IELTS

6 The cognitive processes of taking IELTS Academic Writing Task 1

Authors
Guoxing Yu, Pauline Rea-Dickins and Richard Kiely
University of Bristol

Grant awarded Round 13, 2007

An investigation of the cognitive processes of candidates taking IELTS Academic Writing Task 1 (AWT1) with different graphic prompts at two time points – before short training on how to do AWT1 tasks, and after. It explores the extent to which candidates' cognitive processes are affected by the use of different graphs, their graphic skills and English writing abilities, and by the training.

Click here to read the Introduction to this volume which includes an appraisal of this research, its context and impact.

ABSTRACT

This research investigated the cognitive processes of candidates taking IELTS Academic Writing Task One (AWT1) with different graphic prompts at two different time points – before short training on how to do AWT1 tasks, and post-training. It explored the extent to which candidates' cognitive processes are affected by the use of different graphs, their graphic skills and English writing abilities, and the short training. A grounded and multi-layered case study approach was employed to collect data on candidates' cognitive processes. 24 intending IELTS candidates from a large Chinese university completed eight AWT1 tasks while thinking aloud their processes of doing the tasks (four before training and four after training) under examination conditions. Samples of their English writing abilities and graphicacy were also collected, as well as post-task interviews with all participants.

The think-aloud protocols were analysed to identify the common patterns of cognitive processes. A model of cognitive processes was developed, consisting of three interrelated stages – comprehending non-graphically presented task instructions, comprehending graphic information and re-producing graph comprehension in written discourse in English as a foreign language. This model guided our analyses to address the four research questions: (1) How the participants processed the graphic information and how they followed the graphic conventions to re-produce their graph comprehension in written discourse in English were affected by the types of graphs they read. Such effects of different graphic prompts on the cognitive processes were clearly evidenced in the mean scores of the writings, in the use of vocabulary, and in whether and how they would make comparisons or trend assessments, following the graphic conventions in presentation, interpretation and re-production. (2) Although graph familiarity did not seem to affect task performance in terms of the marks of the writings, the participants clearly expressed some potential psychological impact of graph familiarity on their task performance. (3) There is a strong correlation between AWT1 writing performance and writing ability as measured via topic-based argumentative essays. (4) The influence of the special training was strong, in particular, in terms of whether or not the participants tried to make interpretations, predictions and comments by linking the graphic information with their domain knowledge about the graphs.

The implications of these findings are discussed with reference to AWT1 task design, as well as other language test tasks that use graphs as prompts, particularly for listening, speaking and writing assessments.

AUTHOR BIODATA

GUOXING YU

Dr Yu is Lecturer at the Graduate School of Education, University of Bristol. He earned his PhD in 2005 from Bristol; his dissertation was awarded the Jacqueline A. Ross TOEFL Dissertation Award by Educational Testing Service USA (2008). He was the Principal Investigator of this research project. He publishes in academic journals including *Applied Linguistics*, *Assessing Writing*, *Assessment in Education*, *Educational Research*, *Language Assessment Quarterly*, and *Language Testing*.

PAULINE REA-DICKINS

Dr Rea-Dickins is Professor of Applied Linguistics at the Graduate School of Education, University of Bristol. She has published widely in areas of language testing and assessment, particularly classroom-based assessment and language programme evaluation. She is Director of a major ESRC/DfID research project (2007-2010, RES-167-25-0263): *Student Performance in National Examination: the dynamics of language* (SPINE; www/bristol.ac.uk/spine), which is a collaboration between Bristol and the State University of Zanzibar.

RICHARD KIELY

Dr Kiely is Reader of Applied Linguistics at the Graduate School of Education, University of Bristol. His research interests include language programme evaluation, language teaching and teacher development, and language learning explored from language socialisation and identity perspectives.

CONTENTS

1 BACKGROUND AND RATIONALE 4
1.1 Introduction 4
1.2 Dearth of research into test takers' cognitive processes of completing AWT1 4
1.3 Theories of graph comprehension in cognitive psychology and their implications for research into integrated writing tasks using graph prompts in language tests 5

2 RESEARCH AIMS AND QUESTIONS 8
2.1 Research aims 8
2.2 Research questions 8

3 RESEARCH DESIGN AND METHODOLOGY 8
3.1 Approach 8
3.2 Participants 8
3.3 Data collection 10

4 ANALYSES AND FINDINGS 13
4.1 Participants' graph familiarity and English writing abilities 13
4.1.1 Participants' graph familiarity 13
4.1.2 Participants' writing performances under normal examination conditions 14
4.1.3 Participants' writing performances under think-aloud conditions 15
4.2 Developing a working model of cognitive processes for AWT1 tasks 16
4.2.1 Comprehending non-graphically presented task instructions 17
4.2.2 Comprehending graphic information 18
4.2.3 Re-presenting graphic and non-graphic information in continuous discourse 18
4.2.4 Summary of the working model 19
4.3 Addressing the four research questions 21
4.3.1 Research question 1: To what extent are there differences in the candidates' cognitive processes due to different AWT1 prompts? 21
4.3.2 Research question 2: To what extent are the candidates' cognitive processes affected by their graphicacy? 25
4.3.3 Research question 3: To what extent are the candidates' cognitive processes related to their writing abilities? 28
4.3.4 Research question 4: To what extent are the candidates' cognitive processes influenced by test preparation opportunities offered to them by the research team? 31

5 CONCLUSION 34

ACKNOWLEDGEMENTS 36

REFERENCES 37

APPENDIX 1: CONSENT FORM 40

APPENDIX 2: IELTS AWT1 TASKS WITHOUT THINK-ALOUD (STAGE 1) 41

APPENDIX 3: GRAPHICACY QUESTIONNAIRE 42

APPENDIX 4: THINK-ALOUD TRAINING DOCUMENT 46

APPENDIX 5: THE EIGHT AWT1 TASKS 49

APPENDIX 6: STUDENT EVALUATION OF THE COG-PRO AWT1 TRAINING 57

APPENDIX 7: INTERVIEW QUESTIONS 59

APPENDIX 8: FREQUENCY OF THE RESPONSES TO THE GRAPH FAMILIARITY QUESTIONS (12-46) 60

APPENDIX 9: THE MOST FREQUENTLY USED WORDS IN EACH TASK 62

APPENDIX 10: PERSONAL INTERPRETATIONS AND COMMENTS IN TASK B AND TASK A WRITINGS 67

APPENDIX 11: STUDENT EVALUATION OF THE PROJECT AND ITS TRAINING: GRADE AND QUALITATIVE COMMENTS 73

1 BACKGROUND AND RATIONALE

1.1 Introduction

This study relates to the first broad area of interest identified by the IELTS Joint Research Committee Call for Proposals 2007/2008 (Round 13), namely "test development and validation issues". In particular, this research investigated "the cognitive processes of IELTS test takers" when completing Academic Writing Task One (AWT1). The cognitive processes of taking AWT1 were examined at two different points – one before and one after special short training on how to achieve the best performance in AWT1. As such, this research is also linked to "issues in test impact" in relation to "test preparation practice" – another area of interest identified by the IELTS Joint Research Committee.

This section provides the theoretical rationale and background for this study with specific reference to the effects of the features of graphs and test takers' "graphicacy" (Wainer 1992, p 16) on their AWT1 performances, as explained below. Details of the research design and methodology are provided from Section 2 onwards.

1.2 Dearth of research into test takers' cognitive processes of completing AWT1

In IELTS AWT1 tasks candidates are asked to "describe some information (graph/chart /table/diagram), and to present the description in their own words". It is suggested that candidates should spend 20 minutes on this and write at least 150 words. Candidates are assessed on their ability to organise, present and possibly compare data, describe the stages of a process or procedure, describe an object or event or sequence of events, or explain how something works (*IELTS Handbook* 2006, p 8). AWT1, therefore, can be considered as an integrated writing task, requiring candidates not only to comprehend the graph input, but also to re-present in written English the information accessible to them (various terms such as chart, graph and diagram have been used interchangeably in research; see Friel, Curcio, & Bright 2001, Fry 1981, Wainer 1992). The term "graph" is probably the most widely used in applied cognitive psychology, the key knowledgebase upon which this research will draw, and we will use "graph", hereafter, to represent all the other three terms – table/chart/diagram – that the *IELTS Handbook* (2006) has used.

Graph comprehension is a *sine qua non* for successful performance of the writing task. As a result, the variability in the graph input and the candidates' different familiarities and proficiencies in comprehending the graphs may pose a threat to the validity of AWT1 as a measure of the candidates' academic writing abilities. Surprisingly, to the best of our knowledge, we notice that only two IELTS funded research projects (Mickan, Slater, & Gibson 2000; O'Loughlin & Wigglesworth 2003) have so far investigated some of these issues. As only a very small part of their research focus, Mickan et al. (2000) investigated how test takers interpreted AWT1 task prompts and planned their writings, but they did not examine the effects of characteristics of graphs on the process or the product of the AWT1 tasks. O'Loughlin and Wigglesworth (2003) examined the extent to which the difficulty of AWT1 was affected by the quantity and the manner of presentation of information in graphs. As a primarily product-oriented study via the analyses of the written scripts, they found that the writings produced in the tasks with less information in the graphs were more linguistically complex than those writings produced in the tasks with more information in the graphs, irrespective of the participants' language proficiency level. The results also indicated that there were no substantial differences in the difficulty across the tasks which varied in terms of the quantity and the manner of information presented in the graphs. However, we should point out that only three types of graphs (ie, statistical table, bar chart, line graph) were used in O'Loughlin and Wigglesworth (2003). These graphs were also limited to two topics (ie, number of people attending places of entertainment, women and men in postgraduate studies). All these put in question the generalisability of their finding that there are no substantial differences in the difficulty across the AWT1 tasks. Would similar findings be observed if

other types of graphs (eg, diagrams depicting the sequence of events and statistical tables in conjunction with other visual presentations) of various topics had been used in their study?

Furthermore, "we have no way of knowing what exactly was attended to by the participants while planning for and completing the tasks" (Xi 2005, pp 496-497), and how exactly the different quantity and manner of presentation of the information in the graphs affected the participants' test taking processes. Compared to the enormous quantity of funded research into IELTS Academic Writing Task Two (eg, Mickan & Slater 2003; Mickan, et al 2000; Moore 1993), the dearth of research into AWT1 is striking. Indeed, in the field of language testing, research into the use of graphs in writing or speaking tests is only recently emerging (ie, Katz, Xi, Kim, & Cheng 2004; Xi 2005), although graphs have been as test prompts for years and a rich knowledgebase in applied cognitive and educational psychology has much to offer language testers to understand the relevant issues. As Xi (2005) commented (see above), these few studies in language testing are very much product-oriented and are not very helpful in understanding test takers' cognitive processes when completing tasks using graphs as prompts. Therefore, the need to gain better understandings of the cognitive processes in taking AWT1 is also compelling, in order not only to understand the validity of the AWT1 *per se* but also to make important contribution to the current debate on the effects on language test performance of the features of graphs and test takers' graphicacy, ie, "proficiency in understanding quantitative phenomena that are presented in a graphical way" (Wainer 1992, p 16).

1.3 Theories of graph comprehension in cognitive psychology and their implications for research into integrated writing tasks using graph prompts in language tests

AWT1, as an integrated writing task, involves two basic processes – the comprehending of the information presented in graphs and the re-presentation of the information from graphs in continuous written discourse. We, thus, identify two main areas of research that inform this study – graph theories in cognitive psychology and second/foreign language writing processes. However, in this report we draw more on the findings in graph theories in cognitive psychology than on second language writing processes, because we believe graph theories have been somewhat neglected in studies about the processes and the assessment of second/foreign language writing. In this section, we briefly review the key factors in graph comprehension in psychology and statistics/mathematics education pursuant to a cognitive approach (as opposed to the view of graph comprehension as a social practice, in which graphs are seen not to have meaning *a priori*, but rather the meaning of graphs arises from the contexts of use [Roth 2002, 2003]). We will discuss the implications of these theories for language test development and validation, with specific reference to IELTS AWT1.

In cognitive psychology, several models or frameworks of graph comprehension (eg, Carpenter & Shah 1998; Freedman & Shah 2002; Guthrie, Weber, & Kimmerly 1993; Hollands & Spence 1998, 2001; Körner, 2004; Lohse 1993; Peebles & Cheng 2002, 2003; Pinker 1990; Schnotz, Picard, & Hron 1993; Shah, Freedman, & Vekiri 2005) have been proposed. See also Pinker (1990) and Shah and Hoeffner (2002) for some reviews of graph comprehension models.

Among these models, we find that the knowledge-based construction-integration model proposed by Shah and associates (eg, Carpenter & Shah 1998; Freedman & Shah 2002) is the most illuminative and probably most directly relevant to research on integrated writing assessment using graphs as prompts, because the knowledge-based model is analogous to Kintsch's (1988) Construction-Integration model of text comprehension. Graphs are a special kind of text, and graph comprehension is subsumed into a more general activity of text comprehension. Carpenter and Shah (1998) consider graph comprehension to be an integrated sequence of interactions between conceptual and perceptual processes, eg, "pattern-recognition processes that encode graphic patterns, interpretive processes that operate on those patterns to retrieve or construct qualitative and quantitative meanings, and integrative processes that relate these meanings to the referents inferred from labels and titles" (p 75). Freedman and Shah (2002) developed further this knowledge-based model of graph comprehension. According

to the knowledge-based model, graph comprehension is influenced not only by the display characteristics of a graph, including, for example, different types of graphs, their cognitive demands (Lewandowsky & Spence 1989), and the relative importance and relationships of graphical elements (eg, Feeney, Hola, Liversedge, Findlay, & Metcalf 2000; Shah & Carpenter 1995; Shah, Mayer, & Hegarty 1999; Simkin & Hastie 1987; Zacks & Tversky 1999), but also a viewer's domain knowledge (eg, Curcio, 1987), graphical literacy skills, and explanatory and other scientific reasoning skills [e.g. spatial reasoning, (Feeney, Adams, Webber, & Ewbank 2004; Stewart, Hunter, & Best 2008; Trickett & Trafton 2004, 2006)], and incremental interactions between these factors.

Further empirical evidence supports this knowledge-based model of graph comprehension. Specifically, in terms of the effects of the features of graphs on information extraction and processing, Meyer, Shinar and Leiser (1997) compared the relative efficiency of line graphs, bar graphs and tables, in different types of information processing tasks, and found that bar graphs had the advantage for reading exact values and identifying maxima, whereas line graphs had the advantage for reading trends. Furthermore, participants also tended to describe concrete contrasts in data presented in bar graphs (eg, higher, lower, greater than, less than); whereas when they saw line graphs, they tended to describe trends (eg, rising, falling, increasing, decreasing). In other words, it seems that there may be certain innate features and cognitive naturalness or conventions of different types of graphs that govern how readers would interpret and extract information from graphs. Even the same type of graphs may lead to different cognitive demands, for example, horizontal bar graphs were found to require longer decision making times than vertical bar graphs, and similarly negative number graphs required longer decision making times than positive number graphs (Fischer, Dewulf, & Hill, 2005). Furthermore, as Pinker (1990) pointed out: "different types of graphs are not easier or more difficult across the board, but are easier or more difficult depending on the particular class of information that is to be extracted" (p 111). However, overall, as Carpenter and Shah (1998) noted, "even relatively simple graphical displays require relatively complex cognitive processes" (pp 98-99). Pinker's model of graph comprehension predicts that it would be easier to make discrete comparisons between individual data points from bar graphs, using terms such as higher, lower, greater than, and less than; and easier to assess trends from line graphs, using terms such as rising, falling, increasing, and decreasing. Zacks and Tversky (1999) confirmed that readers had a strong tendency to describe discrete comparisons when they saw bar graphs, and describe trends when they saw line graphs – a phenomenon that Zacks and Tversky called "bar-line message correspondence". They also found that people produced bars to depict discrete comparisons and lines to depict trends – a phenomenon that they called "bar-line data correspondence" (p 1077). These two correspondences in comprehension and production of graphs conform to the principles of "cognitive naturalness" of using space to convey meaning in graphic communication (Tversky 1995).

This knowledge-based model of graph comprehension helps to explain the differences between novices and experts in comprehending graphs, and the relative ease or difficulty of graph comprehension encountered by these two groups. Similarly, Friel, Curcio and Bright (2001) identified four critical factors that appeared to influence comprehension of statistical graphs: (a) the purposes for using graphs, eg, whether for analyses or communication (Kosslyn 1989), (b) task characteristics (see Simkin & Hastie, 1987), (c) discipline characteristics, eg, spread and variation with a dataset, the type and size of the data, and the way a representation provides a structure for data (ie, graph complexity), and (d) characteristics of graph readers, including their prior knowledge or bias of the graphic information (Vernon 1946). These four critical factors are congruent to the key components of the knowledge-based construction-integration model described earlier. The important role that the characteristics of graph readers can play in graph comprehension has been widely supported with empirical evidence. For example, Carpenter and Shah (1998) noted that "individual differences in graphic knowledge should play as large a role in the comprehension process as does variation in the properties of the graph itself" (p 97). With reference to students' academic achievements, Åberg-

Bengtsson and Ottosson (2006) noted that graphic knowledge had the strongest correlation with mathematic/science achievements.

However, Roth (2002) suggested that graphic knowledge or expertise was more complex and we need to move beyond thinking about familiarity because not only knowledge but also experience and expertise of the phenomenon depicted in the graphs affects comprehension. He categorised graph comprehension along a continuum from "transparent", "competent" to "problematic" readings of graphs, based on the familiarity and knowledge of the graph readers. In "transparent reading", the familiar graph provides the readers with a "transparent window onto a familiar world", the graph and the phenomenon have "fused" (p 5). "Readers no longer think of words, or parts of a line curve, but go directly to the things they know them to be about" (p 6). In "competent reading", graphs and their topics are less familiar, "more work is required on the part of a reader" (p 6). In "problematic reading", "people are unfamiliar with graphs, phenomena, or the translation between the two, problems in reading become apparent. Most of the reading activity is then concerned with structuring the graph (and accompanying text) itself …rather than with relating it to some phenomenon" (pp 8-9).

In the field of second or foreign language testing, there are very few studies that address the effects of visual input, such as graphs, pictures and diagrams, on test performance, although graphs have been used quite widely as test prompts. Furthermore, we noticed that among these few studies that addressed the effects of graphic inputs on test performance, they were often conducted in the context of listening (eg, Ginther 2002; Gruba 1997) and speaking assessments (eg, Katz et al 2004; Xi 2005), and much less in writing assessment, except for O'Loughlin and Wigglesworth (2003). The TOEFL Program reported that it discontinued temporarily the chart-graph format with explicit comparison statement in its Test of Written English pending further investigation in early 1990s (Golub-Smith, Reese, & Steinhaus 1993), due to the findings that this format (see *ibid*. 18) produced the highest mean scores, compared to other formats without graphs. However, the findings of these studies in the field of language testing have provided empirical evidence of the potential effects of test takers' graphicacy on test performance. For example, Xi (2005) found that test takers' graph familiarity affected their oral reproduction of the information accessible to them in graphs and that graph familiarity had differing effects on the holistic scores of the speaking test based on the bar and line graph tasks. The individual and probably idiosyncratic differences in terms of the effects of the features of graphs and test takers' graph familiarity raise questions on the validity and fairness of such tasks. Although the analyses on the effects of the quantity and the manner of presentation of information in graphs, as done by O'Loughlin and Wigglesworth (2003), shed some light on the validity of the IELTS AWT1, such studies have not provided insights into the ways the characteristics of the graphs and the test takers' graphicacy and the interactions of these factors contribute to the test takers' performance. The validity of AWT1 tasks should be investigated not only by analysing the written scripts of the tasks but also test takers' cognitive processes when doing the tasks. This, then, has both informed and shaped the focus of this research.

2 RESEARCH AIMS AND QUESTIONS

2.1 Research aims

The overall aim of this study is to explore the cognitive processes of intending IELTS test takers completing AWT1 that uses graphs as test prompts. It aims to understand the validity of AWT1 in a dynamic, rather than "one-off" approach. In other words, this validation study aims to explore the cognitive processes at different time points – before and after test preparation activities. As such, this study will thus also develop insights into the impact of test preparation practice on task performance, so as to explore the validity and fairness issues of intensive preparation for AWT1 tasks. In addition, the possible differential effects of test preparation activities on test takers' cognitive processes will provide invaluable information for understanding the dynamics of the cognitive processes involved (Carpenter and Shah 1998).

2.2 Research questions

The study addressed the following research questions (RQ):

RQ1: To what extent are there differences in the candidates' cognitive processes due to different AWT1 prompts?

RQ2: To what extent are the candidates' cognitive processes affected by their graphicacy?

RQ3: To what extent are the candidates' cognitive processes related to their writing abilities?

RQ4: To what extent are the candidates' cognitive processes influenced by test preparation opportunities offered to them by the research team?

3 RESEARCH DESIGN AND METHODOLOGY

3.1 Approach

We employed a case study and grounded approach to this research, as this was congruent with the focus of the study and the importance of gaining in-depth understandings of the cognitive processes of taking AWT1. We consider the 'case study' approach to be a strength, not a limitation, as it afforded the opportunity, on the one hand, to collect rich and in-depth data on the cognitive processes that large scale product-oriented quantitative studies may not be able to provide, and on the other, to develop hypotheses for further research in a wider number of contexts (eg, with participants whose first language is not Chinese, or/and who have different graph familiarity level). We also notice that the majority of studies investigating test-taking processes in the field of language testing have used a "one-off" approach, and has not – as we proposed – collected data from the same participants at different time points. This last design feature is what is unique about the current research to understand the cognitive processes of taking AWT1 tasks.

We followed the University of Bristol's research ethics and data protection guidelines and regulations during the whole process of this project. Written consent was obtained from each participant.

We set out our research methodology below.

3.2 Participants

This research was conducted in a leading university in China. It was chosen because a large number of its undergraduate and postgraduate students take the IELTS academic module each year. In this respect, this research will be more beneficial for IELTS partners to gain a better understanding of the validity of AWT1 than if a smaller university which has fewer prospective IELTS test takers was selected. With the co-ordination of its Foreign Affairs Office, the Graduate School and the student Society of International Communication and Exchange, calls for participation were sent to all

departments across the university through its administrative intranet which all staff and students access on a daily basis, and to the university's public communication platform specifically for its students. As expected, there was a considerable amount of interest. Over 380 students (and 1 teaching staff member) emailed us their personal details to apply for a place. We envisaged that a large number of students would sign up because of the possible benefits of having free training on how to take AWT1 as offered by the project. However, due to the nature of the case study approach to collecting data on cognitive processes, we could select only a small number of them to participate, by taking into account their background characteristics such as gender, department (science, social sciences, or arts) and academic status (ie, undergraduate or postgraduate including Master and Doctoral students). We applied the following criteria in the order indicated below, by examining the personal details that the applicants provided. These details of the selected participants were re-collected later in the graph familiarity questionnaire (see section 3.3) to double-check their accuracy.

1. For undergraduates, they must be in their final year (or fourth year for medical students who are in 7-year programmes) of their degree programme. For postgraduates at Master level, they must be in their second year (some Master programmes in the university require 3-4 year study); and for postgraduates at doctoral level, they can be any year of their study.

2. The participants must have no experience of taking the actual IELTS test, but are planning to take IELTS at the end of 2008 or early 2009. We anticipated that these participants would be more engaged in this research project. Students who did not provide this information in the application form were not selected.

