

Carleton University

The Role of Preparation in Test Takers' Strategy Use and Test Performance

Christine D. Doe

School of Linguistics and Applied Language Studies

Carleton University, Ottawa

September 2006

A thesis submitted to the Faculty of Graduate Studies and Research

in partial requirements of the degree of Master's of Arts

© Christine D. Doe

2006



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-18257-4
Our file *Notre référence*
ISBN: 978-0-494-18257-4

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

ABSTRACT

In recent years, there has been a focus on test taking strategies and preparing students for high-stakes exams (Alderson and Hamp-Lyons, 1996); however, it is unclear whether test-takers' increased strategy awareness lends itself to a better test performance. The CAEL, a high-stakes university entrance exam taken by 10,000+ test-takers world wide, a criterion-referenced test, that encompasses the four language skills related to first-year university tasks (Fox, 2003). This study investigated whether test-takers' increased strategy awareness from a preparation course resulted in a desired performance. A mixed method approach comprised of preparation course observations, pre-test semi-structured interviews, including a retrospective verbal protocol, post-test interviews and a post-test questionnaire was used. Overall, it was observed that some test takers may have resorted to non-test-wise strategies while taking the live test, and as the result of taking a preparation course strategy use may be different for high (40-90 total bandscore) and low (10-30 total bandscore) level test takers.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Janna Fox. In the last year, I have learned so much about the field of language assessment, and at the same time been given the opportunity to indulge in my love for research, all of which would not have been possible without you. Thank you for believing in me and encouraging me along this academic road even when the thesis was in its roughest, and possibly completely incoherent, form. You are truly an inspiration.

Second, thank you to my friends and family who I will forever be indebted to. Especially since if it were not for some of you, I am sure this small little beast would still be in the infancy stage. Special thanks to Ann, Annina, Karen and Chrissy the four of you provided support and encouragement at critical points during the completion of my thesis. Thanks to my supportive family - my mom, who would remind me of my strengths whenever I was feeling down, Elaine, whose assertion that I would complete a master's was unwavering, and Debbie, whose understanding and desire for academics was always encouraging. Finally, many thanks to my grandmother, "Mothers", thank you for leading by example as a life-long learner and jet-set traveler – when it was not even the thing to do! Although you may no longer be with us, your memory remains.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
TABLES	vi
FIGURES	vii
ACRONYMS AND ABBREVIATIONS	viii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	6
<i>High-Stakes Testing, the CAEL Assessment and Related Research</i>	6
<i>Washback and Test Preparation Research</i>	10
<i>Learning a Second Language and Taking a Test: The Strategies Used</i>	20
CHAPTER 3: METHODOLOGY	31
<i>Participants</i>	33
<i>Instruments</i>	35
<i>Canadian Academic English Language (CAEL) Assessment</i>	35
<i>Post-test Questionnaire</i>	35
<i>Semi-Structured Pre-test Interview – Students</i>	36
<i>Verbal Report</i>	36
<i>Practice Test – Brain Chemistry</i>	37
<i>Semi-Structured Interview – Instructor</i>	37
<i>Post-Test Interview</i>	37
<i>Procedure</i>	38
<i>Observations of the CAEL Assessment Preparation Course</i>	38
<i>Pre-test Interview</i>	38
<i>Retrospective Verbal Report</i>	39
<i>Post-test Questionnaire</i>	39
<i>Analysis</i>	40
CHAPTER 4: RESULTS AND DISCUSSION	44
1. <i>Preparation Course Observations</i>	44
<i>Overall Description of the Course</i>	45
<i>Observations of Strategies Suggested to the Students</i>	46
2. <i>Pre-Test Interviews</i>	55
<i>Verbal Reports</i>	59
3. <i>Post-Test Interviews</i>	81
4. <i>Post-Test Questionnaires</i>	90
CHAPTER 5: CONCLUSION	104
REFERENCES	112

