind that any research could not control for all factors that might exert psychological influence over test takers' performance, nor was it my intention in this study, I shall concentrate on what I attempted to investigate in this research, that is, the implications of time allocation and test-wiseness.

Research has indicated that testing performance under time pressure cannot rival nonpressured performance (see Goodie & Crooks, 2004, for a review). However, given the results of this study, I am tempted to conclude that one of the main factors why the participants under time pressure were not negatively influenced was that they did not face that time pressure in the same way that test takers in actual examination conditions may do.

As to the issue of test-wiseness, as I have discussed in chapter 4, it did not prove to be a contributing factor for test takers achieving better results in this study. Perhaps

this could be better explained if we looked at the profile of the participants in each group. Even though all participants had to reach a satisfactory score in a readiness-test before taking the CAE reading paper, the groups of participants considerably differed in their backgrounds as readers of the English language. On average, the group of native English readers, due to the fact of being proficient native readers, were probably more experienced readers of the English language than the readers in the other groups. Group 3 (i.e., the EFL students taking an MA course in English language and literature), as I have discussed, may be considered the second most experienced group of readers of the English language in this study. Then came group 1 (the EFL students preparing to take the CAE examinations), and finally, group 2 (the advanced EFL readers not familiar with the CAE examination) was supposedly the group whose participants were the least experienced readers of English.

Therefore, in the end, the results of this study lead me to conclude that test-wiseness *per se* does not seem to exert much influence on the outcomes of an EFL reading proficiency examination. Experience with reading in the target language, what may make readers more strategic, seems to have been of major influence in this study.

5.2- Limitations of the Study

It is common knowledge that one of the main aims of scientific research is that of generalising from research results. However, due to some limitations that I will try to discuss here, it may not be safe to generalise from the results of the present study before carrying out more research which could account for the limitations that I will now present.

The first limitation that I notice for this study refers to the sample sizes of the different groups of participants on which I could count. The literature on quantitative

research recommends sample sizes of 30 participants or more for skewed populations so that the means can approximate a normal distribution (e.g., Howell, 1992). However, due to difficulties in finding participants to compose each of the groups, none of them had 30 participants.

Another factor which can constitute a limitation in this study is that of how I planned to approach the issue of test-wiseness. I had only one group under test-wiseness influence. Perhaps what I could have done to verify whether test-wiseness had any influence on the performance of the participants in the two different conditions (i.e., time limit vs. no time limit) was to have given a test-wiseness treatment for half of the participants in each group, and only then to randomly assign the participants to each of the conditions.

5.3- Future Research Directions

To close the chapter, I would like to present some directions for further research concerned with the issues that I approached in this study. The first suggestion is that other researchers carry out research similar to the study presented here, but which can account for the limitations previously discussed, that is, having a larger number of participants in each group, and a treatment condition for test-wiseness, instead of selecting participants expected to possess test-wiseness.

It might also be interesting for research on the assessment of reading comprehension through proficiency tests, to verify the same issues with other instruments, namely, using the reading papers of other examinations, such as the TOEFL (Test of English as a Foreign Language) or the IELTS (International English Language Testing System).

Finally, it might be important for a better understanding of the issue of the implications of time allocation and test-wiseness in an EFL reading examination, to conduct research with participants at different levels of reading proficiency in English. It may be the case that time allocation does play a role at lower stages of EFL reading proficiency.

REFERENCES

- Aebersold, J. A., & Field, M. L. (1997). *From reader to reading teacher*. Cambridge: Cambridge University Press.
- Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, 1, 41-46.
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, J. C. Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Allison, D. (1999). *Language testing & evaluation*. Singapore: Singapore University Press.
- Anderson, R. C., & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. In Carrell, P. L., Devine, J. & Eskey, D. (Eds.). *Interactive approaches to second language reading*. USA: CUP.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bernhardt, E. B. (1991). *Reading development in a second language: theoretical, empirical, and classroom perspectives*. New Jersey: Ablex Publishing Corporation.
- Block, E. (1986). The comprehension strategies of second language readers. *Tesol Quarterly*, 20(3), 463-494.
- Block, E. (1992). See how they read: Comprehension monitoring of L1 and L2 readers. *Tesol Quarterly*, 26(2), 319-343.
- Cambridge Examinations, Certificates & Diplomas. (2003). *CAE Handbook*. Cambridge: University of Cambridge Local Examinations Syndicate.
- Capovilla, A. G. S., & Capovilla, F. C. (2004). *Alfabetização: método fônico*. 3^a ed. São Paulo: Memnon.
- Carrell, P. L., Devine, J., & Eskey, D. E. (1988). *Interactive approaches to second language reading*. New York: Cambridge University Press.
- Carver, R. P. (1997). Reading for one second, one minute, or one year from the perspective of rounding theory. In J. Williams, (Ed.). *Scientific studies of reading. The official journal of the society for the scientific studying of reading*. Vol. 1, nr. 1. New Jersey: Lawrance Erlbaum Associates.
- Clapham, C. (1996). *The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.

- Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson. (Eds.). *Handbook of Reading Research*. Vol. II. New York: Longman.
- Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Davies, F. (1995). *Introducing Reading*. London: Penguin English.
- Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge: Cambridge University Press.
- Eskey, D. E., & Grabe, W. (1988). Interactive models for second language reading: perspectives on instruction. In P. L. Carrell, J. Devine, & D. E. Eskey. (1988). *Interactive approaches to second language reading*. New York: Cambridge University Press.
- Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In van Oostendorp & S.R. Goldman (Eds.), *The construction of mental representations during reading*. New Jersey: USA: LEA.
- Gagné, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning*. Ch. 12: Reading. pp. 267-312. New York: Harper Collins College Publishers.
- Gernsbacher, M.A. (1997). Two decades of structure building. *Discourse Processes*, 23, 265-304.
- Goodie, A. S., & Crooks, C. L. (2004). Time-pressure effects on performance in a base-rate task. *Journal of General Psychology*. Jan. 2004.
- Goodman, K. S. (1969). Analysis of oral reading miscues: applied psycholinguistics. *Reading Research Quarterly*, 5, 9-30.
- Goodman, K. S. (1970). Reading: a psycholinguistic guessing game. In H. Singer & R. B. Rudell (Eds.). *Theoretical Models and Processes of Reading*. Newark, Delaware: International Reading Association.
- Gough, P.B. (1972). One second of reading. In Kavanagh & Mattingly (Eds.). *Language by ear and by eye*. Cambridge: Mass.: MIT Press.
- Henning, G. (1987). *A guide to language testing. Development. Evaluation. Research*. Cambridge: Newbury House Publishers.
- Howell, D. C. (1992). *Statistical methods for psychology*. 3rd Ed. Belmont: Duxbury Press.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

- Jones, C. M. (1995). Construction of a mental model. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 335-352). New Jersey: LEA.
- Jordan, R. R. (1997). *English for academic purposes. A guide and resource book for teachers*. Ch. 5: Evaluation: students and courses. pp. 85-93. Cambridge: Cambridge University Press.
- Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Ch. 15: Individual differences. pp. 453-482. Massachusetts: Allyn and Bacon.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. Ch. 4: Modeling comprehension processes – the construction integration model. (pp. 93-120). New York: Cambridge University Press.
- Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review*, 85, 263-294.
- Lado, R. (1967). *Language testing. The construction and use of foreign language tests*. London: Longman Group Limited.
- Lorch, R. F., Klusewitz, M. A., & Lorch, E. P. (1995). Distinctions among reading situations. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading*. (pp. 375-398). New Jersey, USA: LEA.
- Murray, J. D. (1995). Logical connectives and local coherence. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 107-125). New Jersey, USA: LEA.
- Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. In *Journal of Educational Psychology*. Vol. 91, no. 3, 488-496.
- Nuttall, C. (2004). *Teaching reading skills in a foreign language*. Ch. 13: The testing of reading. pp. 212-228. Oxford: Macmillan Publishers Ltd.
- O'Brien, E. J. (1995). Automatic components of discourse comprehension. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 159-176). New Jersey, USA: LEA.
- O'Brien, E. J., & Myers, J. L. (1999). Text comprehension: A view from the bottom up. In S. R. Goldman, A. C. Graesser & P. Van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso*. New Jersey, USA: LEA.
- Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*. 20 (1), 26-56.
- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316.