After applying the two criteria above, we still had 121 students remaining in our database: 21 doctoral students, 28 undergraduates and 72 M-level students. A follow-up email was then sent to these students asking them their availability between the end of October and early November. This information also helped the research team to have a sense of the students' commitment to the research project. This reduced the number to 53 students who could be available every day during that week. Among these students, we selected randomly 24 as participants. In case there were dropouts during the first session, we operated a waiting list. Prior to data collection, a few students informed us that they were not able to attend the first session for various reasons, so we phoned those in the waiting list and still managed to have 24 participants when the data collection commenced (see Table 1). There were 13 female and 11 male students from different departments across the university such as economics and finance, management, psychology, physics, chemistry, biology and food science, computer science, agriculture, civil and electrical engineering, medicine and pharmaceutical sciences. Among them, 12 were undergraduates, 12 postgraduates (including 5 PhD candidates); 17 in science and engineering subjects and 7 in social science and arts. The subject areas of the selected participants reflected the overall picture of the specialisms of the 380 students who expressed their interest in participation. None of them had taken official IELTS tests, but all of them were planning to take IELTS at the end of 2008 or early 2009. For ease of reference in this report, we assigned each participant a code, from A to X.

		Status			
		MSc	***PhD***	***UG***	***Total***
gender	F	5	1	7	13
	M	2	4	5	11
Total		7	5	12	24

Table 1: Cross-tabulation of 24 participants by gender and academic status

However, at the start of Stage 2 data collection (see section 3.3), 5 dropped out (Students F, M, O, P, V). At the start of Stage 4 data collection, 1 more dropped out (Student A). Therefore the final complete dataset consisted of 18 participants (see Table 2): 11 female, 7 male; 11 undergraduate, 7 postgraduate (incl. 2 PhD candidates).

		Status			
		MSc	***PhD***	***UG***	***Total***
gender	F	4	1	6	11
	M	1	1	5	7
Total		7	5	12	24

Table 2: Cross-tabulation of the final set of 18 participants by gender and academic status

3.3 Data collection

This research collected both qualitative and quantitative data at five stages with different research focuses and instruments, as summarised below in time order.

STAGE 1: Baseline data collection and think-aloud training

1 At the first meeting, the participants were briefed about the purpose of the project and the schedule of data collection. Written consent was obtained from each participant (see Appendix 1).

2 Administration of IELTS Academic Writing Tasks 1 and 2 (see Appendix 2), developed by the research team according to *IELTS Handbook* (2006), to measure the participants' writing abilities without thinking aloud. The AWT1 task asked the participants to summarise, making comparisons where relevant, the main features of a line graph reporting the UK greenhouse gas emissions by different end users between 1990 and 2003. The AWT2 task asked the participants to write on the extent to which they agreed or disagreed with the following statement: "Once children start school, the teachers would have more influence in their intellectual and social development than parents".

3 Administration of the graph familiarity questionnaire to understand the participants' graph familiarity and comprehension (see also Xi 2005). This self-evaluation questionnaire contained questions on: the participants' experience of using graphs in their academic study and reading; their familiarity with different types of graphs; their ability in comprehending and interpreting graphs; and their ability in using graphs in written communication and in assisting learning (see Appendix 3). The first 11 questions were used to re-collect the participants' personal information such as their gender, IELTS experience, specialism, and year group (see section 3.2 for the participant recruitment procedure). As "graph" can potentially mean different types, we provided not only Chinese explanations on what "graph" means but also examples of graphs such as bar graph, line graph, pie chart, statistical table, and flow chart.

4 The participants were trained on how to think aloud (see Appendix 4). They were first asked to think aloud when solving some simple mathematics word problems, and then think aloud when doing an IELTS AWT1 task. An example of think aloud was then provided for them to compare it with their own think aloud protocols to raise their awareness of the importance of keeping thinking aloud. The participants were allowed to think aloud in English and/or Chinese as appropriate and comfortable.

STAGE 2: First round of collection of the participants' cognitive processes (ie, pre-training)

1 At the beginning of Stage 2 data collection, the participants were given further brief training and practice in think-aloud using the first AWT1 task they did the other day. Only when they were comfortable with think-aloud did they start doing the AWT1 tasks (see Appendix 5). None of the graphs used in the AWT1 tasks is 3-D in order to avoid a potentially extra information processing load on the participants (Carswell, Frankenberger & Bernhard, 1991).

2 Administration of the first set of four AWT1 tasks, printed in colour (see Appendix 5, Set A). These graphic prompts included

- a simplified diagram/drawing showing the changes of the landscape or layout of an area from 1937 to 1995 (hereafter A-layout)
- a numerical table showing the top countries from which the USA imported crude oil (hereafter A-oil)
- a stacked bar chart reporting the instruction hours of children in OECD countries (hereafter A-instruction)
- a line graph describing the UK CO_2 emissions trend (hereafter A-UKCO_2).

The AWT1 tasks were assigned to each participant in random order, one at a time (25 minutes for each task, ie, 5 minutes more than the IELTS recommended 20 minutes due to the extra requirement of thinking aloud). The participants were asked to think aloud while completing the writing tasks. Each was given a digital voice recorder to record their think-aloud process. In order to minimize the influence from each other and avoid recording other people's think-aloud in his/her own recorder, several measures were taken to ensure:

- only four participants a time were allowed into a very big classroom
- they were seated in the four corners of the classroom, and
- at no point did participants have the same writing prompt.

Field notes on how the participants were engaging with the tasks were taken by the first author.

STAGE 3: Training on AWT1 tasks

In the short training session of just over two contact hours, the first author presented to the participants an overview of the IELTS test and the types of AWT1 task prompts and rating criteria. Examples of the participants' own recorded think-aloud protocols in Stage 2 were used to illustrate the cognitive processes and strategies they had used to complete AWT1 tasks and the problems as shown in the think-aloud protocols that the participants should try to avoid. In particular, it was emphasized that they should describe only the information contained in the graphs and make comparisons where

necessary but should try to avoid making unnecessary personal conclusions not based on the data presented in the graphs (see also Section 4.3.4). The students were also reminded of the dos and don'ts as suggested by Cambridge ESOL, and some frequently used words and phrases for describing the process and trend of decreasing and increasing, and those for describing comparisons. In addition, they were given 12 AWT1 tasks using a variety of graphs for them to practise after the training session or at least to make themselves familiar with the different types of graphs that might be used in AWT1 tasks. A formal training booklet (which is available from the authors upon request) including the key messages mentioned above was provided for every participant.

STAGE 4: Second round of collection of the participants' cognitive processes (ie, post-training)

1 Following the same procedure of Stage 2, although no training on think-aloud was given this time as they were by then more familiar with this data collection tool, the participants were asked to think aloud while completing the second set of 4 AWT1 tasks (Appendix 5, Set B). The four tasks are:

   – a diagram showing the sequence of the events leading to the broadcast of a documentary by a TV programme (hereafter B-broadcast)

   – a numerical table and the related coloured world map showing the amount of CO_2 emissions by the top eight countries (hereafter B-map)

   – column and pie charts describing grams of CO_2 emissions per passenger/km of different transport methods and the EU funds spent on them (hereafter B-EUfund),

   – a line graph showing the individuals viewing share of the major TV channels in the UK (hereafter B-viewing).

2 Students were also asked to evaluate the training and their experience in this project and provide any suggestions on AWT1 preparation. This was not originally planned in our research proposal, but we feel it can serve the central purpose of understanding how the training may have affected their cognitive process of taking AWT1 tasks (ie, RQ4). The students filled in the questionnaire (see Appendix 6) either right after the final AWT1 task, or completed it at home and returned it the next day when they came for the interviews.

STAGE 5: interviews

Due to six dropouts in total, we decided to modify the original plan of individually interviewing 50% of the 24 participants based on their gender, graphicacy and writing abilities. Instead, we interviewed all 18 participants. The participants were asked to comment on (i) the ways their cognitive processes may have be affected by the different graphic prompts, (ii) the relationship between their graph familiarity and comprehension and writing abilities, and (iii) whether their cognitive processes might have changed due to the training (see Appendix 7 for a list of the questions that guided the semi-structured interviews). The interviews ranged from 45-68 minutes each, conducted in English and/or Chinese as appropriate and audio-recorded.

In summary, this research comprised five distinct stages and adopted a layered and progressive approach to data collection. The data included the participants' academic writing performance without thinking aloud, graph familiarity and comprehension (Stage 1), AWT1 performances of different graph prompts while thinking aloud (Stages 2 & 4). Between Stages 2 and 4 (ie, Stage 3), training on AWT1 was provided to the participants. At Stage 5, interviews with the participants were conducted, aiming to probe further their cognitive processes of taking AWT1 tasks.

4 ANALYSES AND FINDINGS

The data were analysed first of all to understand the participants' graph familiarity and their writing performances under the normal examination conditions and under think-aloud conditions (4.1). Section 4.2 presents an overview of the participants' cognitive processes. Section 4.3 addresses the four research questions one by one.

4.1 Participants' graph familiarity and English writing abilities

4.1.1 Participants' graph familiarity

Data on the participants' graph familiarity was collected via the questionnaire (Appendix 3) at Stage 1 from 24 participants (see Table 1). Except Questions 42-44 (see below), the participants' responses were used to calculate their graph familiarity score. Their responses to all the remaining questions, except Questions 18-20 and 40-41, were added together, with a bigger number representing higher graph familiarity. Responses to Questions 18-20 and 40-41 where a smaller number represented higher graph familiarity (ie, negative questions), were recoded to be consistent with the majority of the questions. Questions 42-44 were quite neutral: responses to bigger/smaller numbers did not necessarily mean higher/lower graph familiarity. They were not used to calculate the total score of graph familiarity, rather they were analysed separately (see section 4.3.2) from the other questions. As a result, there are 32 questions included in the calculation of the total score of graph familiarity. Reliability analyses of the scale (32 items) indicated an Alpha of 0.948. Overall, we trust the participants responded to the questions consistently and the measure of their graph familiarity using this questionnaire achieved satisfactory reliability. Using the 32 items to measure the participants' graph familiarity, the maximum possible score would be 192 (32 x 6) and the minimum 32 (32 x 1). Within this sample, we observed a minimum of 96, maximum of 182, mean of 143.8 (ie, around 75% of the total maximum possible score), std. deviation=22.9. The graph familiarity is close to a normal distribution (Kolmogorov-Smirnov Z=0.65, n.s., see Figure 1). The data suggests that the participants were quite familiar with graphs. As part of their academic study, a third of the participants (8/24) *very often* (ie, the highest of the 1-6 scale) used special computer software to produce graphs, nearly 46% (11/24) *very often* needed to produce graphs, and a quarter (6/24) *very often* needed to interpret graphs. For further details of these participants' graph familiarity, see Appendix 8 which reports the frequency statistics of their responses to each question.

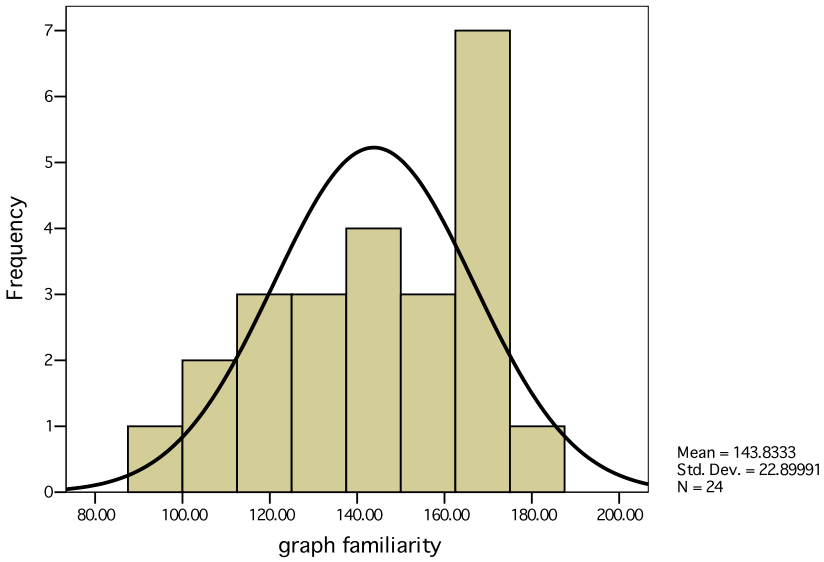


Figure 1: Graph familiarity of the 24 participants

Although we understand this is a small sample, we also tried to examine if there was any meaningful difference in graph familiarity between male and female participants. It was noted that there was no statistically significant difference in graph familiarity between male and female participants: Mean of 11 male participants = 151.4, std. deviation=19.0; mean of 13 female participants (n=13) =137.5, std. deviation=24.7 (t=1.52, df=22, n.s.)

The above statistics included all the 24 participants. However, as six of them dropped out, we did separate analyses using the remaining 18 participants only: very similar findings were noted (see Figure 2). Mean = 143.6, std. deviation = 24.96, min=96, max=182; and the difference between female and male participants was not significant either.

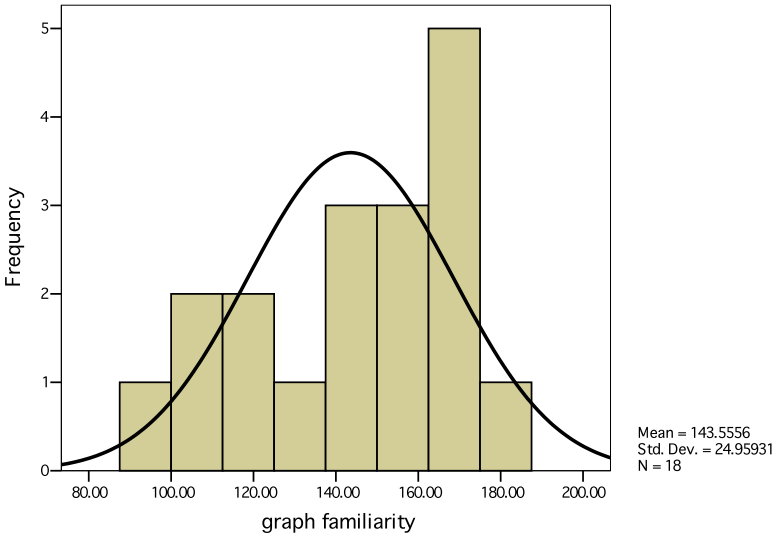


Figure 2: Graph familiarity of the 18 participants

Below we report the participants' AWT1 writing performance under normal examination conditions and when thinking aloud.

4.1.2 Participants' writing performances under normal examination conditions

Under normal examination conditions without think-aloud, the participants completed two writing tasks: AWT1 and AWT2 (Appendix 2). Their written scripts were Word processed before being double marked by the research team according to the IELTS writing band descriptors (public versions). Each researcher independently marked 2/3 of the printed writings randomly assigned to them, to ensure double blind marking (see Table 3). If the difference in marks between two raters was bigger than 1, a third rater marked the writing in question. If the third mark was the average of the previous two marks, the third mark was reported as the final mark of the writing; otherwise, the average of the two most adjacent marks was reported as the final mark of the writing (Note: We retain ¼ and ¾ marks in the analysis). If there was still no agreement among the three raters, we then had face-to-face moderation discussions. The initial agreement between raters before moderation was not satisfactory, largely due to the different interpretations of the first set of rating criteria of AWT1 – "task achievement".

Script ID	Rater 1	Rater 2	Rater 3
1	×	×	
2		×	×
3	×		×
4	×	×	
5		×	×
6	×		×
7	×	×	

Table 3: Double blind marking scheme

At the moderation meetings, we discussed the scripts in question and marked them together. The moderation exercise was an essential step to improve rater reliability. The average inter-rater reliability Cronbach's Alpha was improved (above 0.83).

Marking the AWT2 scripts was much more straightforward and achieved high inter-rater consistency before moderation: Cronbach's Alpha = 0.843 between Rater 1 and Rater 2, 0.881 between Rater 1 and Rater 3, 0.793 between Rater 2 and Rater 3. Only two scripts had difference larger than 1 and were marked by a third rater. The participants' performances in these two writing tasks are reported below.

	N	Minimum	Maximum	Mean	Std. Deviation
AWT1 no think-aloud	24	4.50	7.00	5.69	0.673
AWT2 no think-aloud	24	4.50	7.50	6.32	0.764

Table 4: Performance in AWT1 and AWT2 without think-aloud

As shown in Table 4, the mean of the participants' performance in AWT1 and AWT2 under normal examination conditions without think-aloud was 5.69 (std. deviation=0.67) and 6.32 (std. deviation=0.76) respectively. Below we report the participants' performances in the eight tasks under think-aloud conditions.

4.1.3 Participants' writing performances under think-aloud conditions

The writings of the eight tasks under think-aloud conditions were randomly assigned to and double blind marked by the research team, using the same procedure as for marking the first AWT1 task under normal examination conditions (see section 4.1.1). Overall, the inter-rater reliability Cronbah's Alphas were satisfactory: 0.69 for A-instruction scripts, 0.86 for A-layout scripts, 0.85 for A-oil scripts, 0.90 for A-UKCO_2 scripts, 0.78 for B-broadcast, 0.74 for B-EUfund, 0.71 for B-map, and 0.79 for B-viewing. The participants' performances in the eight tasks are reported in Table 5 below.

	Type of Graph(s)				*Mean Statistic*	*Std. Error*	
Task A: instruction	Bar graph	18	5.00	7.00	6.00	.151	.642
Task A: layout	Linear drawing	18	5.00	6.75	6.10	.115	.486
Task A: oil	Statistical table	18	4.75	7.00	5.93	.164	.696
Task A: UK CO_2	Line graph	18	4.75	7.25	6.03	.150	.635
Task B: broadcast	Diagram of sequence of events	17	4.75	7.25	6.00	.165	.679
Task B: EU fund	Bar graph and pie chart	17	4.75	7.00	6.04	.144	.595
Task B: map	World map and statistical table	18	5.25	7.25	6.36	.107	.456
Task B: viewing	Line graph	18	5.00	7.50	6.50	.181	.770

Table 5: Performance in the eight AWT1 tasks with think-aloud
One student (# T) wrote down what she was thinking aloud, rather than a piece of writing as responding to the AWT1 tasks. Her data were not included in the analyses of the participants' performance, but included when analysing the cognitive processes.

As shown in Table 5, the participants' performance ranged from around 5 to 7.5, with mean scores from 6 to 6.5.

4.2 Developing a working model of cognitive processes for AWT1 tasks

The think aloud protocols and the post-task interviews were transcribed, coded and categorised in a qualitative data analysis computer programme – winMAX (Kuckartz 1998), to identify the patterns of the cognitive processes of completing AWT1 tasks. This grounded approach is appropriate to understand the complexity of the qualitative data of the cognitive processes (see section 3.1). In order to develop a working model of cognitive processes for AWT1 tasks, we analysed the think-aloud protocols, cross-referencing to the other four main sources of data in this research project (ie, AWT1 scripts, interviews, graphicacy, and English writing abilities) and the theories of graph comprehension, especially the knowledge construction-integration model (see section 1.3). Below we explain in detail the working model which emerged from the data.

The AWT1 tasks seemed to require three key consecutive processes:

1 comprehending non-graphic task instructions

2 comprehending (and interpreting) the components of graphs (see Section 1.3 on the key factors that can influence graph comprehension)

3 re-presenting or re-producing the non-graphic and graphic information as continuous discourse in written form in English as a foreign language.

Although these three processes were found to be iterative and to vary among the participants in terms of, for example, the amount of time they spent on each process for different tasks and the differential interactions with and influences on test performance of other factors such as the participants' prior domain knowledge about the information presented in the graphs, the conventions of graph presentation, graphic skills and English writing abilities, it seemed to be very clear that these three processes happened, first and foremost, in time order. However, overall, the whole process for completing AWT1 tasks, as graph comprehension, is iterative and recursive, rather than serial (Carpenter & Shah, 1998).

4.2.1 Comprehending non-graphically presented task instructions

All the participants started to read the non-graphically presented task instructions first, presented before the actual graph(s). However, not all text instructions received the same attention or effort from the participants. They seemed to be very strategic in allocating their time to different components of the task instructions. It is interesting to note that all of the participants skipped the first line (ie, "*You should spend about 20 minutes on this task"*) and the last line ("*Write at least 150 words"*) of the task instructions, although both sentences were clearly marked in bold. This may be largely because the participants were already familiar with these two specific task requirements – time allowance and the expected length of their writings.

What the participants repeatedly read and re-read at different stages of completing the tasks were two sentences. The first was the summary-like sentence about the graph (eg, "The following graph shows the UK CO_2 emissions by end users from 1970 to 2004" in Task A:UKCO_2, "The following map and table show the amount of CO_2 emission by the top 8 countries" in Task B:Map). This introductory sentence gave the participants a clear lead as to what the following graph was all about (ie, the topic or the theme of the graph), and therefore may have facilitated the participants to process the information contained in the graph (see Section 4.2.2). However, the summary-like sentence served only as an entry point, in other words, the participants still had to work out what else they should focus on when looking for the main features and making comparisons. For example, Participant K reflected in the interview on the summary-like sentence in Task A: Instruction:

Extract 1

> "I read the direction, it says different age group. Now that it says different age group in the task direction, so I wondered if it is necessary to figure out and mention the differences with reference to different age groups. After writing about the total length of the bars, ie, the total number of intended instruction hours, I didn't know if it is necessary also to write about the differences between countries in the different age groups." (Participant K)

In Section 4.2.3 we report on how the non-graphic task instructions were re-presented or re-produced in the actual writings.

The second sentence – "summarise the information by selecting and reporting the main features, and make comparisons where relevant" – was in fact the same for all AWT1 tasks. Why didn't the participants skip this sentence as they skipped the sentences on time allowance and expected length because they had already known that this was exactly the same requirement for AWT1 tasks – summarise main features and make comparisons? Contrary to our anticipation that the participants would pay less attention to this sentence after they became more familiar with the AWT1 task requirement, we noticed that there were more occurrences of re-reading of this sentence in the think-aloud protocols of the second set of four tasks than the first four tasks. Data from the interviews and end-of-project evaluation (Appendix 6) indicated that this was largely attributable to the short training (see data collection in Stage 3). Before the training, the participants had a strong tendency to interpret what they observed in the graphs and try to find the reasons behind the data. At the training session, we explicitly pointed out that the main task was to describe rather than to interpret the data, in other words, the main task was to summarise and make comparisons rather than make personal interpretations based primarily on prior domain knowledge. Therefore, the participants were probably made more aware of this requirement in the second four tasks after the training, and they might have used re-reading this sentence as a kind of constant reminder to themselves that their major task was to summarise and make comparisons of what are in the graphs, but not to involve too much personal interpretation (see Section 4.3.4 on the coachability of AWT1 tasks).

4.2.2 Comprehending graphic information

After comprehending the non-graphic task instructions which were placed before the graphs, the participants moved on to read the graphs, look for main features of the graphs, and make comparisons. It seems to be a quite natural cognitive process and transition from focusing on the non-graphic information to the graphic information. The comprehension of graphic information involved a series of iterative and internal processes (see Section 1.3), starting from reading, understanding, deciding on the main features, to interpreting the graphs, prior to re-presenting and re-producing them in continuous written discourse in English as a foreign language (see Section 4.2.3). The think-aloud protocols demonstrated a series of activities involved in graph comprehensions, corresponding to what Curcio's (1987) three levels of graph processing and comprehension: read the data, read between the data, and read beyond the data.