APPENDICES	123
<i>Appendix I: Description of CAEL Assessment Band Scores</i>	123
<i>Appendix II: Ethics Certificate</i>	124
<i>Appendix III: Post-Test Questionnaire</i>	125
<i>Appendix IV: Interview Questions for Pre-test Interviews</i>	126
<i>Appendix V: Practice Test</i>	127
<i>Appendix VI: Semi-Structured Interview Questions – Instructor</i>	135
<i>Appendix VII: Post-test Interview Questions</i>	136
<i>Appendix VIII: Sample Narrative of Classroom Observations</i>	137
<i>Appendix IX: Script</i>	139
<i>Appendix X: O’Malley and Chamot’s (1990) Description of Strategies</i>	140
<i>Appendix XI: Millman, Bishop, and Ebel’s (1965) Outline of Test-wiseness Principles</i>	142
<i>Appendix XII: Transcription Notation Symbols</i>	144

TABLES

Table 1 Characteristics of Test Takers Interviewed	33
Table 2 Comparison of Typical CAEL Test Takers in this Study.....	34
Table 3 Strategies Discussed in the Preparation Course Grouped According to Skill Tested.....	47
Table 4 Test Taking Strategies Discussed in Class: Language Use versus Test-Wisness Strategies.....	52
Table 7 Test Takers' Bandscores for the CAEL Assessment after the Preparation Course	59
Table 8 Summary of Strategy Use by Test Takers	74
Table 9 Test Taking Strategies reported in the Pre-Test Interviews: Language Use versus Test-Wisness Strategies	78
Table 10 Strategies Discussed in the Post-Test Interviews	81
Table 11 Language Use Strategies Recalled During the Post-Test Interviews	88
Table 12 High and Low Test Takers' Strategy Use who did and did not Take a CAEL Assessment Preparation Course.....	92
Table 13 High and Low Level Preparation and Non-Preparation Test Takers who did and did not Take a Preparation Course.....	97

FIGURES

Figure 1 The Differences in Researchers' Definitions and Classification of Strategies... 22

ACRONYMS AND ABBREVIATIONS

CAEL (ASSESSMENT)	Canadian Academic English Language Assessment
EAP	English in use for Academic Purposes
ESL	English as a Second Language
IELTS	International English Language Testing System
L1	First Language
L2	Second Language
OLT	Oral Language Test
SILL	Strategy Inventory for Language Learning
TOEFL	Test of English as a Foreign Language

CHAPTER 1: INTRODUCTION

Tests that have important decision making consequences for stakeholders are considered high-stakes exams (Cheng, 1997), and the presence of these exams leads to behavior from students that would not have been present if the exams did not exist. This phenomenon is often known as washback (Alderson and Wall, 1993). An example of a high-stakes situation is when someone wants to attend an English speaking university, but does not speak English as a first language. In this case, they will most likely have to obtain a particular score on a university language entrance exam, such as the Test of English as a Foreign Language (TOEFL), International English Language Testing System (IELTS) or Canadian Academic English Language (CAEL) Assessment.

This pressure and demand for students to gain entrance into English speaking universities leads to an increase in coaching and preparation practices, evident by the success of companies like Kaplan Inc. and Princeton Review. In fact, Mehrens and Kaminsky (1989) allude to the fact that the higher the stakes of an exam the stronger the desire to take test preparation practices. As more students feel driven to take these types of courses to ‘just’ pass the test, it is important to understand what students are learning from preparation courses. A component of learning a language is the strategies that aid a student in obtaining their language goals, one of which could be obtaining a high enough score to gain full or partial admission to the university.

In this study, I investigated the role that preparation has on test takers’ strategy use and performance on a high-stakes test. To do so, I explored the strategies that students were introduced to in a preparation course, what strategies they used while completing a practice test, as well as what they reported using during a live test. In

addition, I investigated the strategy use of two proficiency levels in relation to whether the students had taken a preparation course.

This study took place in a specific setting and focused on students who wanted to obtain a particular score on the CAEL Assessment in order to gain entrance into an academic program. To better understand this study it is important to describe the educational setting in more detail.