- Paris, S. G., Wasik, B. A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson. (Eds.). *Handbook of Reading Research*. Vol. II. (pp. 609-640). New York: Longman.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly*, 25(4), 273-295.
- Rankin Jr., E. F. (1970). The relationship between reading rate and comprehension. In Farr, R. (Ed.). *Measurement and Evaluation of Reading*. New York: Harcourt, Brace World Inc.
- Retorta, M. S. (2001). Multiple-choice and cloze procedures in reading tests: what do they really measure? *The ESPecialist*. Vol. 22, n. 2. São Paulo: Centro de Pesquisas, Recursos e Informação em Leitura da Pontifícia Universidade Católica de São Paulo.
- Rumelhart, D. E. (1977). Toward an integrated model of reading. In Dominic S. (Ed.). *Attention and Performance IV*. Hillsdale, NJ: Lawrence Erlbaum.
- Scaramucci, M. V. R. (1995). *O papel do léxico na compreensão em leitura em língua estrangeira: foco no produto e no processo*. Unpublished Doctoral dissertation, UNICAMP, Campinas.
- Smith, F. (1971). *Understanding reading: a psycholinguistic analysis of reading and learning to read*. New York: Holt Rinehart Winston.
- Smith, F. (1973). *Psycholinguistics and reading*. New York: Holt Rinehart Winston.
- Smith, F. (1978). *Reading*. Cambridge: Cambridge University Press.
- Spiro, R. J., & Myers, A. (1984). Individual differences and underlying cognitive processes in reading. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson. (Eds.), *Handbook of reading research*. Vol. I. (pp. 471-501). White Plains, New York: Longman.
- Tomitch, L. M. B. (1991). *Schema activation and text comprehension*. In Fragmentos, vol. 3, no. 2, 29-43.
- Tomitch, L. M. B. (1996). *Individual differences in text organization perception and working memory capacity*. Revista da Anpoll, n. 2, p. 71-93.
- Tomitch, L. M. B. (1999-2000). Individual differences in working memory capacity and the recall of predicted elements in the text. *Linguas Modernas*, 26-27, 31-51.
- Tomitch, L.M.B. (2002). *The teacher's role in reading instruction in EFL*. Trabalho apresentado na mesa redonda "English Language Teaching" , no I Congresso Internacional das Linguagens, na URI- Erechim, em 31 de julho de 2002.

- Urquart, A. H., & Weir, C. J. (1998). *Reading in a second language: process, product and practice*. New York: Longman.
- Van den Broek, P., Risen, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.). *Sources of coherence in reading*. (pp. 353-373). New Jersey, USA: LEA.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Weir, C. (1993). *Understanding and Developing Language Tests*. UK: Prentice Hall International Ltd.
- Williams, J. P. (1988). Identifying main ideas: A basic aspect of reading comprehension. *Topics in Language Disorders*. 8(3), 1-13.
- Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.
- Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, 21, 289-327.

APPENDIX A
THE CAE READING PAPER

UNIVERSITY OF CAMBRIDGE ESOL EXAMINATIONS

English for Speakers of Other Languages

CERTIFICATE IN ADVANCED ENGLISH

0150/1

PAPER 1 Reading

Wednesday **10 DECEMBER 2003** Morning 1 hour 15 minutes

Additional materials:

Multiple Choice answer sheet

Soft clean eraser

Soft pencil (type B or HB is recommended)

TIME 1 hour 15 minutes

INSTRUCTIONS TO CANDIDATES

Do not open this booklet until you are told to do so.

Write your name, Centre number and candidate number on the answer sheet in the spaces provided unless this has already been done for you.

This paper requires you to read four texts and answer the accompanying questions.

For each question indicate in pencil **on the separate answer sheet** one answer from the appropriate list of choices.

Read and follow carefully the instructions given on the separate answer sheet.

You may write on the question paper if you wish, but you must **transfer your answers to the separate answer sheet within the time limit.**

At the end of the examination, you should hand in both the question paper and the answer sheet.

INFORMATION FOR CANDIDATES

Questions in Parts 1 and 4 carry one mark.

Questions in Parts 2 and 3 carry two marks.

This question paper consists of 10 printed pages and 4 blank pages.

2

Part 1

Answer questions 1–15 by referring to the reviews of video releases from a magazine on page 3. Indicate your answers **on the separate answer sheet**.

For questions 1–15, answer by choosing from the reviews (A–F). Some of the choices may be required more than once.

Which film

- | | |
|--|----------|
| allows viewers to appreciate the director's technical skills? | 1 |
| is so entertaining that its lack of originality is unimportant? | 2 |
| has a central character whose personality reflects the setting? | 3 |
| is a greater achievement in terms of planning than of artistic merit? | 4 |
| interprets a story in a comparatively straightforward manner? | 5 |
| is criticised for its extravagant production? | 6 |
| features an actor who is sometimes good, although not really suited to his role? | 7 |
| is an older treatment of a recent cinema release? | 8 |
| is criticised for paying too much attention to appearances but too little to the plot? | 9 |
| is criticised for having a poor script? | 10 |
| features characters who care little about the harm they may be doing others? | 11 |
| is described as being rather better than many films of its type? | 12 |
| moves too slowly at the beginning? | 13 |
| has characters reminiscent of those in another director's films? | 14 |
| is described as being more faithful to its source than another film? | 15 |

FILMS ON VIDEO

Film critic Nick James reviews some recent releases

A

Abyss

Long before *Titanic*, director James Cameron made this sweaty, claustrophobic Cold War thriller about oil riggers and navy experts trying to rescue a nuclear submarine stranded many miles beneath water. The banter and self-deprecating bravery of foreman Bud and his men rekindle memories of similarly laconic heroes in movies directed by Howard Hawks. Production design and special effects are hugely impressive. It's only the dialogue and characterisation that creak. For all the craftsmanship which goes into the film-making, the story itself is strictly B-movie material.

B

The Thin Red Line

The video release of this version of the James Jones novel about the battle for Guadalcanal directed by Andrew Marton makes a fascinating counterpart to Terrence Malick's new film. Whereas Malick's approach is mystical and poetic, Marton made a much more conventional war movie, albeit one that is often truer to the book. He concentrated on a single soldier, and on his relationship with his abrasive sergeant. Malick's film is infinitely richer and more complex, but Marton's version has its moments. The flashback sequence, in which the soldier dreams of the wife he longs for, is handled with a harshness which arguably works better than Malick's soft-focus imagery of the woman on the swing.