When reading the data or the information contained in a graph, the participants started to search and locate specific information and try to perceive trends and patterns, ie, they started from local to global search of graphic information (see also Guthrie *et al* 1993), constantly checking the graphic information against their prior knowledge about the subject or the content of the graphs (eg, CO_2 emissions, environmental protection, and crude oil importation) as well as their prior knowledge about the types and conventions of graphs. The discriminated symbols and conventions used in graphs and the ways they were configured were meaningful to graph readers, and anticipated certain type of information processing, eg, fact-retrieval from line or bar graphs, trends from line and bar graphs, making proportion judgements from pie charts, making comparison judgements and determining the slope of the trends from a regression line. In other words, the types and conventions of graphs influenced the way how comparisons can and should be made, following the cognitive naturalness of graphic communication conventions. In addition to the prior knowledge about the content and conventions of graphs, the participants' graphicacy, explanatory skills and other scientific reasoning skills also influenced, in varying degrees, the cognitive processes of completing AWT1 tasks and subsequently their task performance (see section 4.3.2 on the effects of graphicacy).

4.2.3 Re-presenting graphic and non-graphic information in continuous discourse

As described in sections 4.2.1 and 4.2.2, the iterative cognitive processes involved various factors including, for example, reading non-graphic task instructions, reading and interacting with the display characteristics of the graphs (including most prominently title/caption, x-axis, y-axis, source of data, legend, and colour schemes) and their conventions, understanding the graphs, extracting the key information or data points, making comparisons of main features, and trying to find out the underlying reasons for the phenomenon being described (although this is not required by the AWT1 tasks). As responses to the requirements of AWT1 – essentially a writing task, the participants were constantly planning and organising content (including choosing specific words to make trend assessments, discrete comparisons, or predictions, following the conventions of graphic presentations), checking the accuracy of linguistic forms and content, self-monitoring and evaluation of writing (including re-reading what has been written and counting number of words of their writings). The knowledge construction-integration model of graph comprehension (see 1.3) seems to be fit for our purpose of explaining and evidencing some of the cognitive processes of completing AWT1 tasks, especially during the graph comprehension stage. However, AWT1 tasks are unique in the sense that they involve not only candidates' processing of graphs and but also more importantly, re-producing the graphic and the non-graphic information in a continuous written discourse in English as a foreign language. It is this writing process that makes AWT1 a unique integrated reading/writing task. It is this writing process that we noted that in effect facilitated the participants to better interact with graphic information and enhanced their graph comprehension through writing. In other words, the comprehending and the re-producing of graphic information were found to be mutually beneficial.

As AWT1 tasks are a kind of writing activity, primarily knowledge telling, to reproduce meanings already depicted in the task prompts, candidates are to some extent able to copy and deploy the lexical and syntactic components of the non-graphic prompts in their own writings. The importance and the value of the summary-like sentence in the non-graphic task instructions are clearly evidenced in the participants' actual writings. Taking the scripts of Task A:UKCO_2 as an example, we noticed that all the 19 participants almost unanimously started their writing with this summary-like sentence or its slight variation, eg, "The line chart illustrates the trend of the UK CO_2 emissions by different end users between 1970 and 2004 in million tonnes carbon equivalent" (Participant A), "The graph in the paper show UK CO_2 emissions by end user from 1970 to 2004" (Participant B), "The graph in the picture indicates UK carbon dioxide emissions by end user from 1970 to 2004" (Participant C), "From the graph above, we can see that the UK CO_2 emissions by end user has fallen down in 34 years" (Participant D), "It's a graph which shows the UK CO_2 emission by end user from 1970 to 2004" (Participant E). It is the same case for the other seven tasks too. Consistently, the participants started their writings, with almost verbatim copy of this summary-like sentence from the task instructions, and then provided further details in subsequent paragraphs. Asked why they started with this copying exercise at the interviews, all the participants said that there couldn't be a better introductory sentence than this authoritative one in the task instructions.

4.2.4 Summary of the working model

The working model of the cognitive processes for AWT1 tasks is presented in Figure 3. The vertical timeline indicates the time allowed for AWT1 tasks, and the positions of the three key processes along the vertical timeline in the left column represent approximately the average duration of each process. The overlapping between "graph comprehension" and "re-producing graph comprehension in continuous written form in English" indicates that some students developed their understanding of the graphs through writing rather than following the linear process of graph comprehension and then writing. The participants may also change and improve their understanding of the graphs, when they are writing. At this first stage of "comprehending non-graphically presented task instructions" and its associated activities, especially the summative introductory sentence. Comprehension of this sentence directly affects the next stage, ie, graph comprehension, which is further governed by the display characteristics of graphs, the characteristics of graph readers, and their interactions, plus "the purposes of understanding the graphs and the cognitive demands of the writing tasks". At the third process, ie, re-producing graph comprehension in continuous written discourse in English, the key aspects include writing and revising the main features of graphs, writing and revising the comparisons where relevant, writing personal interpretations (sometimes), and constantly monitoring and self-evaluating their writings.

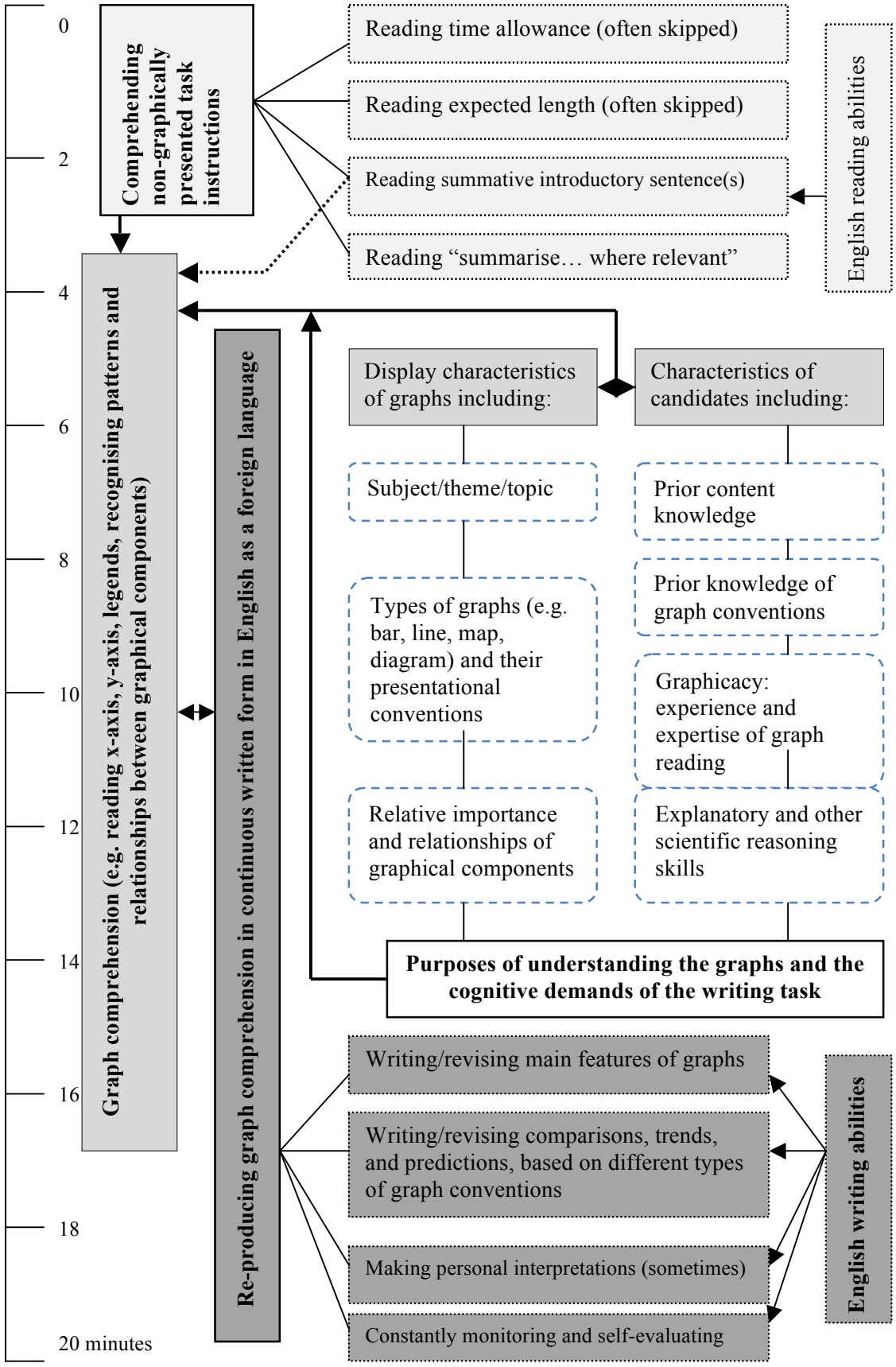


Figure 3: A working model of cognitive processes for taking IELTS AWT1 tasks

Overall, the participants' English writing abilities should play a pivotal role, at least in theory; and at this stage, the more generic second language writing models become more applicable for AWT1 tasks.

This working model will be used to guide our analyses of the empirical data to address the four research questions in Section 4.3.

4.3 Addressing the four research questions

4.3.1 Research question 1: To what extent are there differences in the candidates' cognitive processes due to different AWT1 prompts?

In terms of cognitive processes of completing AWT1 tasks, we identified three key stages: comprehending non-graphic task instructions, comprehending graphic information, and re-producing graphic and non-graphic information in continuous discourse (see Figure 3). Think-aloud protocols, interviews and AWT1 scripts form the three main data sources to investigate the extent to which the cognitive processes were due to the use of different graphic prompts.

As shown in Table 5 (see section 4.1.3) and Appendix 5, we used different types of graphic prompts, including bar graph (Task A:Instruction, Task B:EUfund), linear drawing of the layout of an area (Task A:Layout), statistical table (Task A:Oil, Task B:Map), line graph (Task A:UKCO$_2$, Task B:Viewing), diagram of the sequence of events (Task B:Broadcast), pie chart (Task B:EUfund), and world map (Task B:Map). Two of the tasks used two types of graphic prompts. In Task B:EUfund both bar graph and pie chart were used, and in Task B:Map, world map and statistical table.

As reported in section 4.2.1, there does not seem to be any significant difference in the way the participants processed the non-graphic task instructions. All the participants (except Participant T) started with the non-graphic task instructions. When doing Task A:Layout, Participant T started by analysing the graphic information first, according to her think-aloud. Overall, it is only when comprehending graphic prompts and re-producing the main features of the graphic information in written discourse that the participants demonstrated various differences in their cognitive processes.

We noticed that the participants' performance in the eight tasks varied significantly, in terms of the final scores they received (Chi-square=19.12, df=7, n=17, $p<0.01$). As shown in Table 5 (section 4.1.3), scripts of Task B:Viewing which used a line graph received the highest mean score (6.5) among the eight tasks, and those of Task A:Oil which used a statistical table the lowest mean score (5.93). The participants considered line graphs the easiest as the trends and the slopes are readily perceivable from the visual displays. Line graphs were also the most familiar to them (see Section 4.3.2). This may explain why scripts of Task B:Viewing received the highest mean score. However, it may not be necessarily the case that all AWT1 tasks using line graphs are always the easiest. Task A:UKCO$_2$ which also used a line graph presented a different situation. The average performance of Task A:UKCO$_2$ was not much better than tasks using other types of graphs. The lowest mean score of the scripts of Task A:Oil may well be due to the fact that it used a statistical table which is different from and probably more cognitively demanding than line graphs, as Fry (1981) explained:

> "Graphs pack a high density of information into a small area… are more globally visible than they are detailed, symbolic, and sequential… tend to show the 'big picture' or gestalt…. Often relationships can be seen better with a graph than with a purely verbal or numerical information." (Fry 1981: 388).

Perhaps also as the two 19th century economists commented: "Getting information from a table is like extracting sunlight from a cucumber" (Farquhar & Farquhar 1891), cited in Wainer (1992: 18), statistical tables are challenging to understand. However, not all tables are born equal. As Lohse (1993) pointed out, the cognitive demands of processing tabular information depended on tasks.

> "Tables do not require a complex mapping of syntactic symbols to semantic information. The semantic information is read directly from the table. This greatly reduces the information load in STM (i.e. short term memory, added by the authors) when the queries are simple. However, for complex comparisons of trends, tables require the reader to look at several entries in the table and make mental calculations and comparisons among these values. These intermediate values are held in STM during the comparison. In contrast, line graphs … show trend information without the scanning and interpolation of intermediate data. Slopes are perceived readily from a line graph." (Lohse 1993: 360).

Similarly, Vessey (1991) argued that graphic representations emphasize spatial information, while tabular representations emphasize symbolic information, hence it is more time-consuming to process tabular information. However, Vessey (1991) further suggested that so long as there is cognitive fit (i.e. matching between the nature of the representation, the process and the nature of the task) each representation, whether graphic or tabular, would lead to both quicker and more accurate problem solving. In case of AWT1 tasks, as all participants were required to describe whatever graphic and tabular information they were given, therefore the level of cognitive fit varied from participant to participant, and from task to task. The cognitive demands for processing tabular information may be particularly high for writing tasks such as AWT1 as it requires not only locating and extracting specific information but also integrating and identifying the overall patterns and making comparisons from the tabular information which may not be readily available from the table per se. The participants had to work out the patterns by themselves, through a series of calculations and comparisons in rows and columns of the statistical table. They felt that the statistical table on USA oil imports was not straightforward, as several participants commented, for example:

Extract 2

> "The oil statistics, it takes a long time to understand the information, to understand what it means….I was really puzzled, trying to work out what it means. It is quite exercising." (Participant T)

Extract 3

> "The statistics look very much the same, and I found it very difficult to present the key information from the table. But the line graphs, bar graphs and pie graphs are relatively straightforward and vivid…" (Participant U)

When the tabular information is presented in conjunction with other graphic prompts, eg, Task B:Map which used the world map and a statistical table with fewer data points than the statistical table in Task A:Oil, the cognitive demands for processing tabular numerical information seemed to be reduced. On average, the scripts of Task B:Map received the second highest mean score (6.36) among the eight tasks. The visual assistance of the world map in colour might have assisted the reading and the interpretation of the numerical data contained in the table. In addition, the fewer data points of this table might have rendered itself more manageable than the statistical table about US oil imports from 15 countries in June and July of 2008 and July 2007 and Year to Date of 2007 and 2008. The difficulty of this complex statistical table in Task A:Oil was further increased because of one particular phrase – Year to Date. All the participants knew these three words separately, but none of them knew exactly the statistical meaning of YTD. Similarly in Task B:Viewing, some participants seemed to have difficulty in understanding what MCH (multi-channel) meant. It seems that a single phrase in the accompanying text of a graph could increase substantially the difficulty level of the graph. Coupled

with the higher cognitive demands of statistical tables per se, the participants' lower familiarity with processing tables of numerical data had probably also made AWT1 tasks using statistical tables more challenging than other tasks. The participants reported that they were less familiar with tables of numerical data than other graphs, with the exception of diagrams (see section 4.3.2). In relation to diagrams, however, the use of a simple and linear diagram reporting the sequence of the events leading to the broadcast of a TV documentary in Task B:Broadcast did not necessarily lead to the lowest performance; with a mean score of 6.00, it was higher than the lowest – Task A:Oil (5.93). The substantial accompanying text in Task B:Broadcast might have mitigated the innate high cognitive demands of diagrams.

In addition to the participants' average performance data which provided some insights into understanding the effects of different graphic prompts on the cognitive processes of completing AWT1 tasks, we investigated which words were used frequently in each task as another means of examining the effects of graphic prompts on the participants' lexical decision-making. Are the words used in the writings predictable? We ran word frequency using Wordsmith Tools (the full word list is available upon request). Appendix 9 reports the frequently used words specific to each task. A clear pattern is observed in the content words used for each task. The most frequently used words (ie, excluding non-content words such as "the", "have" and "has") for each task matched very well, as anticipated, with the theme or topic of graphs. In other words, the graphs determine directly and probably naturally the words test takers use; the content of the writing is largely predictable. As Participant W commented:

Extract 4

> "Different graphs would require the use of different words and you may have different level of knowledge of different words… As I said, different tasks would stimulate your vocabulary knowledge, for example, the line graph activates your vocabulary such as increase and decrease." (Participant W)

At the same time, we also noticed that certain words are used frequently anyway, regardless of the topics of the graphs, for example, "difference", "more", "decrease", "increase", and their variations. The use of these four words demonstrates that the participants were constantly making comparisons as required by the task instructions. Another word that also appeared frequently in the writings is "show". Two factors may explain the high frequency of this particular word: firstly because it appeared in all the task instructions and secondly because this is such a neutral word that it can be applied to all tasks, for example, "This diagram shows…", "This table shows…", "This line graph shows…". However, some participants were looking for synonyms of "show", e.g. "reveal", "demonstrate", "display" and so on. Therefore, content-wise, the use of different graphic prompts apparently affected which words were likely to be used and consequently tested; however, process-wise, the participants were doing the similar activity – constantly trying to make comparisons in all the tasks except B:Broadcast where none of the four comparison words were used with any frequency.

The primarily product-oriented analyses above on the average test performance and the lexical features of the writings of each task provided a useful perspective to examine the potential differential effects of different types of graphs on the cognitive processes. Next we examine the think-aloud protocols and interviews to explore further the effects of different graphic prompts on the cognitive processes of completing AWT1 tasks. We report below some of the prominent features of graphs (eg, the display characteristics and conventions of graphs, the prior background knowledge required to understand the graphs) that the participants thought made the tasks particularly challenging or easy for them. It emerged that three types of graphic prompts received the most comments or attention from the participants: Task A: Instruction using stacked bar graph, Task A:Oil using statistical table, and tasks using line graphs.

Task A:Instruction used a stacked bar graph. The participants found it particularly "difficult to find out the length of the second and the third bars, as they are not from the same starting point" (Participant N). The density of the data points of this graph caused further problems as Participant N commented:

Extract 5

> "There are so many bars. It would be better if there were only 5 bars, for example, and it would be ideal if within every part of a bar is indicated with its percentage of the hours of a particular country." (Participant N)

Indeed several participants tried to work out the length of each section of a bar so that they could get a better sense of the number of instruction hours for different stage of schooling. For example, Participant T tried to estimate the percentage of each part of a bar. To her, the difficulty of describing this stacked graph was not just "because there are too many bars"; the main challenge of this graph also came from the fact that "the differences are not easily observable". Similarly, Participant K compared this stacked graph with the line graph in Task B:Viewing, and said:

Extract 6

> "Not like the task about TV individuals viewing share, you can see the change there. It is impossible that I would measure the differences by using some rulers." (Participant K)

Or, indeed as Participant W said:

Extract 7

> "because the trend of all the countries is all the same, only slight difference between the countries, but these differences are not obvious to bare eyes. It does not tell us the exact number of hours." (Participant W)

When comparing Task A: Oil which used statistical table with other tasks using line or bar graphs, the comments made by Participants H and K are typical. They clearly demonstrated the easier accessibility of the visual displays of line and bar graphs, while extra efforts were required to make trend assessments out of information from statistical tables: more calculation and judgements were required to make trend assessment possible.

Extract 8

> "I think the pie or bar graphs are more straightforward, while the statistics tables contain a lot more information. In my study of physics, we often have to process a lot of data from experiment; we can take a couple of hours to interpret data. At first glance of the statistics table, I was a bit lost, not sure what to do with so much information." (Participant H)

Extract 9

> "The line graph task is relatively easier, because I don't have to work out the information, I only need to follow the lines, the trends; I just need to describe what is there, there seems to be set procedure, you just follow the steps, one by one, this seems to be a right template you should follow. But for this type of tasks (Referring to the task with tables), I need to select the key statistics, it is therefore demanding. At first glance of the table, I tried to compare which was bigger or smaller, and also tended to find out the changes, for example, I did some calculations." (Participant K)

The following eloquent arguments, from Participants R and T, on the effects of the different graphic prompts on the processes and the products of AWT1 tasks summarize nicely the views expressed by all the participants on such effects.

Extract 10

> "If it is a line graph, I would first of all look for the changes of patterns, the trends of the lines. If it is a bar graph, I would compare the ones that are in the same category, to compare which number is bigger or smaller, higher or lower. If it is a table with statistics, it won't be as straightforward as the line or bar graphs. When you see a table, you will first of all get a sense of the numbers in the table. However, from the graphs, you don't have a straightforward sense of numbers, but graphs give you a clearer overall picture of the data." (Participant R)
>
> "For example, bar graphs, you would first of all talk about the x and y axis, and they also have some statistics, you would put them into different categories according to their number, e.g. these are below 50, and those are above." (Participant R)
>
> "If it were a table, like this one, I would do in the same way. For example, the USA is high, and China is also high. Both have different colours, and the other countries listed however have almost the same kind of colour, they all look green. I put them into three categories, the highest, the lowest and those in the middle." (Participant R)

Extract 11

> "Each type of graph has its own pattern or convention. For example, line graphs would demonstrate trends of development, for example. Bar chart normally compares the amount of two or three things. When you know the underlying convention and meaning, it is less likely that we may deviate the focus or the main message of the graphs." (Participant T)

In summary, the types and conventions of graphic prompts did matter. They affected how the participants processed the graphic information and how they followed the graphic conventions to re-produce their comprehension in written discourse in English. Such effects of different AWT1 graphic prompts on the cognitive processes were clearly manifested in the mean scores, in the use of vocabulary, and in whether and how they would make comparisons or trend assessments, following the graphic conventions in presentation, interpretation and re-presentation.

4.3.2 Research question 2: To what extent are the candidates' cognitive processes affected by their graphicacy?

In Section 4.3.1, we discussed the effects of different graphic prompts on the cognitive processes and performance of AWT1 tasks, and noted that there might be some dubious interactions between types of graphs and the participants' familiarity with different types and conventions of graphs, as some students may be more familiar with one type of graph than another. In this section, we will focus specifically on how the participants' graphicacy level might affect their cognitive processes and performance, from two perspectives – the participants' views as reflected in the questionnaire and interview data, and their actual writing performance.

In the questionnaire (Appendix 3) we asked the participants about their familiarity with different types of graphs: bar, line, pie, diagram, table with numerical data respectively (Q21-25). The Friedman test indicated that there was a significant difference in students' familiarity among these five different types of graphs (Chi-square=11.33, df=4, $p<.02$). Further analyses showed that the difference was mainly attributable to students' higher level of familiarity with line graphs than other types of graphs. Their familiarity with the different types of graphs is in the order of line, pie, bar, table with numerical data, and diagram. In other words, they were most familiar with line graphs, and least familiar with diagrams. This finding was in line with the participants' comments on the effects of different types and conventions of graphs on their cognitive processes and performance (see Section 4.3.1). In the questionnaire we also asked the participants if they think they "may do better in IELTS Academic Writing Task One using familiar graphs than unfamiliar ones" (Question 42) and if they "would prefer

one type of graph to be used in IELTS Academic Writing Task One" (Question 43). As anticipated, the majority of the participants thought they would do better in tasks using familiar graphs than unfamiliar ones; 95.8% chose 4 to 6 ("strongly agree"). However, the specific ways their familiarity with different graphs influenced their actual AWT1 writing performance may be another matter. Indeed, as shown in their response to Question 43, their belief in the effects of graph familiarity on the processes and performances of AWT1 tasks seemed to be weaker, as fewer students (79.1%) had chosen 4 to 6 ("strongly agree"). As for their belief on whether special training in interpreting graphs would be helpful for getting a higher score in their AWT1 writing, 91.7% had chosen 4 to 6 ("strongly agree"). However, like the perceived effects of graph familiarity on test performance, the actual usefulness and effects of special training on test performance may be another matter.