The Educational Setting

Full-time English support programs differ across the country (Fox, Pychyl, and Zumbo, 1993). In the educational context of the study, the academic language support courses were divided into intensive non-credit English as a Second Language (ESL) courses and credit English for Academic (EAP) concurrent courses.

The intensive ESL courses range from beginner to high intermediate levels with five levels in total. At the two most advanced levels the students can choose a focus: English for personal or professional development. The former are considered pre-credit so that the students will have a seamless transition into the credit courses. All of the courses have a component of English for Academic Purposes. It is important to note that the students in the ESL non-credit courses are unable to have the courses count towards their degrees or to take mainstream academic courses. Therefore, many of the students in the ESL program who intend to continue their studies at the same university are most likely aiming to score at least 40 out of 90 on the CAEL Assessment in order to begin their concurrent studies. Thus, the CAEL Assessment is an exam that acts as a bridge from one program to another - for the students to proceed to the credit EAP program they must first successfully cross the bridge.

The four EAP courses, low, intermediate, advanced and advanced for engineering students, are offered to students who have scored 40, 50 or 60 on the CAEL Assessment. If the students receive a score of 70¹ or higher, they are admitted into mainstream classes without having to take any further English language support courses. Once admitted into the EAP program, the students can take their mainstream credit courses gradually as they progress through the program. For example, if a student receives a 50 on the CAEL Assessment, they will gain entrance into the intermediate EAP course and are allowed to take up to 1.0 credit of academic courses (equivalent to one year-long course).

The CAEL Assessment determines the students' initial course level, but once the students are admitted into the credit program they can progress to the higher levels by obtaining a certain grade or by taking the CAEL Assessment (or another test) again and obtaining the specific bandscore (or score) required to advance or remove the requirement. Furthermore, students are not required to complete every course because those that are capable may move more quickly through the process with the appropriate score on the CAEL Assessment. Each EAP course is one semester long, with the exception of the advanced ESP course that is only offered to engineering students – it is two semesters long.

As is evident above, gaining entrance into the EAP program is highly desirable because the students are permitted to start taking courses for credit that can go towards their degree. Thus, the motivation to do well on the CAEL Assessment is generally quite high for students who are in the ESL program and wishing to gain entrance into the credit

¹ Some graduate programs allow students full admission with a CAEL Assessment score of 60 without requiring any further English language support classes (SLALS, 2005).

program. Students who want to do well on the CAEL Assessment are probably willing to invest time and money by taking a preparation course.

In this study, I investigated the strategies taught in a CAEL Assessment preparation course for ESL students who wanted to gain entrance into the credit EAP program. *Washback* was, therefore, a major component of this study as preparation courses are a direct result of washback. For the purposes of this study, I defined *washback* as testing influencing teaching and learning at the classroom level (Alderson and Wall, 1993) and related *impact* to the effect that testing has on society (McNamara, 2000). The focus was on strategies as a means of looking at what students were doing on the test as a result of taking a preparation course. In this study, I adopted Cohen's (1998) view of *language use strategies* as "steps or actions that learners consciously select in order to accomplish language tasks" (p. 219), and *test-wiseness strategies* as "something that may be dependent on the respondent's knowledge of how to take tests" (p. 219).

This study was an initial inquiry into this research question: What is the role of preparation in test takers' strategy use and test performance? I have further divided the question into four sub-research questions which relate to the four sections of the study:

- 1) What strategies were suggested to the students in a preparation course?
- 2) Did the observed strategies from the preparation course appear in the pre-test interviews with the test takers from the course; and if so, which ones?
- 3) What strategies did the test takers report using from the preparation course and the pre-test interviews in the post-test interview? What was different?

- 4) Was there a difference between the strategy use of high and low level test takers who did take (or did not take) a preparation course?

In the next chapter, I review the literature on washback, strategy use with consideration of the various variables of high-stakes tests, including washback and preparation courses, and strategy related research.