C

On Guard

Loosely based on Paul Féval's 1875 novel, this corny but highly watchable swashbuckler is a cut above most musketeer adventures. It has a consummate villain in Fabrice Luchini's clammy politician, orchestrating death and destruction behind the scenes. Vincent Perez makes an exuberant (if rather short-lived) hero, and while Daniel Auteuil is perhaps too moody a presence for a romp like this, he too has his moments as an acrobat-turned-swordsman. The film-makers peddle costume-drama clichés with so much wit and sparkle it never seems to matter.

D

Character

A handsome but dour tale, set in turn-of-the-century Rotterdam. The excessively detailed production and costume design leave the film looking like a museum piece. Taking his cue from the surroundings, Jan Decler is endlessly morose as the brutal bailiff Dreverhaven, who behaves ruthlessly when evicting tenants. His antagonistic relationship with his son is at the core of the story, but the film-makers seem too busy laying on the period detail to do justice to the dark and vicious parable.

E

The Good, the Bad and the Ugly

This digitally re-mastered video re-release shows off director Leone's craftsmanship to its best advantage. The sound editing, in particular, stands out: every footstep, creaking floorboard or barking dog registers loud and clear. The storytelling is relentlessly cruel and whenever there's a lull, it only takes a burst of Morricone's magnificent music to quicken the pulse. On a moral level, there isn't much to distinguish between the good (Clint Eastwood), the bad (Lee van Cleef) and the ugly (Eli Wallach), all of whom seem equally unscrupulous as they maraud across the post-Civil War West.

F

The Longest Day

'Forty-eight international stars' trumpets the publicity for this three-hour Darryl Zanuck war epic. With four directors and 23,000 extras as well, this is one pudding which is definitely over-egged. The early sequences, in which the battle-hardened veterans wait for confirmation of when the invasion will happen, drag as much for the audience as for the soldiers. On a logistical level (if not an aesthetic one), this is an impressive enough feat but it cries out for the big screen. Panned and scanned on video, it is inevitably a diminished experience.

4

Part 2

For questions 16–21, you must choose which of the paragraphs A–G on page 5 fit into the numbered gaps in the following magazine article. There is one extra paragraph which does not fit in any of the gaps. Indicate your answers on the separate answer sheet.

Where the landscape will do the walking

Despite the growth of tourism in the area, Roger Bray finds there are still undeveloped parts of Cape Cod, an exposed peninsula off the east coast of the USA

On the fragile outer shore of Cape Cod the pervading sense is of a universe in which nothing stands still. The ocean wages its war of attrition against the shifting sand, which rises from the beach into a steep cliff. Gulls wheel on the wind, swallows dart low over the water's edge.

16

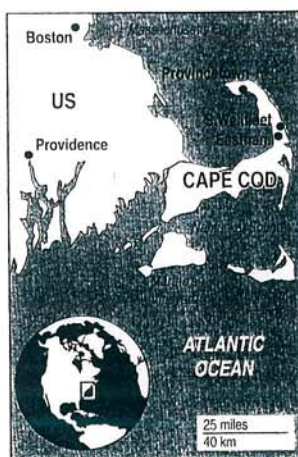
The simple reason is that, here, more than in most places, to get off the roads and away from the most easily accessible beaches is to experience the Cape not just as a holiday retreat for urban Americans but as it has always been.

17

This is mainly because a large swathe of it was established in 1961 as a national park. Our search for recommended hikes took us to the internet – but the maps were hard to follow. We tried bookshops but to no avail. There were books listing walks, to be sure, but the routes they covered were much too short.

18

Following its directions made for superb hiking. To cover the whole of the route we wanted to do would have involved linear sections totalling about 50 kilometres. There were circular itineraries, however, varying in length between about 12 and 20 kilometres, though slow going on soft sand makes them seem longer.



19

One route took us along the Old King's Highway, once a stagecoach route, into the middle of an eerie swamp of Atlantic white cedar, where the sunlight streamed between shaggy barked trunks and where the park management has built a boardwalk and provided nature information.

20

The circuit concluded with an intoxicating hike along the beach. To our right rose the huge sandy cliff, threatening to slide and bury the unwary. Henry Beston, in *The Outermost House*, his lyrical account of a year spent here in the 1920s, describes how, after the cliff was pushed back 6 metres or so by a momentous storm, the long buried wreckage of ships emerged from it, as fruit from a sliced pudding.

21

The shingled Whalewalk Inn was also a delight. It lies behind a white painted picket fence on a leafy road on the fringe of Eastham. It was built in 1830 by Henry Harding, a whaling captain when that industry was at its peak. Later it was used as a farmhouse and a salt works. Nowadays, people also find it a relaxing place to stay.

5

- A It continued to the South Wellfleet sea cliff where Marconi broadcast the first transatlantic wireless message in 1903, sending greetings in Morse code from President Theodore Roosevelt to King Edward VII. The transmitting station was scrapped in 1920 but a model recalls how it looked, its antennae suspended between tall timber masts.
- B If we had sauntered a few kilometres from the car park to stand for a while on that great beach, we might still have felt the whirling of the universe. But without a day of serious hiking to sharpen our appetites, would we have appreciated the food so much?
- C On the other side, however, there was nothing but ocean, jade green inshore, ink blue farther out, between us and the coast of north-west Spain. Although this was a week of near flawless weather in May, we were lucky to encounter only a handful of other walkers. In high summer, when the roads are clogged and there are queues for restaurant tables, it is harder to find an empty stretch of beach.
- D Because, for all the impact of tourism, which nearly triples the population in summer, there are still lonely parts of this storm-scoured, glacial peninsula which have changed little during the last 150 years.
- E We tried several of them. Sometimes we were on woodland trails shaded partly by pitch pine and black oak, sometimes on high windy cliffs overlooking the sea, and sometimes on the foreshore, where we were made diminutive by the huge sky and curving beach of white gold sand.
- F Henry David Thoreau wrote that 'even the sedentary man here enjoys a breadth of view which is almost equivalent to motion'. Perhaps that was why it proved so difficult to find a guide for long hikes. People must wonder why they need to expend effort when they can let the landscape do the walking.
- G Staff at the inquiry desk of the Cape Cod National Seashore's Salt Pond visitor centre were no help, either. But in the centre's bookshop, we struck gold at last. Adam Gamble's *In the Footsteps of Thoreau*, published locally two years ago, has a section tracing the writer's progress in 1849 from Eastham to Race Point Beach, where he turned towards Provincetown, the Cape's outermost community, now a gathering place for whale watchers.

Read the following magazine article and then answer questions 22–28 on page 7. On your answer sheet, indicate the letter **A**, **B**, **C** or **D** against the number of each question 22–28. Give only one answer to each question. Indicate your answers on the separate answer sheet.