At the interviews, we specifically asked the participants to give their opinions about the ways their graph familiarity affected how they processed and wrote about the graphs. The interview data confirmed that these students were more familiar with line graphs than other types. For example, Participant C commented:

Extract 12

> "As for the other types of graphs, e.g. line graph, probably because we use these types of graphs often, whether in study or other activities, therefore we are more familiar with these graphs, and know better how to write about these graphs." (Participant C)

In addition, it was noticed that graph familiarity was considered helpful, at least, psychologically, to boost the participants' confidence that they were doing something they were familiar with.

Extract 13

> "When you are given a task with a familiar graph, you will feel more confident. When you are given a task with an unfamiliar graph, you may feel less secure when writing." (Participant I)

Extract 14

> "I like to read the line graphs; it is probably because I am familiar with this type of graphs." (Participant J)

Extract 15

> "Line, bar and pie chart graphs are more familiar to me; I was a bit surprised when I was given the map task." (Participant S)

Beside the potential psychological boost, the familiarity with graphs was considered to be helpful for guiding the writing process too, as the participants would be able to follow a kind of writing convention or template that they already know or are comfortable with. Unfamiliarity however could probably create a kind of anxiety and insecurity (see also Participant I's comment above).

Extract 16

> "For example, the map task, as I haven't come across this type of graph, I found it difficult. For familiar tasks, you would know how to write the beginning paragraph, the body paragraph and the conclusion paragraph. For unfamiliar graphs, you don't know how to start, how to develop your writing and how to end your writing. You don't know where the entry point for this kind of unfamiliar tasks is. … [Researcher: How would you start then?] If I had not come across a particular type of graph, I would not have a model to follow, and then I would have to follow my instinct to write…" (Participant R)

Furthermore unfamiliarity may also present test takers with some real challenges in processing and writing about the graphic information, particularly when they are trying to integrate their prior knowledge with the graphic information presented. For example, Participant W commented that:

Extract 17

> "Some graphs (e.g. the world map) may also present challenges, especially as I may be lack of some knowledge of geography and history. I was thinking if the different geographical and economic positions of these countries affect their amount of CO2 emissions. As I'm not familiar with the geographical positions of these countries in the world map, this could be a challenge to me. [Researcher: If you had not known Alaska belongs to America, would you think it is another country in the world map?] Yes. In fact, I did ask X [Participant R] yesterday what country is this one (referring to Alaska). To me, when I received the tasks with the map, I was a bit nervous. But apart from that, it was OK." (Participant W)

Participant W further explained her belief that test takers may have different levels of graph familiarity, but it is their writing ability that matters the most in the AWT1 tasks.

Extract 18

> "It is inevitable that some students may be highly familiar with certain graphs, although his or her writing may not be strong. It is possible that this student may do well because he or she is highly familiar with the graph; this kind of situation is possible, and I also think this kind of situation is inevitable. However, generally speaking, his or her writing is a positive reflection of his or her writing ability." (Participant W)

After all, these largely positive attitudes towards the potential impact of graph familiarity on test performance may be due to the fact these participants had high graph familiarity (see Figure 3). Indeed, as Participant T acknowledged: "I don't think there are many graphs which are unfamiliar to us. In the textbooks we study there are graphs". To what extent will this view on the potential relationship between graph familiarity and test performance be evidenced in the actual test performance?

Next we ran a series of regression analyses on graph familiarity and performance of each AWT1 task (including the first one without think-aloud, see Appendix 2), the average performance of the first four tasks, the average performance of the second four tasks, and the average performance of all eight tasks under think aloud conditions. No significant correlation was found between graph familiarity and test performance. We also ran ANOVA to test if each AWT1 task was affected by the participants' familiarity with the particular type of graph used in that task. For example, we ran one-way ANOVA to examine if there was a significant difference in performance in Task A:UKCO2 between participants of different familiarity with line graphs (i.e. factor = question No. 22, see Appendix 3 graphicacy questionnaire), a difference in performance in Task B:Viewing between participants of different familiarity with line graphs (i.e. factor = question No. 22, see Appendix 3), and a difference in performance in Task A:Oil between participants of different familiarity with statistical tables (i.e. factor = question No. 25, see Appendix 3). No significant differences were found either.

In summary, it seems that graph familiarity did not affect these participants' AWT1 task performance in terms of the marks that their writings received, although some potential psychological impact of graph familiarity on task performance was expressed clearly by the participants. This may be largely due to the fact that these participants had a high level of graph familiarity. However, as shown in Section 4.3.1, the participants' understanding and knowledge of the conventions of different types of graphs may affect the ways that they processed and wrote about the graphs. In other words, graph familiarity should be examined from two perspectives: one is the quantified graph familiarity as measured by the graphicacy questionnaire (Appendix 3), as discussed in this section, and the other

is their familiarity with and understanding of the conventions of different types of graphs (see Section 4.3.1).

4.3.3 Research question 3: To what extent are the candidates' cognitive processes related to their writing abilities?

As described in the working model of cognitive processes (Figure 3), test takers' writing abilities played a role when they were re-producing their comprehension of the graphs in written discourse in English as a foreign language. As lexical knowledge is an important aspect of a test taker's writing ability, the use of different vocabulary is clearly an essential indication of the relationship between the cognitive processes of completing the AWT1 tasks and the test taker's writing ability. As reported in section 4.3.1, different graphs activated the use of different vocabulary, in other words, different graphs had different lexical demands (see Appendix 9 and Extract 4). In this section, we report further on (a) the relationships in performance between different AWT1 tasks and between AWT1 and AWT2 tasks, and on (b) how test takers' English writing abilities, their expectations and experiences of academic writing, whether in English or Chinese, might have shaped the way that the comprehended graphic information from AWT1 tasks was re-produced in written discourse in English as a foreign language.

Table 4 (see section 4.1.2) reported the participants' performance in AWT1 and AWT2 tasks in normal examination conditions, that is, without thinking aloud (Appendix 2). It is interesting to note that the correlation between the AWT1 and the AWT2 performances was not significant ($r=0.33$, n.s.). The difference in performance between the AWT1 and AWT2 tasks was statistically significant ($t=-3.73$, $p<0.005$). On the one hand, this insignificant correlation between AWT1 and AWT2 performances raises the question about the extent to which these two writing tasks share an underlying construct; and on the other hand, it demonstrates the necessity of using not only AWT1 but also AWT2 tasks to measure candidates' academic writing abilities (as in IELTS), and also the necessity of using both test scores to analyse the relationships between the so-called English writing abilities and the cognitive processes of doing AWT1 tasks.

The correlation between the participants' performance of the AWT1 task under normal examination conditions and their average performance of the eight AWT1 tasks under think-aloud conditions was statistically significant ($r=0.53$, $p<0.05$, $n=17$), which indicates that the overall effects of think-aloud as a data collection procedure on test performance may be minimal. However, it should be noted that the correlation between the AWT1 task under normal examination conditions and the first four AWT1 tasks under think-aloud conditions before the special training was not significant ($r=0.38$, n.s.); the overall significant correlation is therefore largely attributable to the more significant correlation between the second four AWT1 tasks and the AWT1 task completed under normal examination conditions ($r=0.60$, $p<0.05$). The correlation between the first four and the second four tasks is more significant ($r=0.74$, $p<0.01$). In other words, when the tasks were all under think-aloud conditions, the correlation between AWT1 tasks is stronger than if one was under think-aloud conditions and the other was not. This stronger correlation in performance between the AWT1 without think aloud and the second 4 tasks under think-aloud conditions, and the better performance in the second than the first 4 AWT1 tasks under think-aloud conditions (mean difference $=0.24$, $t=2.9$, $df=16$, $p<0.05$) might be attributable to three factors. Firstly, the participants might have got used to think-aloud towards the end of data collection. Secondly, the training before the administration of the second set of tasks (see Stage 3 in section 3.3) might have helped not only to mitigate the effects of think-aloud on test performance but also improve test performance. Thirdly, think-aloud itself might have helped the participants to concentrate on their tasks and improve their understanding of the graphic information and hence their test performance. Therefore we would argue that it is imprudent to ignore completely the potential effects on the writing performance of think-aloud through which the cognitive processes of these participants were elicited (see section 4.3.4 about the participants' comments on the use of think-aloud).

As indicated earlier, the correlation between the AWT1 and the AWT2 tasks under normal examination conditions (ie, without think aloud) was not statistically significant (r=0.33, n.s.). In other words, AWT1 and AWT2 may not share the same construct. We therefore also used AWT2 performance as another indicator of the participants' writing abilities to predict their performance in AWT1 tasks under think-aloud conditions. It was found that the correlations between AWT2 and the first set of four AWT1 tasks ($r=0.59$, $p<0.01$), between AWT2 and the second set of four AWT1 tasks ($r=0.78$, $p<0.001$), and between AWT2 and the mean score of all eight tasks under think aloud conditions ($r=0.74$, $p<0.001$) were all statistically significant. However, it is puzzling that the participants' performances of the two AWT1 and AWT2 tasks under normal examination conditions were not significantly correlated. We speculate that the score of the AWT1 task using only one type of graph (here a line graph) was less capable, than the mean score of the eight AWT1 tasks using different graphic prompts, of measuring these participants' AWT1 writing abilities. This speculation is in line with our findings regarding the effects of different graphs on the cognitive processes (see 4.3.1), because different types of graphs may activate different lexical demands and therefore involve different English writing abilities. Below we report how test takers' English writing abilities (in terms of grammar and word choice), their expectations and experiences of academic writing might have shaped the way they comprehended the graphic information and re-produced their comprehension in written discourse.

As demonstrated in the think aloud protocols, the participants were constantly monitoring their grammar and word choice. For example, when Participant S was considering which tense to use to describe the sequence of the events leading to the broadcast of a documentary (i.e. Task B:Broadcast), she said: "To complete their things it always take tense will…, information, just use the past tense". Similarly Participant G was also constantly monitoring her use of grammar, whether in present or past tense.

Extract 19

> "Domestic use is less than that, so, how should I say about domestic carbon emission, domestic use is, much, is much less than, no, was, was, was much less than that for industry…" (Participant G)

Word choice is another area that the participants often had to make decisions when completing AWT1 tasks. For example, Participant K must have mis-read the word "disused" as "discussed" in Task A:Layout, he then wrote: "In the South the lake, a railway line which was under discussing in 1937 has already been built up", which of course led to a completely different interpretation from what was intended. Other examples of the participants' decision-makings on word choice are less dramatic. For example, Participant L tried to figure out whether "proportion" would be a more appropriate word than "amount" when describing the biggest source of UK CO_2 emissions in 1970; whether to use "sources", "uses" or "approaches", and whether to use "steadily", "steady", "stead", or "stable" when describing the trend of CO_2 emissions. Finally she decided to use "proportion", "approaches" and "steady increase".

Extract 20

> "Industry in 1970 accounted for the largest amount, no, accounted for the largest proportion in different end users, the largest proportion in the UK CO2 emissions.
>
> Other, UK emitted about 30 million carbon in other sources, in other uses, in other approaches.
>
> Transport, however, transportation, transportation, transportation emitted more and more and more during the 34 years steady, more and more showing a steadily, showing a trend of stead, steadily, stable, steady increase in 2004." (Participant L)

However, when the participants were much less certain about the meanings of unfamiliar words, this can be a long and struggling process. Often their final decision was to "just copy" if these happened to be the words in the accompanying texts of the graphs. For example, Participant G was not quite sure "what is the meaning of million tonnes carbon equivalent". Although she adopted the strategy of "I just copy this", later on, she was still repeatedly asking herself "what carbon equivalent mean". Similarly she kept asking herself "what's the meaning of Year to Date, year to date, year to date, I don't know what is year to date". A "just copy" strategy was also used to "solve" this problem. YTD is clearly written in her script.

When asked about the contribution of their writing ability on the AWT1 task performance, all the participants agreed that it is the writing ability, in particular, lexical knowledge, that matters most. The following comment is typical:

Extract 21

> "This is definitely so, for example, your knowledge of certain words and some sentence structures and so on, these are essential for the completion of the AWT1 tasks. For example, if you only know the words such as decline, increase, but you don't have other more powerful words, your writing will become a bit boring and plain." (Participant J)

Extract 22

> "I think the most challenging inhibitors may be that I could not find the right words straightaway, or even if I can remember of a word vaguely, but I may not be able to spell it correctly, then I will have to use a simpler word instead to be safe." (Participant S)

Another aspect of writing, although not necessarily synonymous to English writing ability proper, the participants' experience and expectation of academic writing, whether in English or Chinese, did seem to affect the way how they interpreted the data and whether they would think it essential to include interpretations in AWT1 writings. Their comments on the purpose of describing graphs (see Figure 3, a key component of the working model: Purposes of understanding the graphs and the cognitive demands of the writing task) are thought provoking and raise some fundamental issues of IELTS AWT1 tasks – what are the purposes of the writing, and how are these communicated to and interpreted by test takers?

Extract 23

> "Usually when we write and describe data, we have a purpose, have an aim, why do we do this, but these tasks, we don't have a clear purpose. I don't know where to start and how to start to describe the information in the graphs, so I feel these tasks do not have a clear purpose, so I feel I can't have a clear idea or the logics of understanding what information is more important and how to summarise and organize them. I don't know what the key messages to summarise are." (Participant L)

Her comment was fully supported by other participants. The following conversation between them on the importance of including interpretations as a natural part of academic writing is interesting.

Extract 24

> "Participant E: I still think we should write interpretations because these are for academic tasks.
>
> Participant U: It is always like that in academic writing that you present the data and explain and interpret why, the reasons why there is such data.

> Participant E: Yes, like that. We should write like that. In fact, some of the facts, we don't need to present them, while the interpretation is important.
>
> Participant L: Yes, this is the sort of structure of academic papers, you first present data and then interpret the data. The purpose of presenting the data is to serve the interpretation and discussions. This is the pattern. It is always like that for academic writing. … Why bother to present data alone?"

In fact, all the participants in this project insisted that they should be allowed to include their interpretation in the AWT1 writings as this is something expected in academic writing. This explains why there were a lot of personal interpretations and comments in the writings of the first four tasks before the special training session which emphasized that they should "describe" rather than "interpret" data.

In Section 4.3.4 below we report how the short test preparation training changed the ways that the participants described or interpreted the main features of graphs.

4.3.4 Research question 4: To what extent are the candidates' cognitive processes influenced by test preparation opportunities offered to them by the research team?

In 4.2.1 we reported that the participants were constantly reminding themselves the second sentence of the task instruction "summarise the information by selecting and reporting the main features, and make comparisons where relevant" to make sure that they keep focusing on the task, rather than being carried away too much by personal interpretations, predictions and comments based on prior knowledge which was not necessarily depicted in the graphs. We view this as the most discernable influence of the test preparation training on the cognitive processes. As presented in Appendix 10, we can see that the participants in the first round of AWT1 tasks made extensive interpretations, predictions and comments. Some were more reasonable than others, while some were simply wrong or irrelevant. In some scripts, more than half of the space were devoted to explaining the potential underlying reasons why a particular phenomenon existed and what actions should be taken. In the second round of AWT1 tasks, ie, after the short test preparation training, the participants tended to refrain from making too many personal interpretations and comments. Although there are still interpretations which were not based primarily on the data of the graphs, it is obvious that the interpretations became more restrained, reasonable and reduced. There are fewer personal interpretations, comments and generalizations in the texts of Tasks B answers than of Tasks A. Overall, the participants had better understanding about the task requirements after the training. For example, Participant B acknowledged in the interview:

Extract 25

> "At the beginning, I did not really understand the task requirements, but after the Wednesday training session, I have a better understanding of the tasks. … For example, as for bar graph, there may be many bars in the graph, and a lot of information, but there is a limit of time, and your choice of words is also limited, you will have to select the most important information, not necessarily to include every bit of information. I think it is particularly useful for me to know this. …. Sometimes we read graphs when doing literature review. And quite often I try to use as much as possible to describe graphs, and try to explore and extract as much information as possible from graphs. As for AWT1 tasks, the first impression I had was I should try my best to interpret what's in the graphs. For example this graph, I try to find out why this is low and why that is high, trying to solve this problem. But I can get it wrong, and it would lead me to a wrong route." (Participant B)

Similarly, Participant U agreed that "After training, it is unlikely that I will still do too much interpretation of the data as I did in the first set of tasks". Participants E and W also held the same view:

Extract 26

> "When I write, e.g. about the line graph, I always think about why there is such a change in the lines, why, this is the question I always ask myself, and add my own understanding/interpretation in my writing. Only when you pointed out in the tutorial that it is not necessary to include my own interpretations of the data, the main thing is to describe the information in the graph, that's it, did I realize that I was not right to always try to interpret the data to find the underlying reasons for the changes in a graph, for example, the lines." (Participant E)

Extract 27

> "Before the Wednesday training, when I wrote the tasks, I tried to explain the reasons, the underlying reasons. For example, there are more and more population, and there are more houses, and the man-made facilities were becoming more and more useful for human beings and the environmental impacts of these are also getting bigger…. If it is your subject specialist area, it is quite natural that you have a tendency to reason and argue this way." (Participant W)

This short test preparation training also seemed to clarify a confusion that the participants had about a similar type of writing tasks in the College English Test in China which in fact required test takers to explain why a particular phenomenon depicted in a graph happened. The explanations by Participants W and Q on their previous experiences shed light on why these students always had a strong tendency to interpret the data when doing IELTS AWT1 tasks.

Extract 28

> "In AWT1 tasks, no matter what kind of graphs or diagrams, you are only asked to describe. However, in the Chinese tests, you are given a picture or a graph, you are not only asked to describe some kind of information at surface level, however, the most important part is to find out and write about the intended message and meaning. However, in AWT1 tasks, you are only asked to describe…. in CET and Gaokao (Note: Chinese university entrance examination) composition writing, you only use 20-30 words max to describe the graphs or pictures, but the following analyses are more important." (Participant W)

Extract 29

> "In essays, normally we present and describe a problem briefly and what follows is more important, you interpret and give some arguments and provides solutions to the problem. It is normally this kind of pattern for argumentative essays. This is the kind of training we have received since secondary education." (Participant Q)

However useful the short test preparation training may have been to help the participants focus on describing rather than interpreting the graphs, we still noticed in think-aloud protocols that interpretations still feature, maybe as a natural process of understanding the graphs. As Participant T commented, although interpretations may not be totally avoidable, she would not include the interpretations in her actual writings.

Extract 30

> "Before the training, I always had the desire to interpret the reasons, e.g. why is there an increase, why is there a decrease? After the training, I had to curb my desire of interpreting, but I still quite naturally, very naturally, or unconsciously try to interpret, though I will not include my interpretations in my writings, as I understand this is not what the task requires, this is not what the markers would be looking for, I have to refrain from writing too much, it has to be within 150 words. I don't need to bother anything not required." (Participant T)

Data were also collected from the participants about their views on the research project and the short training (Appendix 6 Student evaluation questionnaire). Table 6 reports the descriptive statistics of their evaluation on the overall training, learning support, handout, content, teaching quality, and their learning from the training and their own contribution. As can be seen from Table 6, the participants overall had very positive comments on the research project and the training provided.

Score	*Overall training*		*Learning support*		*Handouts*		*Content*		*Teaching quality*		*Learning from training*		*Own contribution*	
	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%
1														
2													1	7.1
3			1	7.1							3	21.4		
4	3	21.4	3	21.4	1	7.1	5	35.7	2	14.3	4	28.6	5	35.7
5	8	57.1	7	50.0	4	28.6	6	42.9	9	64.3	6	42.9	4	28.6
6	3	21.4	3	21.4	9	64.3	3	21.4	3	21.4	1	7.1	4	28.6
Mean	5.0		4.9		5.6		4.9		5.1		4.4		4.7	
STD	0.68		0.86		0.65		0.77		0.62		0.93		1.14	

Table 6: Student evaluation of the project (n=14)

We received very positive comments from the participants regarding the quality of the training and their genuine interest and commitment in participating in the project. The qualitative comments (see Appendix 11) provided further empirical evidence on the effects of the short training on the cognitive processes of completing AWT1 tasks. In particular, the participants noted the benefits of learning special vocabulary for making comparisons and trend assessments (the training booklet is available from the authors upon request). Analyses of the writing performances by the same students on different graphs indicate that there is some re-occurrence of formulaic phrases and verbs; and this is particularly so in the four tasks completed after the short training. This clearly demonstrates the coachability of the AWT1 tasks. However, this coachability may not on its own lead to a high level performance in such tasks, there is a possibility that test takers may produce some formulaic and rigid phrases without necessarily involving true understanding of the phrases or the graphic information.

5 CONCLUSION

This research aimed to understand the cognitive processes of candidates taking IELTS AWT1 with different graphic prompts at two different time points – before short training on how to do AWT1 tasks and post-training. In particular, it explored the extent to which candidates' cognitive processes were affected by the use of different graphs, their "graphicacy" (Wainer 1992, p 16) and English writing abilities, and the short training. A grounded and multi-layered case study approach was employed to collect data on candidates' cognitive processes. Eighteen intending IELTS candidates were recruited from a large Chinese university, and completed eight AWT1 tasks while thinking aloud their processes of doing the tasks (four before training and four after training) under examination conditions. In addition, their English writing abilities and graphicacy were also collected, and post-task interviews were conducted with all the participants.

The think-aloud protocols were analysed using a qualitative computer programme winMAX to identify the common patterns in the cognitive processes. The patterns which emerged from the think-aloud protocols were interpreted with reference to the other four main sources of data – AWT1 scripts, the interviews, graphicacy questionnaire and English writing abilities. From these data sources, a model of cognitive process was developed, consisting of three interrelated stages: comprehending non-graphically presented task instructions, comprehending graphs, and re-producing graph comprehension in written discourse in English as a foreign language. We used this model to guide our analyses to address the four research questions. Below we summarise the main findings of each research question.

1 With regard to the *effects of types of graphs on cognitive processes*, it was found that the types and conventions of graphic prompts did matter. They affected how the participants processed the graphic information and how they followed the graphic conventions to re-produce their graph comprehension in written discourse in English as a foreign language. Such effects of different AWT1 graphic prompts on cognitive processes were clearly evidenced in the mean scores of the writing performances, in the use of vocabulary, and in whether and how they would make comparisons or trend assessments. The graphic conventions or "cognitive naturalness" of graphs (Tversky 1995) affected the processes of comprehending and re-producing graphic information. Candidates had a strong tendency to make trend assessments when describing line graphs, and make discrete comparisons when describing bar and pie charts. When describing a statistical table – the most challenging of the AWT1 tasks in this study, candidates tended to do some calculations of the numbers to develop their reasoning. Different types of graphs also activated the use of different vocabulary types.

2 With regard to the *effects of graphicacy on cognitive processes*, we found that: although graph familiarity as measured via the graphicacy questionnaire did not seem to affect AWT1 task performance in terms of the scores for the writing performances, the participants clearly expressed some potential psychological impact of graph familiarity on their task performance. In addition, the effects of the participants' familiarity with and understanding of graphic conventions also influenced the way they processed and re-produced the graphic information (see 1 above).