CHAPTER 2: LITERATURE REVIEW

This study is based on and situated within three intertwined concepts of (1) high-stakes proficiency testing, the CAEL Assessment and related research, (2) washback and coaching instruction and (3) learning strategies, including language learning and test taking strategies.

High-Stakes Testing, the CAEL Assessment and Related Research

High-stakes tests as defined by McNamara (2000) are something “where a lot hinges on the determinations made in a language test, for example, where it is used to screen for admission to academic or work settings” (p. 48). In the language testing community, as can be imagined, high-stakes tests are frequently researched to ensure the tests meet the highest possible standards, especially in the last decade with an increased awareness of the importance of ethics (see Davies, 1997). This was the case at the 2006 Language Testing and Research Colloquium (LTRC) conference where ten paper and poster presentations dealt directly with well recognized high-stakes tests (based on the authors’ abstracts), whether the focus was impact, validation, raters’ perceptions, as a basis for university selection, test delivery or test-taker characteristics (Saville, 2006; Luxia & Qinsi, 2006; Vongpumivitch, 2006; Douglas & Hegelheimer, 2006; Liao, 2006; Brown, 2006; Eckes, 2006; O’Loughlin, 2006; and Yu, 2006).

Recently, there has been a considerable amount of research regarding well-known entrance exams, e.g., IELTS and TOEFL. In the former, numerous studies have been conducted from a project initiated at Lancaster University with Alderson and from Cambridge ESOL with Saville, Milanovic and Hawkey (see Saville & Hawkey, 2004; Brown, 2006). In the latter, Educational Testing Service (ETS) has supported major

research projects on the new TOEFL Internet-based Test (iBT) and TOEFL 2000 (Douglas & Hegelheimer, 2005, 2006; Cumming, Grant, Mulcahy-Ernt, & Powers, 2005). Thus, research is a major component of any given high-stakes test.

The CAEL Assessment was developed to test students on the academic language that they will face in their post-secondary studies, but with the recognition of different degrees of proficiency so that students are granted gradual admission into the university while taking the appropriate English language support courses (Fox, Pychyl and Zumbo, 1993). It is a criterion-referenced, fully integrated topic-based test (Fox, 2003) that is used to fulfill the language admission requirements at Canadian universities, and as well as at some American Universities and a few universities across Europe (CAEL Assessment, 2006, May 22). The bandscores range from 10/Non-user to 90/Expert user. (See Appendix I for a full description of CAEL Assessment bandscores).

Some of the tasks on the CAEL Assessment are taking notes while listening to academic lectures, reading an academic text, and writing essays, all of which are related to one unified topic; on a separate day there is an oral component given, which is not related to the unified topic of the main test (Fox, 2003; Fraser and Brisson, 2003). The listening for the CAEL Assessment is embedded within two or three readings followed by the writing. Therefore, the order could look like: two readings, one listening, one reading, one essay, and the Oral Language Test (OLT) either one day before or after the written assessment².

There are four types of questions for the reading section: open-ended, multiple-choice where the test-takers might have to choose more than one answer, fill in the chart,

² In the first administration of the questionnaire, the OLT was 7 days after the written portion due to technical difficulties.

and flow chart questions. The listening questions are similar with the exception of a note-taking question: a space is provided for the test-takers to write point form notes about the topic that the lecturer is discussing. In the writing section, the test-taker is required to incorporate information from the readings and listening to support their essay. They are given the writing prompt at the beginning of the test so they can plan their essay or notice information that might be useful for them when they are completing the other sections. For the information that is used from the readings or listening the test-takers are required to cite where it came from. For example, the test-takers should put (reading 1) or (listening) for information used from the respective sections.

The oral component of the CAEL Assessment is tape-mediated and the responses from the test-takers are recorded. The tasks for the oral section are as follows: talk for a minute about themselves in response to one of the two questions that are given in *The test-taker's guide to the CAEL Assessment* (Brisson and Fraser, 2003) and on the CAEL website, summarize the main points of a lecture, give information about a course from a course outline, read an academic text aloud, and give a short presentation on a topic.