Under Sarah's Spell

Sarah Janson is a *trompe-l'oeil* artist whose paintings are designed to deceive the eye by creating the illusion of reality. Here she is interviewed by Joanna Watt

There cannot be many artists who do not sign their work unless they are asked to. Sarah Janson, a *trompe-l'oeil* artist, is one. She is not remotely interested in the concept of the artist as creator, let alone that of the artist as genius: 'It's not the artist who is important, but the work,' she states. Janson is so self-deprecating that she would almost like you to believe that her *trompe-l'oeil* works paint themselves.

All of which does not bode well for a magazine interview. 'I just don't like to shout about myself,' she says, and then covers her face in horror when asked if she minds being photographed for the feature. Cut to her sitting room 30 minutes later (a wonderful space in a block of artists' studios in London, filled with paintings and drawings) and you find two women bent double with hysteria. Her confidence gained, the interview becomes a fascinating, amusing (and sometimes hilarious) encounter.

Janson has been a *trompe-l'oeil* artist for sixteen years, after two years' solid drawing at art school ('the best training any artist can ever have'), a degree in graphic illustration and a stint at a publishing house. But illustration never really satisfied her, and she joined a specialist decorator, Jim Smart ('one of the best in his day'). Smart asked her to do one *trompe-l'oeil*, and that was it. 'Suddenly my interest got channelled,' she says. She left to set up on her own, 'not really knowing where I was going, but feeling that I was on the road to somewhere.' Her instinct was right.

Janson's observational skills and fascination with detail (gained through illustrating) proved essential qualities for a *trompe-l'oeil* artist. 'People often ask me where they can learn *trompe-l'oeil*. But no one can teach you. *Trompe-l'oeil* is the school of life. It's all about observation.' She insists (in that self-deprecating way) that she is still learning. 'The moment you think that you've mastered a field you might as well give up.' She is also brutally honest about her 'failings' ('I can't paint bread; it always looks like grey concrete') and is frank about her mathematical abilities. Faced with a huge commission for the domed chapel ceiling at Lulworth Castle, she became totally confused when calculating measurements. 'I thought to myself, "You're not Michelangelo. Who do you think you are?"' This habit of self-questioning and a reluctance to



openly acknowledge her skill has spawned an oddly distanced attitude to her talent. Janson often speaks in the third person: 'When I finished that ceiling, I thought, "Well I didn't do it, she did".'

Of course, her *trompe-l'oeil* schemes can speak for themselves. Janson's work is in a league of its own, far above those who have jumped on the bandwagon (the art of *trompe-l'oeil* has experienced something of a revival, but not with entirely satisfactory results) and she has a string of major corporate and private commissions behind her.

Much of her work is inspired by architecture or made for architectural settings. There is the *trompe-l'oeil* dining room for one client, based on the façade of a Venetian palazzo, and the painting at the end of a corridor in a flat, which gives the illusion that you can step into two further rooms.

There is always a danger with *trompe-l'oeil*, though, that once you get the joke, your attention is lost, something of which Janson is acutely aware. '*Trompe-l'oeil* has to do two things. First, it must draw you in; it's got to trick you. Secondly, it has to hold you and then engage your imagination. That is the most important part.'

While *trompe-l'oeil* has to be clever, it must also, Janson believes, be personal to the client. 'I love the interaction with clients; that is where the ideas are born,' she says. 'Without the rapport, the job of creating a *trompe-l'oeil* scheme becomes rather difficult. Some clients have firm ideas about what they want; others do not. You have to be willing to listen. You have to get inside a client's imagination.' Many have become friends, not least because Janson practically lives with them if she works on site.

Janson is generous in praise of her clients. 'I am very grateful for the mad ones who have let me loose on their walls,' she confesses. And, they too, seem delighted with her, which is why she is constantly busy – despite her inclination to play down her talent. 'I really don't like to shout about myself,' she repeats at the end. 'Like my work, I am very restrained. I don't want it to shout. You become bored with things that shout.' True, perhaps, but you could never really become bored with Janson or her work. It certainly deserves to become better known, and I am prepared to incur her wrath while I blow her trumpet.

7

- 22 After spending time with Janson, the interviewer concludes that
- A Janson has little faith in journalists.
 - B Janson dislikes interviews in her home.
 - C her initial doubts about Janson were wrong.
 - D her first questions to Janson were threatening.
- 23 What motivated Janson to start her own business as a trompe-l'oeil artist?
- A It was something she was well qualified to do.
 - B She was unhappy with her previous employer.
 - C She was convinced it was what she wanted to do.
 - D It was something that would help her achieve her ambitions.
- 24 What advice does Janson give to people interested in becoming trompe-l'oeil artists?
- A It would be a mistake to become over-confident.
 - B Practice is the only way to improve shortcomings.
 - C Experience in different art forms helps develop essential skills.
 - D A lot can be gained from looking at the work of other artists.
- 25 What point does the interviewer make about Janson's work in paragraph five?
- A It is of an exceptional quality.
 - B Some people regard it as strange.
 - C It is better suited to small locations.
 - D Janson regrets some of the commissions she has taken on.
- 26 What does Janson say about trompe-l'oeil as an art form?
- A It has limited commercial appeal.
 - B The most successful pieces avoid humour.
 - C A small number of people accept it as genuine art.
 - D The difficulty lies in sustaining people's interest.
- 27 What does Janson say about her clients?
- A She prefers to work with clients who have a lot of imagination.
 - B Some clients have ideas which are less practical than others.
 - C She is reluctant to take on commissions if she cannot agree with the client.
 - D A commission is easier if you can discover what kind of ideas the client has.
- 28 What does the interviewer say in the last paragraph?
- A She is puzzled by the way Janson describes her clients.
 - B She realises that Janson may not like what she has written about her.
 - C She is sure that certain types of art soon lose their appeal.
 - D She feels that trompe-l'oeil is unlikely to become a more popular art form.

Part 4

Answer questions 29–45 by referring to the magazine article about writers on pages 9–10. Indicate your answers on the separate answer sheet.

For questions 29–45, choose your answers from the writers (A–E). You may choose any of the writers more than once.

Which writer

- | | |
|--|----------|
| says that he is not the kind of writer who wants a solitary existence? | 29 |
| avoids showing his work to anyone before it has been thoroughly revised? | 30 |
| thinks that some people may have the wrong impression of a writer's life? | 31 |
| no longer feels uneasy about the kind of life writing involves? | 32 |
| points out how much revision can be involved in writing a novel? | 33 |
| says that on some days he knows in advance that writing will be difficult? | 34 |
| says that he has a limited amount of inspiration? | 35 |
| says that it is essential, for a writer's sanity, to spend some time in the company of others? | 36 |
| admits that he does not actually work for the whole time he spends at his desk? | 37 |
| says that he finds it difficult to assess his own writing in a critical way? | 38 |
| forces himself to get something written when he is having difficulties? | 39 |
| thinks that he writes better when working at a fast pace? | 40 |
| draws a contrast between days when it is easy to write and those when it is not? | 41 |
| requires little persuasion to reward himself for work he has done? | 42 |
| says that he feels comfortable with the kind of writing day that he has established for himself? | 43 |
| does not look forward to reading published opinions of his work? | 44 |
| always tries to delay the time when he has to start writing? | 45 |

A Writer's Day

We interviewed five contemporary male novelists to find out how they approach their writing and how they typically spend their day

A

I'm no good at mixed days – it's either work or play. If it's a work day, then I'll start with a huge mug of strong black coffee and then I'll go to my study at the top of the house. It's completely lined with books and has a 'Do Not Disturb' sign on the door.