3 The *effects of the writing abilities on cognitive processes* were manifested in the use of different vocabulary choices for different graphic prompts. There is also a strong correlation between the mean performances of the AWT1 tasks under think-aloud conditions and the AWT2 task of topic-based argumentative writing. Besides the participants' English writing abilities, their expectations and experiences of academic

writing also shaped the way that they interpreted and re-produced their graph comprehension in written discourse in English as a foreign language.

4 The *influence of the special training* was strong; this clearly demonstrated the coachability of the AWT1 tasks. Whether or not the candidates tried to make personal interpretations and comments by linking the graphic information and their domain knowledge about the graphs were clearly influenced by the short test preparation training they received, although they did not necessarily agree with the AWT1 task requirements on "describe".

As we discussed in section 1.1, this research project addressed two broad areas of interest identified by the IELTS Joint Research Committee – (a) "test development and validation issues" in relation to "the cognitive processes of IELTS test takers", and (b) "issues in test impact" in relation to "test preparation practice". The findings of this study have implications for the validation and development of AWT1 tasks from the perspective of test takers' cognitive processes. The working model of cognitive processes (Figure 3) will be a useful framework for designing AWT1 tasks, considering the three interrelated stages of AWT1 cognitive processes – comprehending the non-graphically presented task instructions, comprehending graphic information, and re-producing graph comprehension in written discourse in English. The findings of this research project suggest that when designing AWT1 tasks, we need to consider what accompanying instructions should be provided with the graphs and whether the summary-like introductory sentence should serve as the entry point for levelling the playing field for test takers with varying background knowledge, graphic skills and reading abilities. As AWT1 tasks are essentially reading/writing integrated tasks, it is essential to consider not only the graphic skills of the test takers but also their reading abilities. Although the tasks may require just reading one summary-like sentence in the instructions, the sentence can actually serve as a guide for test takers to grasp the key information of the graphs, and it may become more and more critical when test takers move further down the line to re-produce the main information embedded in the graphic prompts.

It is also important to consider and compare the potential differential effects of different graphic prompts on the cognitive processes of test taking as well as on test performance, because different graphs have different conventions in presentation and interpretation. Beside the types of graphs, we also need to consider the information density, or data points of the graphs, as information density clearly has impact on how and what test takers can extract from the graphs. Other display characteristics of graphs also need to be taken into account when test designers consider the information density of the graphs, eg whether to include caption or title. As display characteristics of graphs can affect task performance, we would be very cautious about O'Loughlin and Wigglesworth's (2003) recommendation: "a variety of presentation types can be encouraged and manipulated" (p 114). Although the final scores that two pieces of writings may receive may be similar, regardless of which graph type they are based on, it is very clear, as shown in our study, that the use of different graphic prompts can activate different forms of "cognitive naturalness" in comprehending and re-presenting graphic information, and can invite candidates to produce different writings, particularly in terms of the use of different vocabulary choices. The quality of a graph is not just innate in its display features (Bertin 1983; Saint-Martin 1990; Tufte 1983); it is in fact determined in interaction between these features and the characteristics of graph readers (eg, their familiarity with and experience in using graphs). These interactions are the starting points for investigating the cognitive processes of AWT1 tasks.

We noticed that AWT1 tasks seemed to be highly trainable and the writings were highly predictable: the participants were easily influenced in their ways of re-producing graph comprehension by the short training. The high coachability of the tasks and the rigid use of the formulaic phrases and words in AWT1 writings without necessarily understanding them bring up the ongoing question of how far AWT1 fosters rote-learning of fixed phrases. They also raise an equity issue. If test takers have not

attended intensive training courses and are not aware of the task requirement of "describing" only, they could be disadvantaged if they tried to integrate the information depicted in the graphs with their prior content and graphic knowledge. Whether and to what extent reasonable "interpretation" should be allowed in AWT1 writings have implications for comparability in marking. The participants' strong view that interpretation of data is an integral and natural part of academic writing also oblige us to consider the extent to which the AWT1 tasks mirror academic writing in target language use domains, for example, when describing and interpreting lab results.

The findings of this research study are useful not only to language test providers and language testing researchers, but also to intending IELTS candidates and English language professionals and teachers to develop a greater understanding of the AWT1 tasks, as well as other language tests (listening, speaking and writing assessments) using graphs as prompts. It contributes to the development of theories and practices in foreign/second language writing, in particular, in relation to our understanding of the roles that non-language knowledge and skills (ie, graphicacy) can play in language assessments using visual input or data. In order to further develop our understandings of tasks using graphic prompts, we acknowledge several limitations of this research study and present two pointers for future studies.

Firstly, although sufficient training for think-aloud was provided to the participants in this research project, the effects of think aloud on test performance may never be removed completely, it is important to examine in detail the extent to which think-aloud may have affected the participants' cognitive processes and their actual writing performances. Detailed discourse analyses on the changes between think-aloud protocols and what the participants finally put down in their writings would be a useful perspective to analyse the effects of think aloud on composition and more importantly on the cognitive processes of re-producing graph comprehension. Investigation of content coverage would be another useful perspective to understand the cognitive processes of AWT1 tasks, for example, what is included and what is not included, how and why, and the extent to which the inclusion or exclusion is related to the test takers' English writing abilities and graph familiarity.

Secondly, our analyses on the participants' comprehension of graphic information did not focus specifically on the sequence that the different graphic components were noticed and comprehended. Instead we focused on the participants' overall comprehension of graphs. As Kennedy (1974) argued, "sometimes we read a label or caption before looking at the picture, but more often, probably, we notice the picture first and recognise the pictured object without any help from the accompanying words" (p 7). Although our data indicated that the participants read or skipped the accompanying textual instructions before focusing on the graphs, it would be useful to systematically analyse which graphic components were noticed and comprehended first, and whether the comprehension of different graphic components was affected by the participants' graphic skills and types of graphs. In addition, it would be useful to investigate more precisely the timeline of AWT1 task completion, for example, the time spent in reading and comprehending the graphs, time spend in re-producing the graphic information in writing, checking and re-checking understandings, and re-writing. Future studies may examine the sequence of comprehending graphic components, as the findings could have clear implications for AWT1 task design.

ACKNOWLEDGEMENTS

The authors acknowledge the role of the IELTS Partners in making this study possible: The British Council provided the research grant which enabled us to conduct this study. Any opinions, findings, conclusions or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the British Council, its related bodies or its partners.

The authors acknowledge the administrative support of Zhejiang University. Special thanks are also due to the student participants who wish to be acknowledged by name, including Hu Zhinan, Li Kangkang, Li Wenting, Liu Yuguo, Ni Shuang, Zhang Kejun, plus others who wish to remain anonymous. Their commitment and enthusiasm made this project possible.

REFERENCES

Åberg-Bengtsson, L, and Ottosson, T, 2006, 'What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement', *Journal of Research in Science Teaching*, vol 43, no 1, pp 43-62

Bertin, J, 1983, *Semiology of graphics: Diagrams, networks, maps*, University of Wisconsin Press, Madison, WI

Carpenter, PA, and Shah, P, 1998, 'A model of the perceptual and conceptual processes in graph comprehension' *Journal of Experimental Psychology: Applied*, vol 4, no 2, pp 75-100

Carswell, CM, Frankenberger, S, and Bernhard, D, 1991, 'Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data', *Behaviour and Information Technology*, vol 10, no 6, pp 459-474

Curcio, FR, 1987, 'Comprehension of mathematical relationships expressed in graphs', *Journal for Research in Mathematics Education*, vol 18, no 5, pp 382-393

Farquhar, AB, and Farquhar, H, 1891, *Economic and industrial delusions: A discourse of the case for protection*, Putnam, New York

Feeney, A, Adams, J, Webber, L, and Ewbank, M, 2004, 'Individual differences in graphical reasoning', *Lecture Notes in Computer Science*, 2980, pp 271-285

Feeney, A, Hola, A, Liversedge, S, Findlay, J, and Metcalf, R, 2000, 'How people extract information from graphs: Evidence from a sentence-graph verification paradigm' in *Theory and application of diagrams*, eds M Anderson, P Cheng & H Haarslev, pp 149-161

Feeney, A, & Webber, L, 2003, 'Analogical representation and graph comprehension' in *Smart graphics*, eds A Butz, A Krüger & P Olivier, pp 212-221

Fischer, MH, Dewulf, N, and Hill, RL, 2005, 'Designing bar graphs: Orientation matters', *Applied Cognitive Psychology*, vol 19, no 7, pp 953-962

Freedman, EG, and Shah, P 2002, 'Toward a model of knowledge-based graph comprehension' *Diagrams*, pp 18-30

Friel, SN , Curcio, FR, and Bright, GW. 2001, 'Making sense of graphs: Critical factors influencing comprehension and instructional implications' *Journal for Research in Mathematics Education*, vol 32, no 2, pp 124-158

Fry, E, 1981, 'Graphical literacy', *Journal of Reading*, vol 24, no 5, pp 389-392

Ginther, A, 2002, 'Context and content visuals and performance on listening comprehension stimuli', *Language Testing*, vol 19, no 2, pp 133-167

Golub-Smith, M, Reese, C, & Steinhaus, K, 1993, *Topic and topic type comparability on the test of written English*, Educational Testing Service, Princeton, NJ

Gruba, P 1997, 'The role of video media in listening assessment', *System*, vol 25, pp 335-346

Guthrie, JT , Weber, S , and Kimmerly, N 1993, 'Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations', *Contemporary Educational Psychology*, vol 18, no 2, pp 186-221

Hollands, JG , & Spence, I, 1998, 'Judging proportion with graphs: The summation model', *Applied Cognitive Psychology*, vol 12, no 2, pp 173-190

Hollands, JG , & Spence, I, 2001, 'The discrimination of graphical elements', *Applied Cognitive Psychology*, vol 15, no 4, pp 413-431

Katz, IR , Xi, X, , Kim, H -J , and Cheng, PCH, 2004, *Elicited speech from graph items on the test of spoken English*, ETS Research Report 74, Educational Testing Service, Princeton, NJ

Kennedy, JM, 1974, *Psychology of picture perception*, Jossey-Bass, San Francisco

Kintsch, W, 1988, 'The role of knowledge in discourse comprehension: A construction-integration model', *Psychological Review*, 95, no 2, pp 163-182

Körner, C, 2004, 'Sequential processing in comprehension of hierarchical graphs' *Applied Cognitive Psychology*, vol 18, no 4, pp 467-480

Kosslyn, SM, 1989, 'Understanding charts and graphs', *Applied Cognitive Psychology*, vol 3, no 3, pp 185-225

Kuckartz, U, 1998, *Winmax: Scientific text analysis for the social sciences* (Version 98 Pro) VERBI Software-Consult-Research GmbH (Germany), Berlin

Lewandowsky, S, and Spence, I, 1989, 'The perception of statistical graphs', *Sociological Methods Research*, vol 18, nos 2 and 3, pp 200-242

Lohse, GL, 1993, 'A cognitive model for understanding graphical perception', *Human-Computer Interaction*, vol 8, no 4, pp 353-388

Meyer, J, Shinar, D, and Leiser, D, 1997, 'Multiple factors that determine performance with tables and graphs', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol 39, no 2, pp 268-286

Mickan, P, and Slater, S, 2003, 'Text analysis and the assessment of academic writing' in *IELTS Research Reports Volume 4*, ed R Tulloh, IELTS Australia Pty Limited, Canberra, pp 59-88

Mickan, P, Slater, S, and Gibson, C, 2000, 'Study of response validity of the IELTS writing subtest' in *IELTS Research Reports Volume 3,* ed R Tulloh, IELTS Australia Pty Limited, Canberra, pp 29-48

Moore, PJ 1993, 'Metacognitive processing of diagrams, maps and graphs', *Learning and Instruction*, vol 3, no 3, pp 215-226

O'Loughlin, K, and Wigglesworth, G, 2003, 'Task design in IELTS academic writing task 1: The effect of quantity and manner of presentation of information on candidate writing' in *IELTS Research Reports Volume 4*, ed R Tulloh, IELTS Australia Pty Limited, Canberra, pp 89-130

Peebles, D, and Cheng, PCH, 2002, 'Extending task analytic models of graph-based reasoning: A cognitive model of problem solving with cartesian graphs in act-r/pm', *Cognitive Systems Research*, vol 3, no 1, pp 77-86

Peebles, D, and Cheng, PCH, 2003, 'Modeling the effect of task and graphical representation on response latency in a graph reading task', *Human Factors*, vol 45, no 1, pp 28-47

Pinker, S, 1990, 'A theory of graph comprehension', in *Artificial intelligence and the future of testing*, ed R Freedle, Lawrence Erlbaum Associates, Hillsdale, NJ, pp 73-126

Roth, WM, 2002, 'Reading graphs: Contributions to an integrative concept of literacy', *Journal of Curriculum Studies*, vol 34, no 1, pp 1-24

Roth, WM, 2003, *Toward an anthropology of graphing: Semiotic and activity-theoretic perspectives*, Kluwer Academic Publishers, Amsterdam

Saint-Martin, F, 1990, *Semiotics of visual language*, Indiana University Press, Bloomington, IN

Schnotz, W, Picard, E, and Hron, A, 1993,'How do successful and unsuccessful learners use texts and graphics?', *Learning and Instruction*, vol 3, no 3, pp 181-199

Shah, P, and Carpenter, P A, 1995, 'Conceptual limitations in comprehending line graphs', *Journal of Experimental Psychology: General*, vol 124, no 1, 43-61

Shah, P, Freedman, E G, and Vekiri, I, 2005, 'The comprehension of quantitative information in graphical displays' in *The cambridge handbook of visuospatial thinking*, eds P Shah and A Miyake, Cambridge University Press, New York, pp 426-476

Shah, P, and Hoeffner, J, 2002, 'Review of graph comprehension research: Implications for instruction' *Educational Psychology Review*, vol 14, no 1, pp 47-69

Shah, P, Mayer, RE, and Hegarty, M, 1999, 'Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension', *Journal of Educational Psychology*, vol 91, no 4, pp 690-702

Simkin, D, and Hastie, R, 1987, 'An information-processing analysis of graph perception', *Journal of the American Statistical Association*, vol 82, no 398, pp 454-465

Stewart, B, Hunter, A , and Best, L, 2008, 'The relationship between graph comprehension and spatial imagery: Support for an integrative theory of graph cognition', *Diagrams*, pp 415-418

Trickett, SB, and Trafton, JG 2004, 'Spatial transformations in graph comprehension', *Diagrams*, pp 1-19

Trickett, SB, and Trafton, JG, 2006, 'Toward a comprehensive model of graph comprehension: Making the case for spatial cognition', *Diagrams*, pp 286-300

Tufte, ER, 1983, *The visual display of quantitative information* Graphics Press, Cheshire, CT

Tversky, B 1995, 'Cognitive origins of graphic conventions', in *Understanding images*, ed FT Marchese, Springer-Verlag, New York, pp 29-53

Vernon, MD, 1946, 'Learning from graphical material', *British Journal of Psychology*, vol 36, no 3, pp 145-158

Vessey, I, 1991, 'Cognitive fit: A theory-based analysis of the graphs versus tables literature', *Decision Sciences*, vol 22, no 2, pp 219-240

Wainer, H 1992, 'Understanding graphs and tables', *Educational Researcher*, vol 21, no 1, pp 14-23

Xi, X, 2005, 'Do visual chunks and planning impact performance on the graph description task in the speaking exam?' *Language Testing*, vol 22, no 4, pp 463-508

Zacks, J , and Tversky, B, 1999, 'Bars and lines: A study of graphic communication', *Memory and Cognition*, vol 27, no 6, pp 1073-1079

APPENDIX 1: CONSENT FORM

Dear Participant,

Thank you for agreeing to participate in the Cog-Pro research project funded by British Council through its IELTS Research Programmes (www.ielts.org) and carried out by the consultants from the University of Bristol (www.bristol.ac.uk/education) in October and November 2008. This project aims to gain better understanding of IELTS Academic Writing Task One (AWT1) that uses graphs as test prompts. The data collection for this research would involve you (in the time order):

- taking IELTS academic writing tasks 1 and 2,
- answering a questionnaire measuring your graph familiarity and comprehension,
- taking IELTS AWT1 tasks, while thinking aloud your test taking process,
- having free training on how to take IELTS AWT1 tasks provided by the consultants,
- taking IELTS AWT1 tasks again, while thinking aloud your test taking process,
- (some of you) being interviewed on a one-to-one basis on how you took the AWT1 tasks

Your think-aloud and interviews will be audio-recorded. Your participation is voluntary and will not be paid, but we will provide you with free training and assessment on AWT1 during data collection (see above). As a potential IELTS test takers, you will benefit from participating in this research. You have the right to withdraw your participation any time if so you wish without any consequences, but we would like to encourage you to work your best until the end of the project to maximize your learning benefits.

We would like to ask for your consent formally, as recommended by the ethical guidelines for the conduct of research of International Language Testing Association (www.iltaonline.com) and British Association for Applied Linguistics (www.baal.org.uk). All data collected for this research (including your age, gender, graph familiarity, test performance, audio-recorded think-aloud protocols and interviews) will be anonymised and used solely for this research in a fair and respectful manner, in its research report and subsequent academic publications and disseminations. Your data will be protected in accordance with the Data Protection Act 1998.

We would be very grateful if you could read this consent form carefully and sign below, and if you do sign them, to indicate the manner in which you would like your contribution to be acknowledged in the research report and any publications and disseminations based on this.

Your Chinese name [in print]_____ Signature_____ Date_____

Please select either A **or** B for acknowledgement of contribution to this research.

- I would like acknowledgement and thanks expressed generically, i.e. to the students at Zhejiang University. OR [please tick here ____]
- I would like acknowledgement and thanks expressed to mention me explicitly, i.e. to the students at Zhejiang University, which includes (my name). [please tick here ____]

If you have any queries about the Cog-Pro project or this consent form, please get in touch.

Best wishes, Guoxing Yu, Graduate School of Education, University of Bristol

APPENDIX 2: IELTS AWT1 TASKS WITHOUT THINK-ALOUD (STAGE 1)

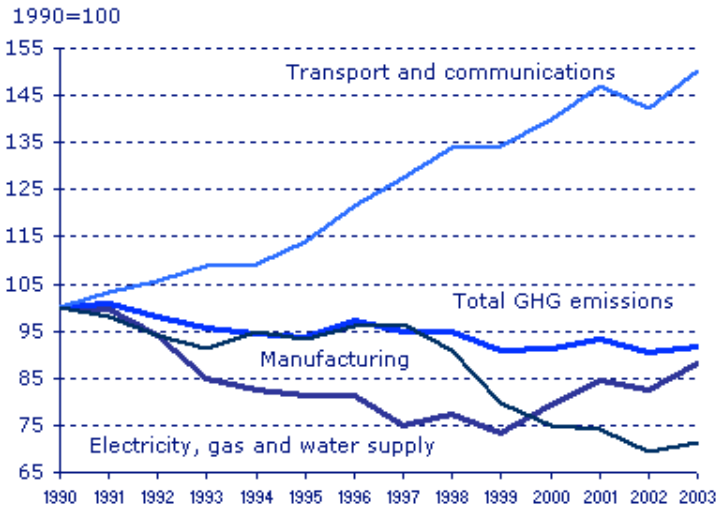
WRITING TASK ONE

You should spend about 20 minutes on this task.

The following graph shows the total UK greenhouse gas (GHG) emissions between 1990 and 2003 in comparison to 1990 as 100 in different end users

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



WRITING TASK TWO

You should spend about 40 minutes on this task.

Write about the following topic:

Once children start school, the teachers would have more influence in their intellectual and social development than parents.

To what extent do you agree or disagree?

Give reasons for your answer and include any relevant examples from your own knowledge or experience.

Write at least 250 words.

APPENDIX 3: GRAPHICACY QUESTIONNAIRE

This questionnaire will collect your personal information and your experience, familiarity and understanding of graphs (图表、数字统计图、数字统计表格、示意图、流程图等, including bar, line, chart, diagram, and table with numerical data). You are asked to provide **ONE** answer by ticking the relevant box or filling the blank which describes best your **OWN** situation. Please answer them **independently and honestly**. There is no right or wrong answer.

For example: Male [√]

If you don't fully understand a question, please ask the tutor for an explanation.

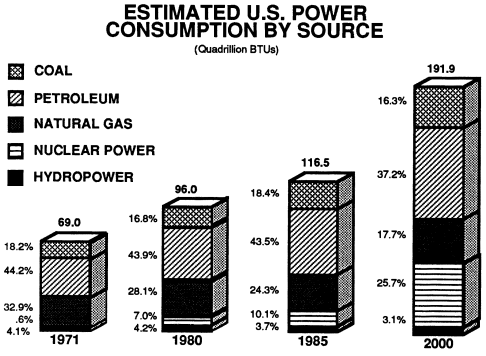
Personal information

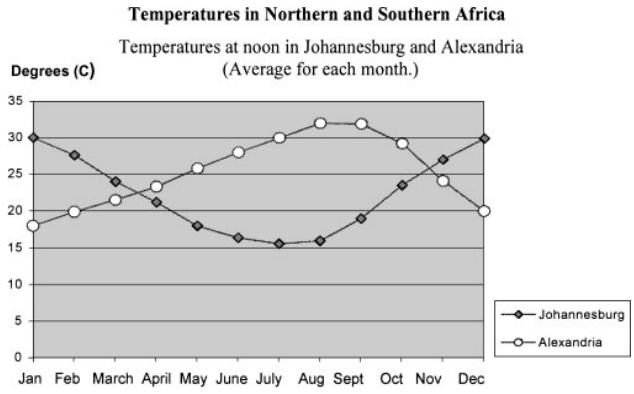
1. Your contact mobile phone number_____
2. Your email address_____ (Please in print)
3. Your CHINESE Name_____ (Please in print)
4. Gender: Male [ ] Female [ ]
5. Faculty/Department/Specialism_____
6. Year Group: [1] [2] [3] [4] [5] [6] [7]
7. Status: Undergraduate [ ]

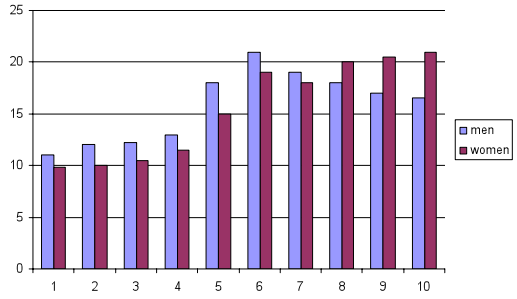
   Postgraduate [ ] If postgraduate, Master [ ] or PhD [ ]
8. Have you taken an IELTS test? Yes [ ] No [ ]
9. If YES to Question 8, when did you take your last IELTS test? [yyyy/mm]
10. If NO to Question 8, are you planning to take IELTS test? Yes [ ] No [ ]
11. If YES to Question 10, when are you planning to take IELTS test? [yyyy/mm]

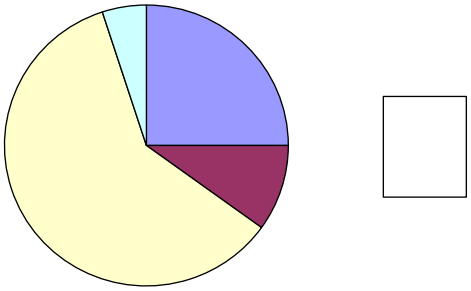
Questions on your graphicacy

Below are several statements concerning your experience, familiarity and preference in using graphs (图表、数字统计图、数字统计表格、示意图、流程图等, including bar, line, chart, diagram, and table with numerical data). Six **examples** of these graphs are given below. We will use GRAPHS as a generic term covering all these different types of graphs in this questionnaire, so your answer should reflect the **AVERAGE** of using these different types of graphs, unless otherwise stated.