I've learned to start writing early and to have a scene hanging over from the day before. I'm obsessive about silence. I can't talk in the middle of work – if I talk, the morning is over. Momentum is important to my novels – if I'm too leisurely, I can feel the tension fading away. Dialogue is the best – blissfully easy, like hearing voices in your head and taking dictation. A few years ago, I was writing 5,000 words a day – now, though, it's only half that.

When I go out, I do all the things you're supposed to as a writer, like going out to London clubs. But when people see you at book launches they forget that being a writer is also about that little thing in between – sitting on your own all day. But you've got to have contact with the outside world and real people or you can go completely mad.

B

I'm completely envious of people who write in the mornings and do what they like in the afternoon. I work through the day and treat writing like an office job. If it's not going well, I keep pushing at it and get it sorted out. I don't get a lot of ideas. I tend to get just one and then run with it. Towards the end of a book, when I think I've got to get an idea for the next one, I start to feel panicky. But something always comes along.

My office is in a flat about ten minutes from our house. It's good to have a geographical break between home and work. I arrive about 9 am, have a coffee and then I'll just get on with it and work through until lunchtime. There's a definite post-lunch dip – that's when I have another coffee. But in the end, the only way I get concentration back is by pushing it.

My wife picks me up about 6.30 and we go home together. I've been doing this for ten years now. It's a routine that suits me and, to be honest, I'm always a little worried about breaking it.

C

My seven-month-old daughter, Matilda, gets me up around 6.30 and I'll play with her for a couple of hours, then go to my desk. I officially sit there for three hours, but I'll do an hour's work.

Like a lot of writers, I tend to get a great sense of achievement very easily. One good sentence entitles me to half an hour off – two or three lines means I can watch daytime TV. My study is at one end of the flat and my wife and daughter are at the other. In theory, no congress takes place until lunchtime, but actually we pop in and out all the time. I've never been one of those writers who likes being isolated – I want people around me all the time. At the moment, I'm plotting my next novel and am in a dreamy state. It's hard to convince people that I'm actually working, but this is a crucial part of the process.

A book takes me about three years in all. I always start out very slowly and then gather speed towards the end. I don't think this is a good way to write at all.

D

I have a really slow start to the day. I'll do anything to put off starting work. I have toast, read newspapers – I have to do the crossword every morning – and deal with my post. I write quite slowly and not in chronological order. I've structured the story before I start, so I can hop around, which I think keeps my writing fresh. Sometimes I wake up and just know it's not going to work – because I'm just not in the right mood – but I know that it's only temporary. Once you've got the first draft down, you know that it's going to be OK. It's awful having to relinquish it to my editor, but I'm curious to know what other people think. I find it impossible to be objective about my own words.

When I started writing and just stayed at home I felt incredibly guilty but now it feels normal. Lots of my friends are creative and don't go to offices, which helps. When we go out we don't talk about work – we gossip about the people we know instead. But if I want to use anything my friends have told me, I always ask.

E

I start writing at about 10.30 am. I don't open any mail before that so I haven't got anything on my mind and the desk is clear. I write until 1.30 pm. Sometimes getting the words out is like pulling teeth - other days it all spills out freely.

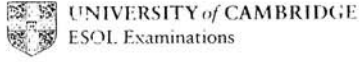
Someone once said that there was no such thing as writing, only re-writing. For my first book, I did no less than 12 drafts. With the first draft of a book, I just try to capture the energy and try to get something down which I can work with. I would die of embarrassment if anyone saw it.

At lunchtime I like to get out of the flat. It's odd going into the outside world - you feel as if you're in a light trance. But after a ten-minute walk in the drizzle I'm usually all right. I spend 90 percent of my time on my own. My contact with the outside world generally happens in great bursts, when I go abroad to publicise my books.

I've just finished my third novel and it's a nerve-racking time. I really dread being at the mercy of book reviewers. But when it comes down to it, I know what my book is like - I don't need to be told by other people.

APPENDIX B

CAE READING PAPER ANSWER SHEET



Candidate Name
If not already printed, write name
in CAPITALS and complete the
Candidate No. grid (in pencil)
Candidate Signature

Centre No.

Examination Title CAE
Centre

Candidate No.
Examination
Details

0	0	0	0
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9

Supervisor:
X If the candidate is ABSENT or has WITHDRAWN shade here:

Multiple-choice Answer Sheet

Use a pencil.
Mark ONE letter for each question.
For example, if you think C is the right answer
to the question, mark your answer sheet like this:

Rub out any answer you wish to change with an eraser.

1	A	B	C	D	E	F	G	H	I
2	A	B	C	D	E	F	G	H	I
3	A	B	C	D	E	F	G	H	I
4	A	B	C	D	E	F	G	H	I
5	A	B	C	D	E	F	G	H	I
6	A	B	C	D	E	F	G	H	I
7	A	B	C	D	E	F	G	H	I
8	A	B	C	D	E	F	G	H	I
9	A	B	C	D	E	F	G	H	I
10	A	B	C	D	E	F	G	H	I
11	A	B	C	D	E	F	G	H	I
12	A	B	C	D	E	F	G	H	I
13	A	B	C	D	E	F	G	H	I
14	A	B	C	D	E	F	G	H	I
15	A	B	C	D	E	F	G	H	I
16	A	B	C	D	E	F	G	H	I
17	A	B	C	D	E	F	G	H	I
18	A	B	C	D	E	F	G	H	I
19	A	B	C	D	E	F	G	H	I
20	A	B	C	D	E	F	G	H	I

21	A	B	C	D	E	F	G	H	I
22	A	B	C	D	E	F	G	H	I
23	A	B	C	D	E	F	G	H	I
24	A	B	C	D	E	F	G	H	I
25	A	B	C	D	E	F	G	H	I
26	A	B	C	D	E	F	G	H	I
27	A	B	C	D	E	F	G	H	I
28	A	B	C	D	E	F	G	H	I
29	A	B	C	D	E	F	G	H	I
30	A	B	C	D	E	F	G	H	I
31	A	B	C	D	E	F	G	H	I
32	A	B	C	D	E	F	G	H	I
33	A	B	C	D	E	F	G	H	I
34	A	B	C	D	E	F	G	H	I
35	A	B	C	D	E	F	G	H	I
36	A	B	C	D	E	F	G	H	I
37	A	B	C	D	E	F	G	H	I
38	A	B	C	D	E	F	G	H	I
39	A	B	C	D	E	F	G	H	I
40	A	B	C	D	E	F	G	H	I

41	A	B	C	D	E	F	G	H	I
42	A	B	C	D	E	F	G	H	I
43	A	B	C	D	E	F	G	H	I
44	A	B	C	D	E	F	G	H	I
45	A	B	C	D	E	F	G	H	I
46	A	B	C	D	E	F	G	H	I
47	A	B	C	D	E	F	G	H	I
48	A	B	C	D	E	F	G	H	I
49	A	B	C	D	E	F	G	H	I
50	A	B	C	D	E	F	G	H	I
51	A	B	C	D	E	F	G	H	I
52	A	B	C	D	E	F	G	H	I
53	A	B	C	D	E	F	G	H	I
54	A	B	C	D	E	F	G	H	I
55	A	B	C	D	E	F	G	H	I
56	A	B	C	D	E	F	G	H	I
57	A	B	C	D	E	F	G	H	I
58	A	B	C	D	E	F	G	H	I
59	A	B	C	D	E	F	G	H	I
60	A	B	C	D	E	F	G	H	I