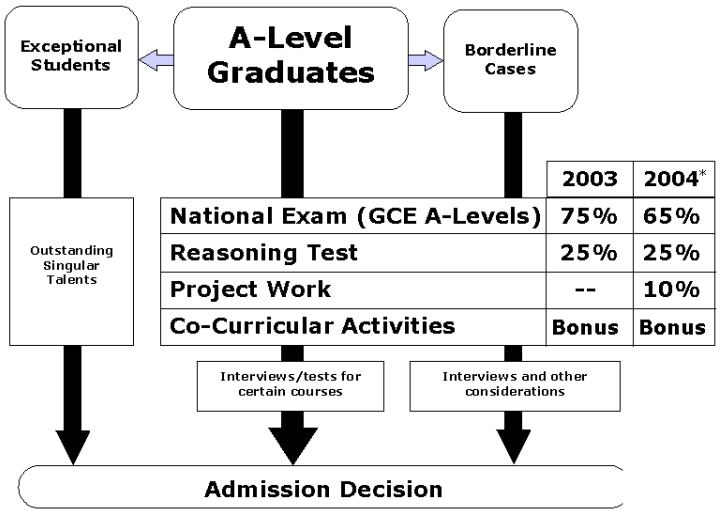








Countries	packaging	
	Tonnes exported in bags	Tonnes exported in containers
China	652	2001
India	4361	5002
New Zealand	82	44032



QUESTIONS START HERE

Please tick **ONE** number which best describes your own situation. There is no right or wrong answer.

Never ⟶ Very often

[1] → [2] → [3] → [4]→ [5] → [6]

I use special computer software to produce graphs. [1] [2] [3] [4] [5] [6]

As part of my academic study, I need to produce graphs. [1] [2] [3] [4] [5] [6]

As part of my academic study, I need to interpret graphs. [1] [2] [3] [4] [5] [6]

I read graphs in the popular press (e.g. magazines, newspapers). [1] [2] [3] [4] [5] [6]

When I read a graph, I try to identify the main trend or the overall pattern that the graph is trying to convey. [1] [2] [3] [4] [5] [6]

When I read a graph, I try to think about the possible underlying reasons for the main trend or the overall pattern of the data presented. [1] [2] [3] [4] [5] [6]

When I read a graph, I do not study it in detail. [1] [2] [3] [4] [5] [6]

When I encounter a graph in a text in popular press (e.g. magazines, newspapers), I tend to ignore/skip it. [1] [2] [3] [4] [5] [6]

When I encounter a graph in an academic paper, I tend to ignore/skip it.

[1] [2] [3] [4] [5] [6]

Strongly disagree ⟶ Strongly agree

[1] → [2] → [3] → [4]→ [5] → [6]

I am familiar with reading bar graphs. [1] [2] [3] [4] [5] [6]

I am familiar with reading line graphs. [1] [2] [3] [4] [5] [6]

I am familiar with reading pie charts. [1] [2] [3] [4] [5] [6]

I am familiar with reading diagrams representing a process. [1] [2] [3] [4] [5] [6]

I am familiar with reading tables with numerical data. [1] [2] [3] [4] [5] [6]

I notice errors or misinterpretations in graphs presented in the popular press.

[1] [2] [3] [4] [5] [6]

I notice errors or misinterpretations in graphs presented in academic papers in my field

[1] [2] [3] [4] [5] [6]

I recognize the different components of a graph (e.g. X and Y axes, legends, colours).

[1] [2] [3] [4] [5] [6]

I recognize how the different components of a graph (e.g. X and Y axes, legends, colours) are combined to represent the data. [1] [2] [3] [4] [5] [6]

I understand the relationships between a graph and the numerical data it represents.

[1] [2] [3] [4] [5] [6]

I can identify the relationships or the patterns displayed in one graph.

[1] [2] [3] [4] [5] [6]

I can identify the relationships or the patterns displayed in a few graphs about one similar theme

[1] [2] [3] [4] [5] [6]

I can tell when one type of graph is a better representation of the data than another.

[1] [2] [3] [4] [5] [6]

I can identify a poorly constructed graph. [1] [2] [3] [4] [5] [6]

I can revise and improve a poorly constructed graph. [1] [2] [3] [4] [5] [6]

I can describe the general trend or overall pattern of a graph in words. [1] [2] [3] [4] [5] [6]

I can use a graph to describe/convey the general trend or overall pattern of numerical data.

[1] [2] [3] [4] [5] [6]

I find graphs useful to vividly represent the numerical data. [1] [2] [3] [4] [5] [6]

I find graphs helpful for me to remember the key information in the numerical data.

[1] [2] [3] [4] [5] [6]

Graphs are a waste of space in a text. [1] [2] [3] [4] [5] [6]

I am concerned that I can not fully demonstrate my writing ability in IELTS Academic Writing Task One because I am not good at interpreting graphs. [1] [2] [3] [4] [5] [6]

I may do better in IELTS Academic Writing Task One using familiar graphs than unfamiliar ones.

[1] [2] [3] [4] [5] [6]

I would prefer one type of graph to be used in IELTS Academic Writing Task One.

[1] [2] [3] [4] [5] [6]

Special training on how to interpret graphs would be helpful for me to get a higher score in IELTS Academic Writing Task One. [1] [2] [3] [4] [5] [6]

Overall, on a scale of 1-6, how would you rate your own experience in using graphs?

[1] [2] [3] [4] [5] [6]

not experienced at all ⟶ very experienced

Overall, on a scale of 1-6, how would you rate your own ability in interpreting graphs?

[1] [2] [3] [4] [5] [6]

very weak ⟶ very strong

ADDITIONAL COMMENTS you want to make about your experience, familiarity and proficiency of using graphs. You can respond in English and/or Chinese.

Thank you for completing this questionnaire.

APPENDIX 4: THINK-ALOUD TRAINING DOCUMENT

Purpose of collecting think-aloud protocols:

In the Cog-Pro project, you will be asked to think-aloud while doing the IELTS Academic Writing Task One. The main purpose of collecting your think-aloud protocols is to understand your test-taking process. It will also help you to identify where you could and should improve through listening to your own think-aloud protocols later; and we are also going to use some think-aloud protocols as examples in the group training session to demonstrate what strategies you have used in completing the AWT1 tasks and how best to complete such tasks.

Practise thinking aloud

- The most important thing is to keep talking, ie, verbalizing what you are doing during the whole process including:

what you are reading,

what you are thinking and

what you are writing.

- You can use English and/or Chinese.
- Don't worry about grammar or sentence structure at all.
- **The tutor will remain silent unless you stop talking for more than 10 seconds**. In that case, the tutor will show you a white paper/card with TALK written on it.
- You are required to start recording with:

My name is _____.

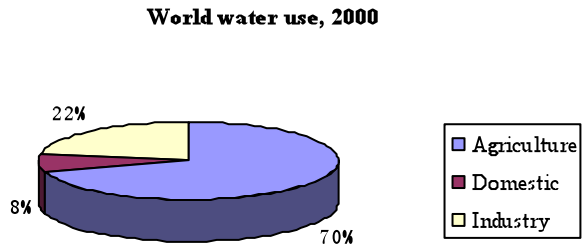
It is now ___ o'clock, on ___ of October 2008.

The task is _____ (Read the first paragraph of the task: eg, The following graph shows _____)

Then think-aloud when you do the AWT1 task.

Examples:

- 124+3546 = ?
- 124 x 378 = ?
- The average mass of 3 parcels is 6 kg. Two of the parcels have a mass of 4.6 kg and 6 kg. Find the mass of the third parcel.
- Try to think-aloud when doing the following AWT1 task.



Water use, selected countries, 2000

	China	India	New Zealand	Canada
Agriculture	69%	92%	44%	8%
Domestic	9%	5%	46%	12%
Industry	22%	3%	10%	80%

Below is an example of a think-aloud protocol when doing the task above.

(Taken from Li (2006) a MSc dissertation supervised by the lead author at the University of Bristol)

… The chart and table below give information about the way in which water was used in different countries in 2002, 2000. So we have a chart and a table. And then I have to summarise the information by selecting and reporting the main features, and make comparisons where relevant. So the tasks are the same. But we have different graphs. One is a pie chart, the other is a table. Now I am looking at the pie chart. The title of the pie chart is World Water Use, 2000. And we have 3 parts, agriculture, domestic and industry. And…ok, and each has different proportions. And then I am looking at the second graph which is a table. It is about water use in selected countries in the year 2000. So we have China, India, New Zealand, and Canada in the agriculture, domestic, industry. Ok. So it seems that there is connection between these two graphs. One is about the world water use in 3 categories. The other one is about eh, water use in selecting countries in these 3 areas. Yes. Now we have got a rough picture of what picture like. Now I am trying to find the main features.

In the first one, obviously, agriculture took the main proportion, 70%, and then it is industry and then domestic use. In the second graph, I found that in different countries, the proportion of the three categories are different. Like in India, agriculture took about 92% while Canada is only 8%**.** Now I can compare these figures. But there is much more information in the table than in the pie chart. So I will concentrate more on the table. Ok, I think I am going to give, make a very very brief draft. I will firstly give a opening paragraph, and then I am going to talk about, yeah, the first pie chart and then second table. If I have time, I will draw a conclusion. If I don't have time, just forget it. Ok, now start.

Opening paragraph, eh… we have some sentence pattern like report. Ok, it is reported that, ok, it is reported that in the pie chart that, oh, no no I made a mistake. Because I just directly to the second paragraph. I should have the opening paragraph. So, I should firstly say the chart and table, eh…I see, chart and table, below, yes, ok, below, describe, we use describe instead of giving information about, describe, how water was used in the whole world. I am changing the wording and paraphrasing in the world as well as in five different countries in the year 2000. Because it is in year 2000, actually it is past tense. This is very important…past tense. And then, describing the first graph. As something shows, as the pie chart shows, eh, agriculture, I should have a phrase here, take, account for is better. Agriculture accounted, past tense, accounted for 70% of the world water use in 2000. Now I am comparing so I can use the link word while, while domestic, while industry, industry took 22% and

domestic use, here use is a noun, domestic use, 8%. Because I only give the figure, and then I should give some comments, so the amount of the water used by agriculture was, say, was more than twice. Because 70% and together, eh, agriculture took 70% but together domestic and industry only took 30%. So you can say the amount of water used by agriculture was more than twice the amount of industrial and domestic use. Ok, seems enough.

Now, I am moving on to the table. This is more difficult and more complex. Therefore, I should pay more attention to it. Let's say. According to the table, I think I can, Ok, I think I can firstly give a brief account of the main information given in the table like in the five countries. No, four countries, one two three four, make a mistake, four countries, not five countries, just now miscounted. One of the four countries, yes, China and India used more water in agriculture while New Zealand and Canada not. And ok, so in China and India, agriculture accounted for the most water user while in New Zealand and Canada, it is not the case. Ok, ok, we will start with this. According to the table, India and China, agriculture, oh, maybe I should just refine wording in the last sentence of the second paragraph. I should say, the amount of the water used by agriculture was more than twice the amount of industrial and domestic use together. I can add together to make it more accurate. Now I am continuing. According to the table of India and China, agriculture, eh, took the largest proportion of water use in the year 2000. eh, yes, ok, yes, right…ok…took the largest proportion of water use in the year 2000. India used 92% while China, I am comparing, while China 69%. Eh, however, in New Zealand and Canada, because Canada use the smallest amount of water in agriculture, I can put Canada first. In Canada and New Zealand, oh, I should change, it didn't mention, because it is different, in New Zealand, domestic use accounts for biggest proportion in Canada, it's industry. I should change. I should describe it one by one. However, in Canada, industry took the, I should, I should change the wording, paraphrase, industry took the, took the most amount of water use. The most amount, the largest amount. Industry took the largest amount of water use reaching 80%. While in New Zealand, I can use the sentence pattern to emphasize. It was, while in New Zealand, it was domestic use that accounted for the biggest water use, ok, so about largest water user in the four countries.

And now, I should discuss, I think, I can discuss the smallest water user. Yeah, ok, Eh, In India, industry only used 3%, thus, eh, thus becoming the smallest water user. And domestic only, only, a kind of, this is a kind of redundant. Use the same phrase again and again. In India industry only used 3%, thus becoming the smallest water user and domestic just 5%. China, China dedicated 22% of this water to industry while only 90% in domestic. New Zealand, New Zealand, eh, actually, I think I should talk about, more about the water use in New Zealand. Because the amount of water use in agriculture and domestic were all the same. I think, I can add information, ah, I can add information here to New Zealand. That is water use in domestic is ok. While in New Zealand it was domestic use that accounted for the biggest water use, making 46%...making 46%. Eh, I can say, It is noted that the water use in agriculture was 44% and close to industry in New Zealand. It is good. New Zealand…only…I am continuing writing about the smallest water user. New Zealand only use 10% of water in its industry. Ok, last and least turn to Canada. Canada, Canada gave a very small proportion of its water to agriculture. As low as 8% and only 12% for domestic use. Ok, Ok. So I am almost done. I still have four minutes. I can just give a conclusion. Eh, in conclusion, the proportions of different water uses in the world vary from country to country. It appears that in developing countries, more water was used in agriculture while in developed countries, more in industry in the year 2000. Ok, I think, that is…I think I add one more sentence, as I still have one more minute. In conclusion, the proportions of different water uses in the world vary from country to country. Although, agriculture account, accounted for a significant majority in world water use. it appears that in developing countries, more water was used in agriculture while in developed countries, more in industry in 2000. Ok.

APPENDIX 5: THE EIGHT AWT1 TASKS

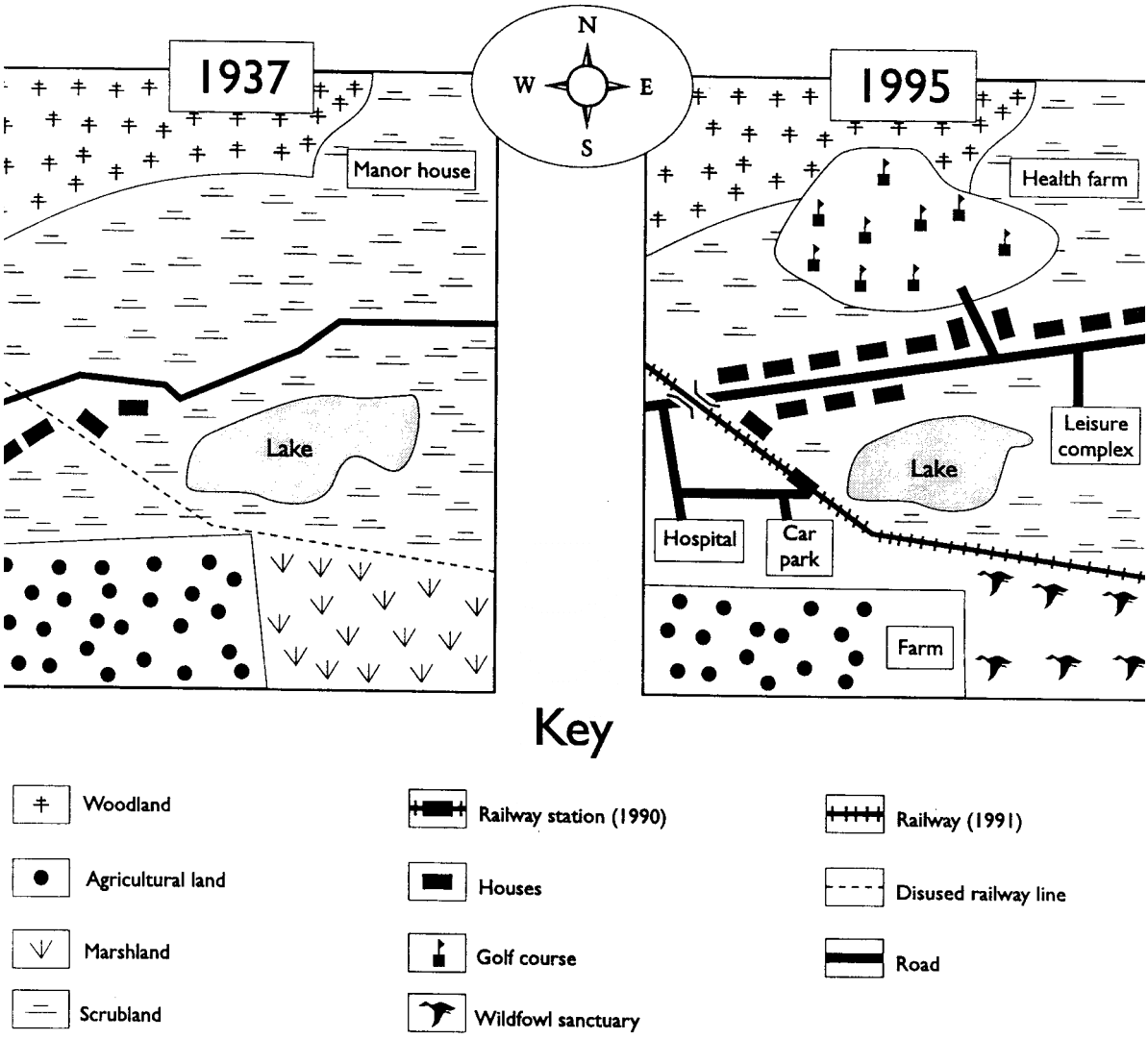
Set A (four tasks)

You should spend about 20 minutes on this task.

The following graph shows the layout of an area in 1937 and 1995.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



You should spend about 20 minutes on this task.

The following table shows the latest statistics of the top 15 countries that the US imports crude oil per day on average.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.

Crude Oil Imports (Top 15 Countries), Thousands Barrels per day on average					
Country	Jul-08	Jun-08	YTD 2008	Jul-07	YTD 2007
Canada	1,960	1,883	1,899	1,818	1,872
Saudi Arabia	1,661	1,479	1,543	1,434	1,411
Mexico	1,200	1,124	1,194	1,469	1,457
Venezuela	1,187	1,085	1,038	1,167	1,117
Nigeria	741	946	993	890	1,003
Iraq	696	693	677	460	473
Angola	640	636	517	392	542
Brazil	241	280	224	147	156
Algeria	232	269	306	537	500
Ecuador	226	178	197	159	189
Russia	202	228	123	99	130
Colombia	178	177	182	207	122
Azerbaijan	134	53	57	68	49
Kuwait	122	179	205	197	194
Chad	108	107	102	61	68

Source: US Energy Information Administration

Notes:

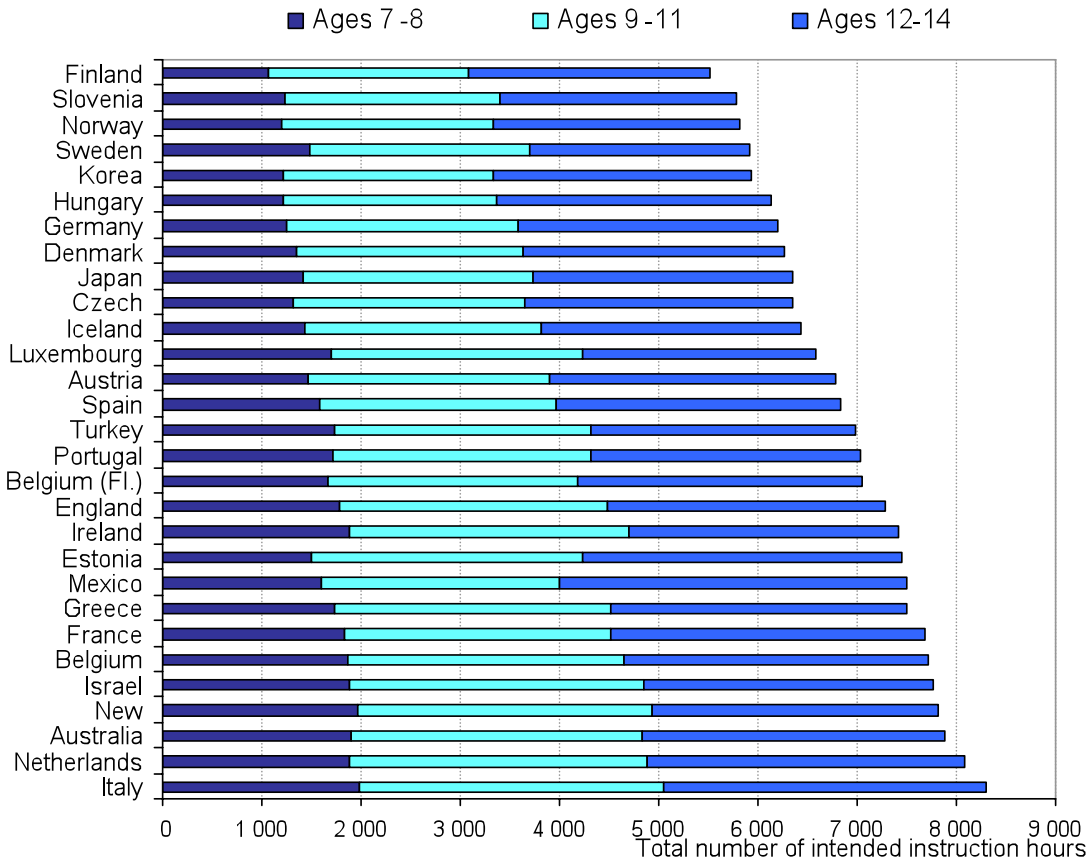
YTD=Year to date.

You should spend about 20 minutes on this task.

The chart below shows total number of intended instruction hours in public institutions between the ages of 7 and 14 (2005). Countries are ranked in ascending order of total number of intended instruction hours.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



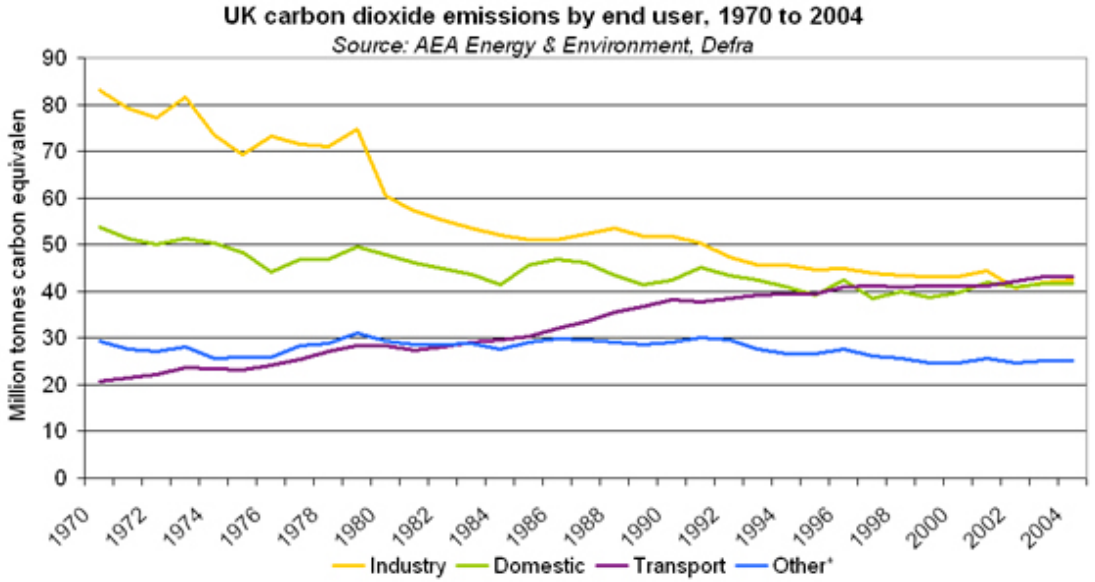
Source: Organisation for Economic Co-operation and Development

You should spend about 20 minutes on this task.

The graph shows the UK CO_2 emissions by end user from 1970 to 2004.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



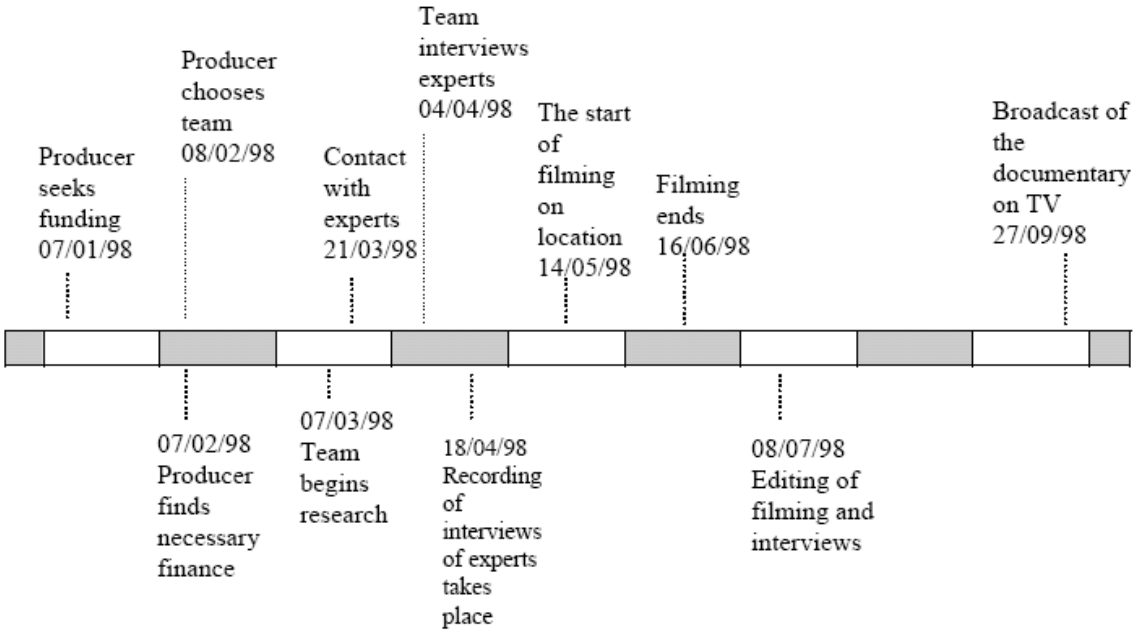
Set B (four tasks)

You should spend about 20 minutes on this task.

The following diagram shows the sequence of the events that led to the broadcast of a documentary by a TV programme.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.

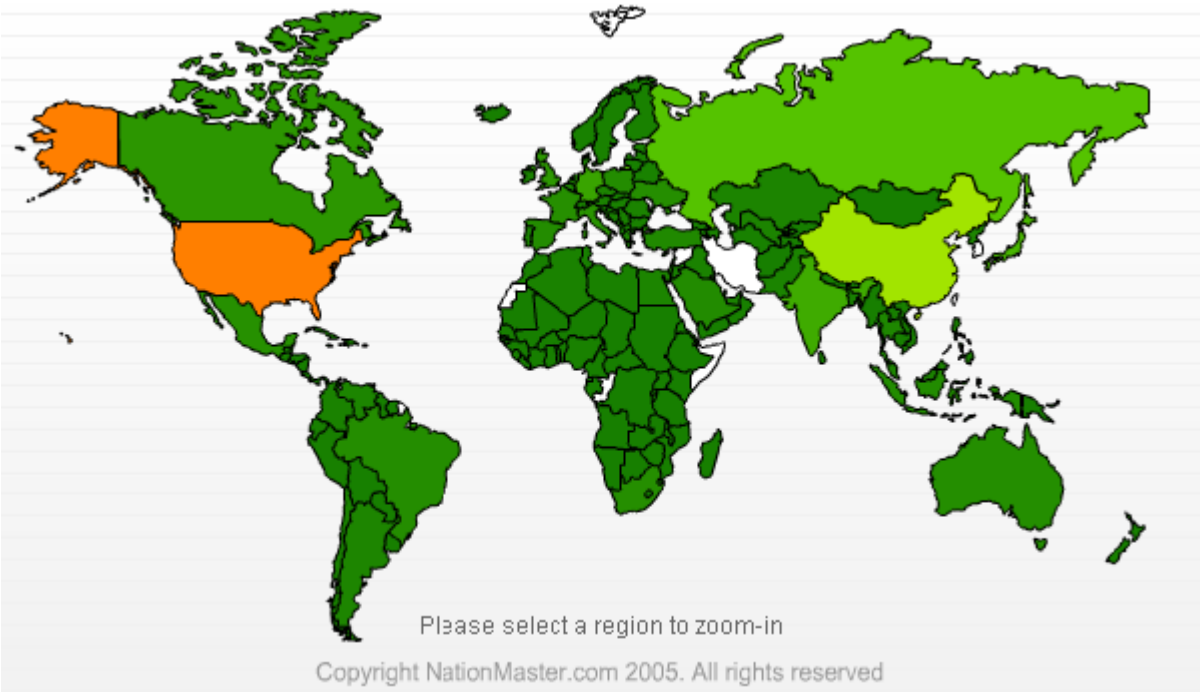


You should spend about 20 minutes on this task.

The following map and table show the amount of CO_2 emission by the top 8 countries.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



Legend: Top Middle Bottom (No data)

5.76 million 128,255 4

Rank	*Countries*	*Amount (top to bottom)* *Units: thousand metric tonnes of carbon dioxide*
#1	United States	5,762,050
#2	China	3,473,600
#3	Russia	1,540,360
#4	Japan	1,224,740
#5	India	1,007,980
#6	Germany	837,425
#7	United Kingdom	558,225
#8	Canada	521,404

CO_2 emissions

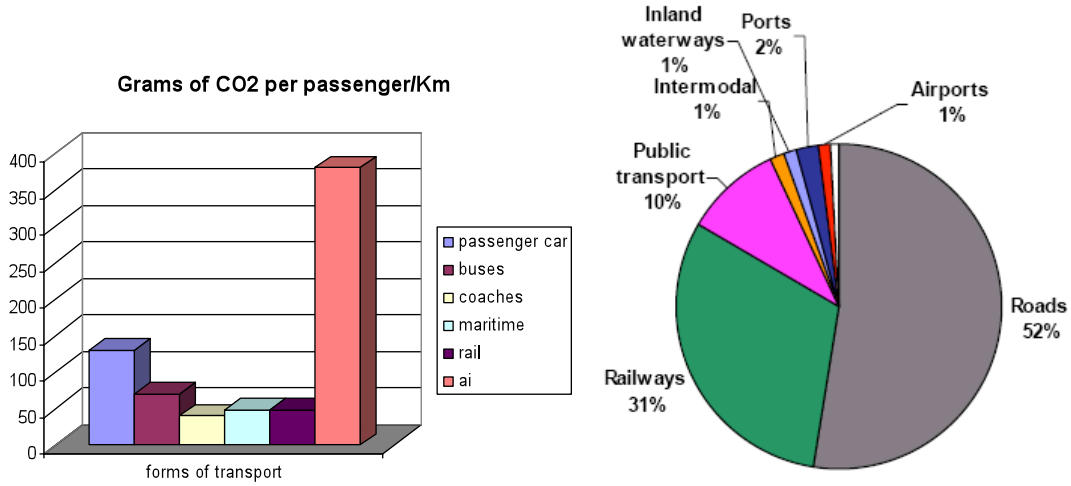
Source: World Resources Institute (2003)

You should spend about 20 minutes on this task.

The column chart below shows CO_2 emissions for different forms of transport in the European Union. The pie chart shows the percentage of European Union funds being spent on different forms of transport.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.

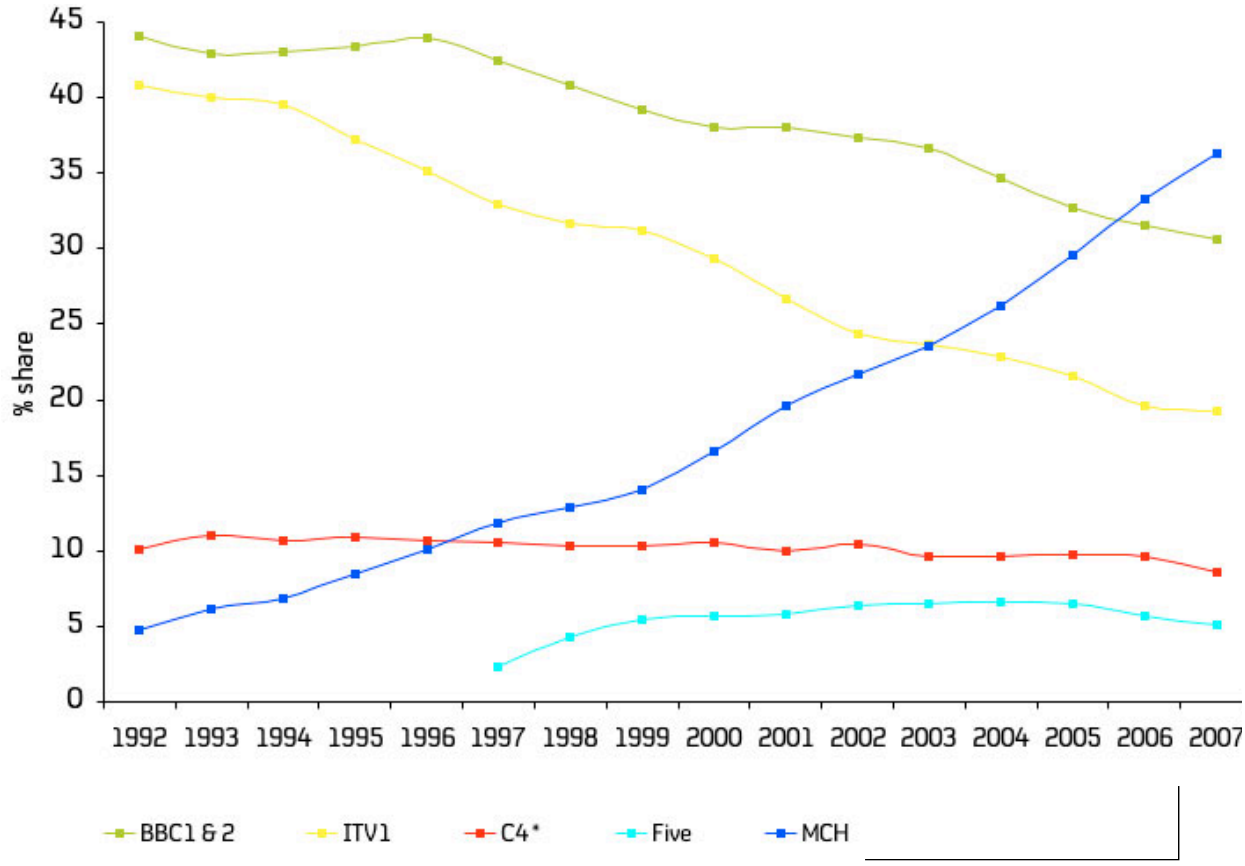


You should spend about 20 minutes on this task.

The following chart shows the individuals viewing share of the major TV channels in the UK.

Summarise the information by selecting and reporting the main features, and make comparisons where relevant.

Write at least 150 words.



Source: BARB/TNS Ratings Analyser and InfosysTV, Network Homes

APPENDIX 6: STUDENT EVALUATION OF THE COG-PRO AWT1 TRAINING

This questionnaire is an opportunity for you to record your assessment of our teaching and your learning experience on the IELTS AWT1 training provided by the Cog-Pro project funded by British Council. Please return it to the tutor. Thank you.

Please rate, by ticking the appropriate box on the right and use the space below each question to provide more information in CHINESE and/or ENGLISH.

When ticking the boxes below, please note:

Excellent/Really Good Disappointing

6…….5 ……..4……..…3…….2……1

1. THE TRAINING OVERALL [6] [5] [4] [3] [2] [1]

Overall, the training was useful for my preparation for IELTS test.

Comment:

2. LEARNING SUPPORT OVERALL [6] [5] [4] [3] [2] [1]

Overall, the learning support was helpful for my preparation for IELTS test.

Comment:

2.1 LEARNING SUPPORT: handouts [6] [5] [4] [3] [2] [1]

The handouts provided were useful for my preparation for IELTS test.

Comment:

2.2 LEARNING SUPPORT: content [6] [5] [4] [3] [2] [1]

The content of the training covered was useful and relevant.

Comment:

eg, Which areas are MOST useful?

Which areas are LEAST useful?

What areas do you wish the training should have covered?

2.3 LEARNING SUPPORT: the quality of teaching [6] [5] [4] [3] [2] [1]

The training was interesting and effective.

Comment:

3. YOUR LEARNING SO FAR from the training [6] [5] [4] [3] [2] [1]

Your self-assessment of your learning from the training.

Comment:

4. YOUR OWN CONTRIBUTION to the training [6] [5] [4] [3] [2] [1]

How would you evaluate your own contribution to the delivery of the training?

Comment:

5. ADDITIONAL COMMENTS

Thank you for completing this questionnaire

APPENDIX 7: INTERVIEW QUESTIONS

1. Briefing the purpose of the interview: to better understand your thinking process when doing IELTS AWT1 tasks.
2. Asking the students to talk about their experience of doing the AWT1 tasks, in particular, what is their general impression of the tasks, which task(s) do they find more challenging and why?
3. In what ways, do you think your AWT1 writing process may be affected by different graphs/prompts? Did you work differently for different graphs?
4. In what ways, do you think your AWT1 writing process may be affected by your familiarity and comprehension of graphs?
5. In what ways, do you think your AWT1 writing process may be affected by your writing ability?
6. In what ways, do you think your AWT1 writing process was changed due to the group training we had the other day? Did you do the tasks differently before and after the group training? What do you want/think the test preparation/training for AWT1 tasks should look like?
7. Any other comments

Notes:

- All participants to be interviewed.
- The recorded think-aloud protocols may be revisited if necessary at the interviews.
- The interviews are to be recorded.
- An interview will last around 45-60 minutes.

APPENDIX 8: FREQUENCY OF THE RESPONSES TO THE GRAPH FAMILIARITY QUESTIONS (12-46)

Question	Statement	Frequency						mean	Std. deviation
		1	2	3	4	5	6		
12	I use special computer software to produce graphs.		4	4	5	3	8	4.29	1.52
13	As part of my academic study, I need to produce graphs.		1	2	4	6	11	5.00	1.18
14	As part of my academic study, I need to interpret graphs.	1		4	4	9	6	4.58	1.28
15	I read graphs in the popular press (e.g. magazines, newspapers).	1	5	2	9	2	5	3.88	1.51
16	When I read a graph, I try to identify the main trend or the overall pattern that the graph is trying to convey.		1	2	6	7	8	4.79	1.14
17	When I read a graph, I try to think about the possible underlying reasons for the main trend or the overall pattern of the data presented.		2	2	8	10	2	4.33	1.05
18	When I read a graph, I do not study it in detail (data recoded)		2	1	5	13	3	4.58	1.06
19	When I encounter a graph in a text in popular press (e.g. magazines, newspapers), I tend to ignore/skip it. (data recoded)	3		2	6	9	4	4.25	1.51
20	When I encounter a graph in an academic paper, I tend to ignore/skip it. (data recoded)		1		3	7	13	5.29	.999
21	I am familiar with reading bar graphs		1	4	5	7	7	4.63	1.21
22	I am familiar with reading line graphs		1	1	7	4	11	4.96	1.16
23	I am familiar with reading pie charts			1	10	4	9	4.88	.99
24	I am familiar with reading diagrams representing a process		1	5	6	8	4	4.38	1.14
25	I am familiar with reading tables with numerical data		1	1	10	7	5	4.58	1.02
26	I notice errors or misinterpretations in graphs presented in the popular press		9	3	7	3	2	3.42	1.35
27	I notice errors or misinterpretations in graphs presented in academic papers in my field	2	4	5	3	9	1	3.67	1.47
28	I recognize the different components of a graph (e.g. X and Y axes, legends, colours)				7	9	8	5.04	.81
29	I recognize how the different components of a graph (e.g. X and Y axes, legends, colours) are combined to represent the data		1	3	5	9	6	4.67	1.13
30	I understand the relationships between a graph and the numerical data it represents			1	6	10	7	4.96	.86
31	I can identify the relationships or the patterns displayed in one graph			3	6	10	5	4.71	.96

Question	Statement	Frequency						mean	Std. deviation
		1	2	3	4	5	6		
32	I can identify the relationships or the patterns displayed in a few graphs about one similar theme	1		4	6	11	2	4.33	1.13
33	I can tell when one type of graph is a better representation of the data than another		2	3	3	9	7	4.67	1.27
34	I can identify a poorly constructed graph	1	3	7	6	6	1	3.67	1.24
35	I can revise and improve a poorly constructed graph		2	5	10	6	1	3.96	.999
36	I can describe the general trend or overall pattern of a graph in words		2	3	7	9	3	4.33	1.13
37	I can use a graph to describe/convey the general trend or overall pattern of numerical data			3	6	11	4	4.67	.92
38	I find graphs useful to vividly represent the numerical data			4	2	7	11	5.04	1.12
39	I find graphs helpful for me to remember the key information in the numerical data		2		5	5	12	5.04	1.23
40	Graphs are a waste of space in a text (data recoded)					7	17	5.71	.46
41	I am concerned that I can not fully demonstrate my writing ability in IELTS Academic Writing Task One because I am not good at interpreting graphs		6	5	5	5	3	3.75	1.39
42	I may do better in IELTS Academic Writing Task One using familiar graphs than unfamiliar ones			1	8	9	6	4.83	.87
43	I would prefer one type of graph to be used in IELTS Academic Writing Task One	1	3	1	11	5	3	4.04	1.30
44	Special training on how to interpret graphs would be helpful for me to get a higher score in IELTS Academic Writing Task One			2	1	13	8	5.13	.85
45	Overall, on a scale of 1-6, how would you rate your own experience in using graphs?		4	4	5	11		3.96	1.16
46	Overall, on a scale of 1-6, how would you rate your own ability in interpreting graphs?		1	7	11	5		3.83	.82

N=24.

APPENDIX 9: THE MOST FREQUENTLY USED WORDS IN EACH TASK

Task	*No. of tokens*	*No. of types*	*Average word length*	*Frequently used words specific to the task*	*Frequency*	*%*
A: Instruction	3263	402	4.46	Age/ages/aged	121	3.71
				Between/betweens	38	1.16
				chart	29	0.89
				children	34	1.04
				countries/country	73	2.24
				difference/differences/different	49	1.51
				Finland	28	0.86
				Group/groups	15	0.46
				Hours	116	3.56
				Institions/institutions/institution	33	1.01
				Instruct/instruction/instructions/instructing	83	2.54
				Intend/intended/intending	73	2.23
				Italy	38	1.16
				More	37	1.13
				Number	63	1.93
				Public	36	1.10
				Show/shows/shown/showed	16	0.49
				Spend/spent	29	0.89
				Total	37	1.13
				We	22	0.67
A: Layout	3510	493	4.23	Agriculture/agricultural	16	0.46
				Area	64	1.82
				Becomes/become/becoming	13	0.37
				Build/builded/building/built	38	1.08
				Change/changed/changes	36	1.03
				Complex	20	0.57
				Course	24	0.68
				Farm	26	0.74
				Hospital	22	0.63
				Lake	27	0.77
				Land/lands	33	0.94
				Layout/layouts	16	0.46
				Leisure	20	0.57
				Marshland	20	0.57
				More	49	1.40
				Park	19	0.54
				People	21	0.60
				Picture/pictures	17	0.48
				Railway	30	0.85
				Road	42	1.20
				Scrubland	35	1.00
				Smaller/small	17	0.48
				Wildfowl	17	0.48
				Woodland	17	0.48
				Year/years	17	0.48

Task	*No. of tokens*	*No. of types*	*Average word length*	*Frequently used words specific to the task*	*Frequency*	*%*
A: Oil	3178	406	4.32	Algeria	14	0.44
				Average	43	1.35
				Azerbaijan	14	0.44
				Barrel/barrels	32	1.00
				Canada	31	0.98
				Countries/country	93	2.93
				Crude	72	2.27
				Day	52	1.64
				Decrease/decreases/decreased	12	0.38
				Diffent/different/difference/differently	19	0.59
				Export/exportation/exports/exporter/exported	24	0.75
				Import/importation/imports/imported	77	2.42
				Increase/increases/increased/increasing	21	0.65
				July	29	0.91
				More	25	0.79
				Per	51	1.60
				Saudi	17	0.53
				Table	27	0.85
				Thousand/thousands	27	0.85
				Top	27	0.85
				US	62	1.95
				Year/years	22	0.69
				YTD	14	0.44
A: UKCO2	3461	506	4.42	By	58	1.68
				Carbon	15	0.43
				Change/changed/changes	12	0.35
				Decrease/decreases/decreased/decreasing	31	0.90
				Different/difference	13	0.38
				Domestic	50	1.44
				During	17	0.49
				Emissed/emission/emissions/emitted	123	3.56
				End	31	0.90
				Fluctuation/fluctuations/fluctuate	10	0.30
				Graph/graphs	22	0.64
				Increase/increasing/increased/increases	22	0.64
				Million	33	0.95
				More	23	0.66
				Other/others	39	1.19
				Same	17	0.49
				Show/shows/showed/shown	18	0.53
				Transport/transportation/transports	58	1.68
				UK	36	1.04
				Use/user/users/using	45	1.30
				Year/Years	47	1.36

Task	*No. of tokens*	*No. of types*	*Average word length*	*Frequently used words specific to the task*	*Frequency*	*%*
B: Broadcast	3046	385	4.37	After	28	0.92
				Began/begin/beginning/begins	32	1.05
				Broadcast/broadcasted/broadcasting	47	1.54
				Choose/chose/choosing/chooses/chosen	20	0.66
				Contacts/contact/contacting	19	0.62
				Day/days	12	0.40
				Diagram	18	0.59
				Documentary	52	1.71
				Edit/editing/edits	23	0.76
				End/ends/ended	13	0.42
				Event/events	28	0.92
				Experts/expert	43	1.41
				Film/filming	53	1.74
				Finance	16	0.53
				First/firstly	14	0.46
				Funding/fundings	19	0.62
				Interview/interviewing/interviews/interviewed	52	1.84
				Led	15	0.49
				Location	16	0.53
				Make/making	12	0.39
				Months/month	64	2.10
				Necessary	16	0.53
				Produce/producer/production/producing	37	1.20
				Program/programmes/programme	33	1.09
				Record/recording/recorded	16	0.52
				Research	17	0.56
				Seek/seeked/seeks/seeking	19	0.63
				Sequence	18	0.59
				Show/showed/shows/shown	13	0.42
				Start/started/starts	23	0.76
				Step/steps	16	0.53
				Take/takes	15	0.49
				Team	43	1.41
				TV	44	1.44