APPENDIX C

ANSWER KEY TO THE CAE READING PAPER

Task 1	Task 2	Task 4
1- E	16- F	29- C
2- C	17- D	30- E
3- D	18- G	31- A
4- F	19- E	32- D
5- B	20- A	33- E
6- D	21- C	34- D
7- C		35- B
8- B	Task 3	36- A
9- D	22- C	37- C
10- A	23- C	38- D
11- E	24- A	39- B
12- C	25- A	40- A
13- F	26- D	41- E
14- A	27- D	42- C
15- B	28- B	43- B
		44- E
		45- D

APPENDIX D

PARTICIPANTS' MARKS AND TIMES

IN THE CAE READING PAPER

Table D1: Group 1.1- Participants' Marks and Times in the CAE Reading Paper

Group 1.1	Mark	Time (min)
1	74,14	75,00
2	53,45	75,00
3	27,59	75,00
4	63,79	75,00
5	58,62	59,00
Means	55,52	71,80

Table D2: Group 1.2 – Participants' Marks and Times in the CAE Reading Paper

Group 1.2	Mark	Time (min)
6	62,07	65,00
7	82,76	57,00
8	25,86	68,00
9	89,65	70,00
Means	65,09	65,00

Table D3: *Group 2.1 – Participants' Marks and Times in the CAE Reading Paper*

Group 2.1	Mark	Time (min)
10	48,27	75,00
11	51,72	75,00
12	60,34	75,00
Means	53,44	75,00

Table D4: *Group 2.2 – Participants' Marks and Times in the CAE Reading Paper*

Group 2.2	Mark	Time (min)
13	62,07	83,00
14	60,34	114,00
15	32,76	98,00
16	72,41	82,00
Means	56,90	94,25

Table D5: *Group 3.1 – Participants' Marks and Times in the CAE Reading Paper*

Group 3.1	Mark	Time (min)
17	96,55	69,00
18	67,24	75,00
19	60,34	75,00
20	48,27	72,00
21	60,34	75,00
22	84,48	59,00
23	82,76	63,00
24	67,24	68,00
Means	70,90	69,50

Table D6: Group 3.2 – Participants’ Marks and Times in the CAE Reading Paper

Group 3.2	Mark	Time (min)
25	44,83	96,00
26	86,21	84,00
27	74,14	49,00
28	82,76	68,00
29	60,34	100,00
30	81,03	100,00
Means	71,55	82,83

Table D7: Group 4.1 – Participants’ Marks and Times in the CAE Reading Paper

Group 4.1	Mark	Time (min)
31	84,48	55,00
32	79,31	68,00
33	82,76	59,00
34	81,03	64,00
35	89,65	75,00
36	68,96	68,00
Means	81,03	64,83

Table D8: Group 4.2 – Participants’ Marks and Times in the CAE Reading Paper

Group 4.2	Mark	Time (min)
37	70,69	95,00
38	89,65	54,00
39	91,38	64,00
40	93,10	85,00
Means	86,21	74,50

APPENDIX E
THE READINESS TEST

8

Part 4

You are going to read a magazine article in which five people talk about their characters. For questions 22-35, choose from the people (A-E). The people may be chosen more than once. When more than one answer is required, these may be given in any order. There is an example at the beginning (0).

Mark your answers **on the separate answer sheet**.

Which person or people state(s) the following?

I used to avoid giving my opinions at work.

0	E
---	---

Taking time off for your professional development can make you feel more self-assured.

22	
----	--

I never thought I'd be a confident person.

23	
----	--

I'm not influenced by people's opinions of me.

24	
----	--

Everyone gets nervous at times.

25	
----	--

26	
----	--

Initially, I misunderstood what confidence was.

27	
----	--

I find making notes very supportive in my work.

28	
----	--

A certain event changed the course of my life.

29	
----	--

30	
----	--

I've worked on having a confident appearance.

31	
----	--

I am realistic about my abilities.

32	
----	--

33	
----	--

My behaviour helps others relax too.

34	
----	--

Getting things wrong can have a positive result.

35	
----	--

Confident people

What's their secret?

Confident people may look as though they were born that way, but most will tell you that it's a skill they've learned because they had to. Nina Hathway asks five people how they did it.

A Jenny

When I left school I was very shy and I always thought I'd stay that way. I was about twenty-five when I was asked to help out at my daughter's school. I was sure I wouldn't cope, but I surprised myself by doing well and someone there suggested that I should do a university course.

There was a huge knot in my stomach the day I turned up for my first lecture. But my confidence gradually grew – I became more outgoing. Looking back, working at the school was the turning point in my life that has helped everything else fall into place.

B Michaela

It all started four years ago when my father became ill and I had to take over the family business. I was so scared, I went over the top and became a bit too aggressive and impatient. I thought that was what confident people were like, but gradually I learned otherwise. To be confident you've got to believe in yourself.

If things get too demanding for me at work, I don't let myself feel guilty if I save a number of tasks until the next day. When I'm confronted with something difficult, I tell myself that I've got nothing to lose. It's fear that makes you lack confidence, so I'm always having quiet chats with myself to put aside those fears!

C Carol

People think I'm very confident but, in fact, the calmer I look, the more terrified I really am. I've had to develop the ability to look confident because it's the most vital thing in TV. Interviewing people has helped me realise that most – if not all – of us get tense in important situations, and we feel calmer when we speak to someone who's genuinely friendly. The best ever piece of advice came from my mother when I was agonising as a teenager about wearing the right clothes. She simply cried, 'Who's looking at you? Everybody's too busy

worrying about how they look.' I've found that's well worth remembering.

I also think you gain confidence by tackling things that scare you. When I took my driving test I was so nervous, but I passed. After that I felt sure that I'd never feel so frightened again, and I never have.

D Barbara

My confidence comes naturally from really enjoying the work I do, but it's something that I've built up over the years. If you just get on with it and learn from any mistakes you make, you're more confident the next time round. I work hard and I'm popular in the restaurant, but it's probable that one out of ten people doesn't like me. I don't let that affect me. You've got to like yourself for what you are, not try to be what others expect.

My company runs a lot of training courses, and going on those has built up my self-esteem. The company also encourages employees to set manageable targets. It helps no end if you can see you're achieving something tangible, rather than reaching for the stars all at once, and ending up with nothing but air!

E Elaine

After I left college I worked for years as a secretary and would sit in meetings, not always agreeing with what was being said, but too scared to speak up. Eventually, I summoned up the confidence to start making my point. Even so, when I first worked in politics, I'd never spoken in public before and always used to shake like a leaf. I would say to myself, 'Don't be so silly. People do this every day of their lives, so there's no reason why you can't.' I also found it helpful to jot a few things down to refer to – rather like having a comfort blanket!

I don't think there is anyone who isn't a little shaky when it comes to talking publicly. The real secret of confidence lies in telling yourself over and over again, 'Nothing is impossible.'