Task	*No. of tokens*	*No. of types*	*Average word length*	*Frequently used words specific to the task*	*Frequency*	*%*
B: EU Fund	3511	374	4.40	Air	29	0.83
				Airport/airports	14	0.39
				Buses	21	0.60
				Cars/car	24	0.69
				Chart/charts	77	2.19
				Coach	28	0.80
				Column	23	0.66
				Different	45	1.28
				Emit/emission/emissed/emissions	58	1.65
				EU/Europe/European	59	1.68
				Forms	63	1.79
				Fund/funds	71	2.02
				Gram/grams	52	1.48
				Inland	15	0.43
				Intermodal	14	0.40
				Km	41	1.17
				Large/larger/largest	22	0.63
				Maritime	20	0.57
				More	18	0.51
				One	20	0.57
				Only	17	0.48
				Other/others	24	0.69
				Passenger/passengers	71	2.02
				Per	49	1.40
				Percent/percentage/percentages	34	0.97
				Pie	31	0.88
				Public	17	0.48
				Rail/railway/railways	40	1.13
				Road/roads	22	0.63
				Second	23	0.66
				Show/showing/shows/showed	26	0.74
				Spend/spends/spent	52	1.48
				Take/taken/takes	27	0.77
				Transport/transports	100	2.84
				Union	27	0.77
				Waterway/waterways	15	0.43

Task	*No. of tokens*	*No. of types*	*Average word length*	*Frequently used words specific to the task*	*Frequency*	*%*
B: Map	3132	409	4.31	Amount	56	1.79
				Canada	22	0.70
				Carbon	18	0.57
				China	40	1.28
				Color/colour/colours/coloured	25	0.80
				Countries	106	3.38
				Differ/different/difference/differences	19	0.61
				Dioxide	18	0.57
				Emission/emissions/emit	97	3.09
				German/Germany	16	0.51
				Green	14	0.45
				India	20	0.64
				Japan	19	0.61
				Large/larger/largest	20	0.64
				Map	45	1.44
				Metric	44	1.40
				Million	32	1.02
				More	16	0.51
				Russia	28	0.89
				States	22	0.70
				Table	32	1.02
				Thousand	36	1.15
				Tonnes	44	1.40
				Top/tops/toppest	42	1.34
				United	31	0.99
				World/worlds	30	0.96
B: Viewing	3140	434	4.12	BBC	33	1.05
				Change/changes/changing	12	0.39
				Channel/channels	85	2.71
				Chart	24	0.76
				Decrease/decreased	18	0.57
				Drop/drops/dropped	11	0.35
				Five	34	1.08
				Increase/increased/increasing	11	0.35
				Individual/individuals	43	1.37
				Major	26	0.83
				MCH (multi-channel)	38	1.21
				More	16	0.51
				Only	13	0.41
				Remain/remained/remains	13	0.41
				Rise/rised/rises/rising	12	0.38
				Share	106	3.38
				Shown/showed/shows	19	0.60
				Some	12	0.38
				Stable/stabilized	12	0.38
				Trend/trends	14	0.45
				TV	38	1.21
				UK	20	0.64
				View/viewing	54	1.72
				Year/Years	61	1.94

Note: Words spelt wrong are also included in this analysis

APPENDIX 10: PERSONAL INTERPRETATIONS AND COMMENTS IN TASK B AND TASK A WRITINGS

Task	*Personal interpretations and comments*	*Participant*
B: Broadcast	This report suggested that the prepared work and the support work were the most important job in the broadcast.	C
	The hole [whole] task of finishing the documentary was difficult and need a lot of money and labours, and also with complex process. It would be benefitful for them to do some research first so they finished it in early March.	H
	Interestingly, we find that most of the process of the broadcast had been done comparatively long before it was broadcasted on TV.	K
	In conclusion, the broadcast of a documentary by a TV programme takes a lot of time and work, and it also needs financial support to make all of these happen.	L
	Finally, the greatest moment came. The documentary was broadcasted on TV on 27/09/98. The diagram gives us a complete and clear picture about the sequence of the events that led to the broadcast of a documentary by a TV programme.	Q
	Doing an broadcast of the documentary on TV is an hard work. We need a plan and do just as the plan.	X
B: EU Fund	From the two charts, we can conclude that the transport receiving the largest proportion of EU funds is not the one that emits the most CO_2 for different forms of transport.	G
	So it may be possible to speculate that we would promote the development of buses, coaches, rail and so on.	J
	When comparing the bar chart and pie chart, we may find that, generally, the form make more grams of CO_2, the less percent of fund it would receive.	K
	The conclusion of my report is: The amount of CO_2 in different forms of transport are quite different, and the pecentiage [percentage] of funds be used in different kinds of transport are also very different from each other, and the coast of money in one kind of transport may not be linked with its CO_2 emission.	N
	(think-aloud): I don't actually know the meaning about the forms of the transport, I can't find the corresponding parts in the pie chat for each form of transport. … Then I make some conclusion. The emission cause the funds increase or the funds down cause the emission increase. What is the relationship between them? I can't decide it up. And I don't know the efficiency about the funds. I can only get the trends that if pay more money on the transport, the emission will go down. Actually, the deep relationship between them, I can't say much clearly.	T

Task	*Personal interpretations and comments*	*Participant*
	Maybe the government need to expend more funds on the air transport to reduce the CO_2 emissions of air to control the pollution.	X
B: Viewing	we can draw a conclusion that, the five major TV channels got different individuals viewing share and had different dropping or growing trends year by year.	B
B: Map	To sum up, the amount of CO_2 emission is different among countries all over the world. And, the difference is a significance.	B
	In a word, North America and Asia were the largest CO_2 emission countries". (Note: summary based on previous knowledge)	C
	The square of Japan is much smaller than other countries in the table, but it has the rank of 4 in CO_2 emission, even more than India's 1,007,980. (involving prior knowledge of the size of Japan)	D
	There are two countries in North America, three countries in Asia and three countries in Europe. The Middle East countries, like Iraq, just have no data, which means that these countries have rarely no CO_2 emission" (involving prior knowledge of continents. It misinterpreted the Iraq information)	E
	Unfortunately, China is the second country after America, over 3.47 million thousand metric tonnes of CO_2 emissions. In the map the area of China is described in light green. …. Hopefully the color of China would turn into dark green in the future.	G
	It is astonishing that Japan ranks 4^{th}, since its total area of the country is not very large, compared to the countries rank at the tops.	H
	… and all of the top 8 countries of CO_2 emission are developed countries except China and India (involving knowledge about developed and developing countries)	J
	The most contribution of CO_2 output countries are located in north America and Asia. (involving prior knowledge about continent)	N
A: Instructions	In the chart, we can see that the countries which have fewer instruction hours like in North Europe, such as Finland, Norway, Sweden and so on". (involving prior knowledge about geography)	C
	From the chart, the top three countries Finland, Norway and Sweden are all from the North Europe. These countries all do very well in instructing the children in public institutions. So the children can meet less problem when they are in these public institutions with no adults. (involving prior knowledge, but wrong interpretation re: *with no adults*)	E
	So I think if we want the children to do more instructions, we should encourage them to do it when they are at the earlier ages from 7-8, since maybe earlier will be better. (personal judgement and evaluation)	H
	So the hours shows a steady increase in these 3 age groups, which shows a quick development in the ability of acceptable learning time of children.	L
A: Oil	Last, I want talk about the Iraq, because of the Iraq War. There is a big growth from 2007 to 2008. I think, after the Iraq War, US imported more oil from Iraq.	B

Task	*Personal interpretations and comments*	*Participant*
	First, I see there are both developed and developing countries above.	D
	Canada is US's neibough, so the US imports the largest crude oil from Canada… The US imports more crude oil from American countries than those from other continents, especially those from Middle East, or Africa. There are not many European countries which have crude oil exports to the US, China, neither. Maybe the transport cost is very high, if the US wants to import crude oil from these countries or because these countries have no crude oil production.	E
	However, the earlier the data is the bigger the disturbances are, which means economies in differences changes differently, some developed fast and so they need more oil and some developed slow, so they don't need as much oil as others did. (wrong interpretation of the data: considering countries importing oil)	I
	There are some countries like Colombia, Azerbaijan, Kuwait, Chad import crule oils even less than 200, I think there're some reasons. One is that they are not very developed, as those countrys like Canada, and the other, they also can not provide enough crude oil with their industry. They're not rich, they don't have enough money to import the crude oil. I just skip the Venezula, I'm not familiar with it. And its statits is similar with Mexio, so I conclude that, there are three or four groups. One is that they import larger than 1000, one is that they import less than 200, the other is that they import between 200 and 100. Different reasons and facets in different countries cause the differences. For example, the Canada and Saudi Arabia import large every year, because they are well-developed country, and they have modern industry, which need more crude oil. And another example is Russia, it is also the developed country, but it has its own crude oil field and it doesn't need to import as much as Canada. And the third example is, I find that country, the Chad, the small country with poor industry, they don't have to import many crude oil, its industry can't consume as many oil as Canada. (Wrong interpretation)	T (from think-aloud protocol)
A: UK CO2	Furthmore, in my point of view, the measure to control the amount of CO_2 emissions is successful these years, especially to control the CO_2 emissions by industry. The measure can be done in the future. However the measure to control the transport CO_2 emissions is not so successful, I think better measure may be taken to solve the problem. And I think the problem may be because of the number of cars owned by personal user becomes larger and larger year by year. So we may use some way to enough people go out to work or to have great time by bus, by train or other public transport. Another way to solve the problem is to use new energy which produce less emmison than the fuel we use nowadays. (a large proportion of her writing is about finding out the underlying reasons)	B
	The reasons of the graphs reveals that more and more people began to own their private cars and transport make big contribution in CO_2 emissions in the UK. The industry and domestic graphs decrease because we use a lot of energy like electricity or nuclear power than 1970s. (a large proportion of her writing is about finding out the underlying reasons)	C

Task	*Personal interpretations and comments*	*Participant*
	The industry has been always developing, so I think the UK government must have took some methods and give a policy to cut down CO_2 emission in industry. And the CO_2 emissions by transport grows slowly during the years. It's because that the UK is using more transport, maybe its population grows or it has more travellers from the world. So the Government has to do well in his job to offer enough transport. I see that CO_2 emission by domestic changes a little but it seems to be periodic, that is it decrease in a year or a period, and in the next years, it increases. So is there relationship between domestic and economic because the economic development is also periodic. The industry produces less and less CO_2 while the transport produces more and more, I think the UK government was trying to change its economic…, to develop industry as a smaller part and to develop its … as a bigger part. (almost the whole piece of writing is about explaining the reasons)	E
	In conclusion, CO_2 emissions by industry is less and less than before, while we have seen an increasing emissions by transportation in the last 30 years. That's what we should care about.	G
	Interestingly, I think now the transport should be taken consideration carefully. I think it will be the big problem to our living environment. Because in this graph it is the only line climbing high step by step. We can see from the graph, in 1970 the CO_2 emission is about 20, but in 2004, the CO_2 emission excess the 40 million tonnes, just increased 2 times as before. All in all, I think the transport problem we should cope with earlier before it deterirating.	R
	Now I think it must be the reason why the government restrain the personal car and set up more rules go the licences. It may be used, I thought; Oh. It's not relation to the graph. I got back to it. Plus them, distinguish the difference among them and what will I do next. Maybe I'll find them a reason . The industry cut down its emission for the process and technology moderate, maybe the transport can cut down it emission at the same way. Like enwide the road, redesign the program of traffic lights, transport emission may go down by those measures.	T (from think aloud protocol)
	There is no end for developments of industry and domestic because the technology is still developing. So I predict that the UK CO_2 emissions of industry and domestic will increase in the future when CO_2 emission of transport increase to some level.	W
A: Layout	In my point of view, the development of world make great changes in the area of the picture. Firstly, in modern world, there's not enough space in cities for people to do sport, and we also like the woodland and scrubland which is hardly seen in the cities, so a golf course is built. Secondly, just I stated above to the space in the cities is not enough. As a result, more people build their houses here, as well as new hospital, car park, leisure complex and so on. In order ot go the places easily, people broaden the road and build the railway. Because of the use of land, our lake area and agricultural land become smaller. Furthermore, people now take more care of wild animals, so I think that's why the marshland in 1937 is replaced by the wildfowl sanctuary.	B
	Although there are a lot of buildings in 1995, the lake and the farm was becoming smaller and smaller. It's the cost of economic and social development.	C
	The economy level should have a great improve.	D

Task	*Personal interpretations and comments*	*Participant*
	In a word, I think that the land has been used to its maximum and more comfortable for people to work and live. Transport, entertainment and medical care are all getting better. The road also reaches the leisure complex, which adds more happiness to the living.	E
	The precious manor house had become a health farm in 1995. ….We can concluded that more people moved to this area in 1995 than before. It's certain that development would cost something.	G
	As the span of the time was about 60 years, and it was normally that there're so many changes in the same area, as the society was improved day by day. Then we can get the conclusion that. As the society improves day by day, there are much more modern buildings and constructions in the area. But as the scrubland and the woodland or the agricultural land are all important to our environment and will be benitful to the atmosphere which we rely to survive. And according to the changes during the past several years I think the trend of the changes in the future in this area will be beniful to the citizens reside in this area.	H
	Newly built roads would lead people to the hospital, car park, golf course and the leisure complex. Compared with the old road, the new is more straight, so that the trip would be shortened.	I
	And the land has become more convient; comfortable and beautiful than it was in 1937. I think a lot of people may come to this area for leisure activities. And it will become a famous land for wildlife.	J
	In conclusion, this area has changed a lot during the last sixty years and it is hard for us to reveal its original face.	K
	In conclusion, the traffic system in this area has greatly improved, and also several facilities have built to enrich the life of this area, which all turn this area into a more suitable place for people living and working.	L
	The changes of the layout also show the change of the society. The population increased and then the public organization.	Q
	All in all, from 1937 to 1995, this land become more suitable for people living. But I prefer the land of 1937, since of the more nature. For example, more lake, more farm and so on. …. The hospital, I think, is more advanced than the clinic in 1937, so it do good to our health. … The people in 1995 became more wealthy than in 1937.	R
	And 1995, there is a new hospital, car park and golf course, seems that there are some better life for the people who live on the land. The houses are along the road and they are 4 times as the numbers in 1937. The lake is shrinked, maybe caused by construction. The main change of the land, may have some reasons. Transport develop the agriculture and leisure consummation, so they may have the golf course and son on. The farmland didn't shrink, or cut down, means they may need more supplement for the new town. May call them the town, because there are some road and hospitals, car park, railway, it have developed to a town. The house are more, and the facilities is constructed to satisfied the need of the people in there: with land redesigned and rebuild, the people live in there have to adapt the change and make it more fitable to live there.	T (from think-aloud protocol)

Task	*Personal interpretations and comments*	*Participant*
	Although more buildings have been built, people are concerned more on the environment of this area, the plants in the woodland are well protected, so is the farm. In a word, this area is like a leisure place for tourists to enjoy their time, the great nature in 1995.	U
	We may see from those two picture that the layout of 1995 is more colorful and reasonable. It utilize land greatly. What's more, this layout offer wildfowl some better land, which can improve the balance development of the nature. At the same time, this layout make our life better. We have many and more different land to live.	W
	When the health of the people become more and more important, the hospital become. As the economy developed, more and more people have their own cars, so car park is needed on the north area…	X

APPENDIX 11: STUDENT EVALUATION OF THE PROJECT AND ITS TRAINING: GRADE AND QUALITATIVE COMMENTS

Note: blank means no additional comments were provided by the student.

1	5	if there are some good examples to analyses, cause what we see here is written by Chinese, that would be better
2	5	I now have an overview of the requirements of IELTS AWT1 tasks, how to analyze graphs and what I should pay attention to, what I should avoid, e.g. it's necessary to describe what is in graphs, but not to further reasoning
3	4	partly, I haven't take a view of the whole test
4	6	Through this training, I know a lot more about AWT1
5	4	
6	4	I'm not quite sure about the long-term effect yet, but think-aloud is very useful, it helps us to pay attention to the problems in our writing process and find some solutions to these problems.
7	5	To me this training is extremely useful, this kind of training is much better and effective than other intensive preparation courses and self-study
8	5	Now I have systematic understanding of AWT1
9	6	The lecture was very useful
10	5	
11	6	I think it is beneficial to both my writing and speaking ability
12	5	I have learned some special words and idioms for describing the trend or extent
13	5	The training made me aware of the way I think when doing the graph writing, and helped me develop myself. It also gave us some useful tips
14	5	Maybe it's a bit short. If the period can be arranged a little longer, such as 2 weeks, and the training be more scattered, the effects will be much better

1	6	
2	5	I now know a lot more useful words to describe graphs
3	4	partly, I wasn't ready psychologically, it was quite new to me, I need time to adjust myself and make the best use of the resources provided
4	5	It let me know what is AWT1
5	5	
6	4	The tutor listened to our think-aloud and pointed out some problems we had; this is really useful for individuals to notice our problems.
7	5	(same as for Q1:To me this training is extremely useful, this kind of training is much better and effective than other intensive preparation courses and self-study)
8	5	
9	6	The only pity is that there could be more feedback so that we know what our weaknesses are
10	4	
11	6	

12	5	
13	5	It gave us much better understanding of all types of graphs, background knowledge, it is eye opening, broaden my horizon.
14	3	The support can be given more individually

1	6	It is well organized and there are a lot of useful information
2	6	Excellent content coverage and well delivered lecture: I believe it will be very useful for my preparation for IELTS writings
3	4	Some of the information contained in the training package can be found online; but I will continue using think-aloud to train myself for the test
4	6	I want to be provided much more
5	6	
6	5	It explained in much detail about all aspects of AWT1, provided constructive feedback and suggestions, provided some samples and exercises; it is really helpful for us to understand and grasp AWT1
7	6	The materials are well prepared, with detailed and rich content, very good for test preparation
8	5	
9	6	
10	6	
11	6	
12	5	
13	6	These materials tell us what AWT1 tests, and the constructive and analytical tutorials, which are all very helpful preparation for AWT1 tasks
14	5	Although I've only read part of the handouts, I find it really useful. The tips were helpful to me.

ID	*Grade*	*Most useful content*	*Least useful content*	*Wish list of content*
1	6	Don’ts and dos based on my action		feedback of the essays we wrote
2	5	Especially the lecture on the third night	We think faster than speak/write. When thinking and write, we have time to re-organize ideas; but at think-aloud & write, we don't have time to re-organize, we say what we think, and sometimes our thinking can be interrupted, maybe it's because of my low English ability	I think it might be better for training if we wrote without think-aloud first, then think-aloud or re-visit our thinking
3	5	think-aloud training, it helps me to find out my weaknesses	I've already known some information contained in the handouts	It would be good to have more comments from the tutor on each individual writing
4	4	How to think aloud and describe a chart/table/diagram…		For specific graph or chart, how to describe and in what sequence we should organize our ideas and

				thoughts
5	4			
6	4	the third training session	Although think-aloud has great potential, some of the problems detected through think-aloud are the problems I'm already aware of, therefore I don't think it is the most efficient strategy to improve writing	how to analyse all sorts of different graphs, provide exemplars, and explain good reasoning strategies
7	5	interaction during the training	NONE	It seems that we have already had all.
8	4	interaction		the training could be longer
9	5	wish the project is longer, with more examples of writings		
10	6			
11	5	Think-aloud training		analysis of every task
12	5	Magic words	the history of IELTS	the samples of the given titles
13	6	the dos and don'ts, magic and useful words, think-aloud	NA	more detailed guidance on how to analyze different types of graphs
14	4	It provides me a new way to look at IELTS, especially AWT1. The task is new and up to date	No.	We may want more feedback from the tutor, as we didn't receive individual help from the tutor

ID	*Grade*	*Comments on teaching quality*
1	5	
2	5	It isn't necessarily interesting, but really attractive as it helps to improve my writing or to know more about IELTS writing test
3	5	It is interesting and effective, but a little rush; we need more time to digest what is there the helpful information
4	6	Good! Excellent!
5	5	
6	5	YES! The tutor is very patient and attentive; the training was very detailed and complete
7	6	the interaction
8	5	I haven't done similar training, I feel it is interesting.
9	5	
10	5	
11	6	The teacher broadened my horizon of the IELTS AWT1 and I felt this training unique and useful
12	4	
13	5	helps to know different types of graphs, background knowledge, very interesting. The tutor is passionate about what he teaches and humorous, which encouraged us to do the best in the project
14	4	Due to the long hours, I really felt a little tired when I needed to record what I'm thinking and writing

1	6	Through the handout and the training class I can describe the graphs much better than before
2	4	The training is too short, but it is specifically for AWT1, helpful to know how to do AWT1. I also got a chance to practise my writing
3	3	At the beginning, I haven't adjusted myself well; but later it was ok. I was there on time every day and finished all the tasks; gradually I am in the tasks now.
4	3	A bit improved, because the training time is so short
5	4	
6	4	it is a long process to learn English well, although in such a short time we know well AWT1 now ,it still needs more time and practise to integrate what we've learned; however, this is a very good direction, absolutely essential good direction and beginning
7	5	I feel I improved
8	5	
9	5	
10	3	
11	5	
12	5	
13	5	helped me to understand how I think when doing the tasks, my weaknesses and strength, as a result improving how I do AWT1 tasks, and my way of thinking
14	4	Before taking a test, I need to figure out why the examiner designed the question like this. By figuring it out, I can get a better score more easily.

1	6	I worked hard
2	4	My speaking and writing are not very good, but I've done the best I can; I don't think I've made much contribution to the training, but I am serious and sincere to be part of the project, alas, my English is not very good.
3	5	I'm working hard and I have full passion on it.
4	6	
5	4	
6	4	I've done the best I can during the training, as the teacher instructed, learned some methods and techniques; but my English is not very good, I apologize
7	5	I am serious and did the best to be part of the project, I hope my participation is useful to the project
8	4	I have done all the tasks – full participation
9	6	I think I was very diligent.
10	2	
11	5	
12	6	
13	5	I was actively engaged in this research project, did my best to complete all the tasks, according to the tutor's instructions
14	4	I hold onto this training and I really benefit from this training. But sometimes I felt tired and didn't do my best to finish the task.

ID	Additional comments
1	Thank you for the training and the whole experience provided to us.
2	Speaking for myself, I would suggest asking us to write first and when we finish writing we think-aloud
3	
4	Apart from AWT1, we would like to be provided more parts of IELTS
5	
6	AWT1tests academic skills in analyzing, it seems more important to learn the techniques/strategies. Think-aloud is very helpful, I wish to have one-to-one tutorials from the teacher; it is a pity that it is not possible
7	
8	
9	
10	
11	Thanks to this training, I have promoted my writing ability and speaking ability
12	I've begun to notice how many words to write, but sometimes I find too much to write. I can't end it because I still find something important to notice. Maybe I wrote a lot sentences that doesn't belong to major features. I should pay attention to my handwriting
13	the training is short and we don't have enough time to prepare, our improvement may not be immediate and obvious
14	some reward e.g. a little gift for the participants may encourage and stimulate the participants; and one-to-one communication to understand each one's problems

Note: This is an open question asking for the participants' additional comments, there is no grade.