Read the following magazine article and answer questions 19–24 on page 7. On your answer sheet, indicate the letter **A, B, C** or **D** against the number of each question, 19–24. Give only one answer to each question. Indicate your answers on the separate answer sheet.

The Cabinet-Maker

Charles Hurst makes a living from perfectly crafted furniture.

Joanna Watt meets him

Charles Hurst gives the impression of being a man in a hurry. I arrive at his workshop, tucked under a railway arch in East London, and am greeted with a quick handshake and the words: 'Well, fire away then!' Whether this brusqueness is real or a front hiding a shy streak is not immediately apparent. But a glance around the workshop reveals that Hurst is obviously busy, with good reason not to waste a minute of his time.

The arched space is full of half-made pieces of furniture and planks of wood in an amazing array of natural colours. Hurst has been a cabinet-maker for ten years and has built up a very nice reputation for himself. His order book is always full for several months in advance, despite the fact that he does not really promote himself. Word has spread that if you want a decent cupboard or table, bookcase or kitchen units, Hurst is your man.

Of course, finding a furniture-maker is not that taxing a task. Wherever you live in the countryside, the craft is alive and well. But finding a cabinet-maker who prides himself on making beautifully crafted furniture with clean, simple lines is less easy. 'There are few real cabinet-makers now. People call themselves furniture-makers,' Hurst says wearily. As a craftsman who sets himself exacting standards, he is continually disappointed by some contemporary furniture. 'I am amazed by what some furniture-makers get away with, and saddened by what people will put up with.' He rails against shoddy, mass-produced furniture, and craftsmen who churn out second-rate pieces.

Such a quest for perfection is obviously a key to Hurst's success. That and his talent. This man is not coy about his ability. Indeed, his blatant self-confidence is as surprising as his initial brusque manner. 'I have a huge natural ability,' he says, with a deadpan expression. 'I have always been good at making things.' If it were not for the self-deprecating mood into which he slipped towards the end of our interview, I would have believed his conceit to be wholly genuine.

Hurst is self-taught. So how did he learn his craft? 'I asked the right questions and picked it all up,' he says nonchalantly. Almost all of his commissions come from private individuals ('I used to do some commercial work for companies but it was soul-destroying'). Some clients have returned time and again. 'You end up doing the whole of their house. That is very satisfying.' But he is honest enough to admit that relationships with clients do not always run smoothly. 'The most infuriating clients are those who don't know what they want, and then decide they do when it's too late ... my favourite clients are the exacting ones.'

If Hurst has every reason to be pleased with himself, he is also gracious in his praise for others – where it is due. With a sudden shot of modesty, he says: 'There are people far better than me. I can admire other people. After all, I wasn't trained at Parnham' (the leading college of furniture design). However, he is also unremittingly critical of those craftsmen who 'are trying to be artists and take a year to make one piece.' He also has little time for degree shows, in which students exhibit their work but at the same time are 'trying to make fashion statements. That can be pretentious. A piece of furniture is not about making a statement. It has to be something that people really can use.'

Confident Hurst may be, even brusque, but you could never call him or his work pretentious. Indeed, his parting shot displays a welcome down-to-earth approach to his craft and a streak of humility strangely at odds with his earlier self-confidence. 'After all, I am only making furniture,' he says as I make my exit.

- 19 When she arrived at the workshop, the writer
- A was not sure if her first impression of Hurst was accurate.
 - B was offended by the way Hurst introduced himself.
 - C thought that Hurst was pretending to have a lot to do.
 - D thought it was obvious that Hurst did not want to speak to her.
- 20 Hurst has few problems selling his furniture because he
- A advertises locally.
 - B is known to be a skilled craftsman.
 - C uses only natural materials.
 - D has a reputation for being fair.
- 21 What does Hurst think has led to the decline in the craft of cabinet-making?
- A It is a difficult skill to learn.
 - B It is only popular in rural areas.
 - C Consumers will accept poor quality furniture.
 - D Simple designs do not appeal to modern tastes.
- 22 The writer says that when Hurst describes his 'talent', he
- A has a tendency to exaggerate.
 - B reveals a natural sense of humour.
 - C becomes more animated than he usually is.
 - D appears more arrogant than he really is.
- 23 Hurst believes that it is essential for craftsmen to
- A create original furniture.
 - B exhibit to a wide audience.
 - C produce functional designs.
 - D invest extra time in perfecting their work.
- 24 The writer's final impression of Hurst is that he
- A has an unusual attitude to his work.
 - B believes in the special nature of his work.
 - C enjoys being interviewed about his work.
 - D has the ability to put his work into perspective.

6

Part 3

You are going to read an article written by someone who lives in a house in a valley. Seven sentences have been removed from the article. Choose from the sentences **A-H** the one which fits each gap (**16-21**). There is one extra sentence which you do not need to use. There is an example at the beginning (**0**).

Put your answers on the separate answer sheet.

LIVING IN THE VALLEY

We had been living in our valley for sixteen months when we first realised the dangers that could exist. **0** **H** Until then, we had felt safe and sheltered in our valley.

Soon snow began to fall. Within a day it lay some 15 centimetres deep. **16** But on the neighbouring heights the snow was much deeper and stayed for longer. Up there the wind blasted fiercely. Deep in our valley we felt only sudden gusts of wind; trees swayed but the branches held firm.

And yet we knew that there was reason for us to worry. The snow and wind were certainly inconvenient but they did not really trouble us greatly. **17** It reminded us of what could have occurred if circumstances had been different, if the flow of water from the hills had not, many years before, been controlled, held back by a series of dams.

In a short time the snow started to melt. Day after day, we watched furious clouds pile up high over the hills to the west. Sinister grey clouds extended over the valleys. **18** We had seen enough of the sky; now we began to watch the river, which every day was becoming fuller and wilder.

The snow was gradually washed away as more and more rain streamed from the clouds, but high up in the hills the reservoir was filling and was fast approaching danger level. And then it happened – for the first time in years the reservoir overflowed. **19**

The river seemed maddened as the waters poured almost horizontally down to its lower stretches. Just a couple of metres from our cottage, the stream seemed wild beneath the bridge. **20** For three days we prayed that it would stay below its wall. Our prayers were answered as the dam held and the waters began to subside.

On many occasions through the centuries before the dam was built, the river had flooded the nearby villages in just such a rage. Now, though, the dam restricts the flow of the river and usually all is well; the great mass of water from the hills, the product of snow and torrential rain, remains behind its barrier with just the occasional overflow. **21** We can feel our home in the valley is still secure and safe.

7

- | | |
|---|--|
| <p>A It was the river, the Ryburn, which normally flowed so gently, that threatened us most.</p> | <p>E We can thus enjoy, rather than fear, the huge clouds that hang over the valley, and can be thrilled by the tremendous power which we know the river possesses.</p> |
| <p>B And yet the immense power of all this water above us prevents us from ever believing ourselves to be completely safe in our home.</p> | <p>F It almost completely blocked our lane and made the streamside path slippery and dangerous.</p> |
| <p>C They twisted and turned, rising eastwards and upwards, warning of what was to come.</p> | <p>G There in the heights it was like the Niagara Falls, as the water surged over the edge of the dam and poured into the stream below.</p> |
| <p>D It was far deeper than we'd ever seen it so near our home, lunging furiously at its banks.</p> | <p>H It was the year when the storms came early, before the calendar even hinted at winter, even before November was out.</p> |

APPENDIX F**ANSWER KEY TO THE READINESS TEST**

Task 1	Task 2	Task 3
22- D	19- A	16- F
23- A	20- B	17- A
24- D	21- C	18- C
25- C	22- D	19- G
26- E	23- C	20- D
27- B	24- D	21- E
28- E		
29- A		
30- B		
31- C		
32- B		
33- D		
34- C		
35- D		

APPENDIX G

THE PRACTICE MATERIAL

You are going to read an extract from a book about collecting animals. Eight sentences have been removed from the article. Choose from the sentences **A-I** the one which fits each gap (**15-21**). There is one extra sentence which you do not need to use. There is an example at the beginning (**0**).

Mark your answers on the separate answer sheet.

A Zoo in My Luggage

Some time ago, I made a six-month trip with my wife to Bafut, a mountain grassland kingdom in West Africa. **0** I We wanted to collect our own zoo.

For many years I had been financing and organising expeditions to many parts of the world to collect wild animals for various zoos.

15 After months of giving the animals every care and attention, you had to part with them.

If you are acting as mother, father, food-provider and protector to an animal, half a year is enough to build up a very real friendship with it. **16** Then, just when this relationship should begin to bear fruit, when you ought to be in a unique position to study the animal's habits and behaviour, you are forced to part company.

There was only one answer to this problem as far as I was concerned, and that was to have a zoo of my own. **17** The zoo, of course, would have to be open to the public so that, from my point of view, it would be a sort of self-supporting laboratory in which I could keep and watch my animals.

18 I, like many other people, have been seriously concerned by the fact that year by year, all over the world, various kinds of animal are being slowly but surely wiped out in the wild, thanks directly or indirectly to human interference.

19 Often this is because they are small and generally of no commercial or touristic value, but, to me, wiping out any animal is a criminal offence, just like the destruction of anything we cannot recreate or replace, such as a great painting or an ancient monument.

In my opinion, zoos all over the world should have as one of their main objectives the conservation of these rare and threatened species. **20** For many years I had wanted to start a zoo with just such an objective in view, and now seemed the ideal moment to begin.

Any reasonable person would have established the zoo first and obtained the animals afterwards. But throughout my life I have rarely, if ever, achieved what I wanted by handling it in a straightforward fashion.

21 This was not so easy as it might seem on the face of it, and looking back now I cannot believe I seriously tried to achieve success in this way. It explains why, for some considerable time, my animal collection travelled around with me.

- A** Financial considerations aside, there was another and, to my mind, more urgent reason for creating a zoo.
- B** I could then bring my animals back knowing what sort of food and treatment they were going to receive, and happy in the knowledge that I could go on studying them to my heart's content.
- C** So, naturally, I went and did the opposite, getting the animals first and then setting about the task of finding a zoo.
- D** The difficulties arose when I began my search for a suitable cage to house the animal while it raised its young in a way as similar as possible to its natural behaviour in the wild.
- E** Then, if it is unavoidable that an animal should be wiped out in the wild, at least we have not lost it completely.
- F** While many worthy and hard-working societies are doing their best to tackle this problem, I know a great number of animals which are not receiving adequate protection.
- G** The creature trusts you and, what is more important, behaves naturally when you are around.
- H** Bitter experience during this time had taught me that the worst and most heart-breaking part of any collecting trip came at the end.
- I** Our reason for going there was, to say the least, somewhat unusual.

You are going to read a magazine article about memory. For Questions 22-35, choose from the sections of the article (A-E). The sections may be chosen more than once. When more than one answer is required, these may be given in any order. There is an example at the beginning (0).

Mark your answers on the separate answer sheet.

In which section(s) of the article are the following mentioned?

remembering previous instances of regular activities	0	E	
things that we do not wish to remember	22		
being unable to prove whether or not people have remembered things correctly	23		
a belief that we forget certain types of information more quickly than others	24		
something that might bring back a memory	25		26
recalling activities that are unusual for you	27		
the fact that writing down an event does not mean that you will remember it	28		
a belief that we never really forget anything	29		
the way that one memory can confuse another	30		
the idea that your ability to remember an event does not always depend on how long ago it happened	31		
a belief that we are always gaining and losing memories at the same time	32		33
the enormous amount of information that the human mind is capable of holding	34		
the idea that memories which survive a certain length of time may never be forgotten	35		

What makes us forgetful?

- A** It would require a considerable effort of memory to recall all the theories that explain why we forget. At one extreme is the 'bathtub theory', with new memories pouring in through the tap and the old ones out through the plughole at a more or less constant rate. At the other extreme lies the theory of an infinite lake, where all our memories are stored until we die. All it takes is the right bait on the hook (a smell, a sound, a picture) to fish them out again.
- B** That both theories are confirmed by evidence is enough to suggest that neither is wholly true. People who keep diaries frequently look back and find that, despite the clues, they have absolutely no recall of the events recorded in them. It is obvious that we do completely lose touch with some events from the past. And yet we are all familiar with the experience of the suddenly recovered memory. You come across an old photograph and there he is again, your old friend John, last seen at the end-of-term party. You had totally forgotten him. Except, of course, you hadn't.
- C** It's obvious that our brains store more than we can recall at any given time. So what's going on? In physical terms, scientists are tending towards a view that sounds remarkably like the bathtub theory. The process which fixes memory in the brain must be balanced by an opposite process to erase them. Otherwise the brain over time would become flooded with unwanted memories and, despite its vast capacity, would be unable to soak up any more.
- D** Testing memory is not easy. Researchers can't confirm the accuracy of individual recollections or objectively measure how much has been lost. Checking the memories of former university students for the names and faces of their classmates and the geography of their university towns does, however, suggest that different things may be forgotten at different rates. The memory for names and faces, for example, seems to decline fairly steadily over the first 30 years, then to fall off rapidly between 35 and 50. With academic learning, on the other hand, rapid early forgetting seems to be followed by a remarkably strong long-term retention. With foreign languages, for example, there is a very steep decline over the first three or four years, but any material that lasts longer may stay with you for the rest of your life.
- E** The passage of time is clearly a factor in the process of forgetting, but not the whole story. If it were, all memories would decay at the same rate, and each year of advancing age would wipe away another year's memories. Other factors must be involved – for example, whether something that happened made a strong impression on you at the time, the number of times you have thought about it, and 'interference'. Interference most often occurs when the memory of a recent experience muddles the memory of an earlier, similar one. You can easily check the truth of it yourself. Fix upon something you only do infrequently – going to the theatre, perhaps. No matter how long ago it was, your memory of it is likely to be strong. Now think about something you do every week – for example, visiting the supermarket. Your memories of the last visit will be clear, but beyond that, they will merge into a blur. Usually, a recent memory will interfere with an older one, but occasionally it works the other way round – a distant memory gets mixed up with a new one.

APPENDIX H**ANSWER KEY TO THE PRACTICE MATERIAL**

Task 1	Task 2
15- H	22- C
16- G	23- D
17- B	24- D
18- A	25- A
19- F	26- B
20- E	27- E
21- C	28- B
	29- A
	30- E
	31- E
	32- B
	33- C
	34- C
	35- D