Canadian universities often allow students who speak a first language other than English full admission if the applicants attended a Canadian high school for more than three years, thereby not requiring any English support courses. The logic is that the students are proficient enough to be successful at university, though Fox (2005) revealed that students without English support classes are much less likely to succeed than those who have “timely interventions” (Roessingh & Field, 2000, p. 110) which is when students are provided with a form of English support at the appropriate time in their studies. Further to the point, she came to the conclusion that students who had no English

support courses were more likely to drop their classes altogether. This study provides strong support for the use of the CAEL Assessment as a language admission entrance exam that provides the opportunity for the students to receive academic support, which in the long term will benefit the university.

Non-language testing experts sometimes consider university language entrance exams a predictor of academic performance, though they are only to measure a test takers' English ability. However, Fox's (2004) study revealed that the CAEL Assessment was fairly successful in placing test takers with the appropriate support courses and predicting their success at university. These findings are not surprising when considering two of the participants in her doctoral dissertation: Ofelia who had a high level of English proficiency but possessed a low level of academic expertise, whereas Diana had a low level of English proficiency but a high level of academic expertise (Fox, 2001). With Ofelia's essay the raters felt that she needed more drafts and recognized that she drew on limited resources, while Diana's plan for her essay was excellent and well thought out, she could not express her ideas because of her limited level of English. Therefore, to do well on the CAEL Assessment essay a test taker needs to possess a high level of English proficiency, but also some degree of academic expertise.

In a washback related study, Aminkova (2004) drew on questionnaires, diaries and interviews to examine test takers' perceptions of the CAEL Assessment as a high-stakes test. The author concluded that individual characteristics, language proficiency, reasons for taking the test, perceived performance on the test and motivation for taking the test play into the test takers' perceptions of the CAEL Assessment. In the next

section, I will explore the concept of washback, how it is defined, and what research has been conducted on coaching practices and test preparation courses.

Washback and Test Preparation Research

It is important to discuss the definitions before providing an overview of the literature surrounding the topic of washback. As noted by Cheng (2005) and Cheng, Wantanabe and Curtis (2004) *washback* does not appear in any dictionary, yet it is the term most commonly used by language testers to refer to the phenomenon of test influence on teaching, learning and society. On the other hand, Hughes (2003) stressed that the term *backwash* does appear in the dictionary and remained firm in his use of it over *washback*. Spolsky (1994; as cited in Cheng 1998, 2005) stated that though *backwash* relates to the test impact, the effects are incidental and are not the intended outcomes of the test. Therefore, the term should only be applied when this condition appears.

Commonly, the phenomenon described as washback is the influence that testing has on teaching and learning (Alderson and Wall, 1993; Cheng & Curtis, 2004; Cheng, 1997). This definition of washback is inclusive of both positive and negative washback, often *positive* or *negative* precedes *washback* to denote the type of effect. Other terms are sometimes used to describe a larger concept that encompasses washback. Bachman and Palmer (1996) and McNamara (2000) referred to *test impact* in a broader sense in that tests have an effect at the individual and societal level, where they incorporate *washback* as a facet of the larger phenomenon. Messick (1989, 1995, 1996) describes *consequential validity*, which includes both positive and negative aspects of washback, as

an integral component of validity³, which includes both positive and negative aspects of washback. An example of this is provided in Messick (1995): low scores should not be the result of something that is missing from the main test construct that could possibly enable the test takers to demonstrate their true ability.

In discussing washback more descriptively, researchers will use terms that convey either positive or negative washback. The term *measurement-driven instruction* is based on the idea that tests can narrow the syllabus (Haladyna, Nolan & Haas, 1991; Popham, 1993), which should be seen as being positive in nature and sometimes is referred to as *systemic validity* (Frederiksen & Collins, 1989). *Curriculum alignment* (Shepard, 1990; Cheng, 2005), on the other hand, is generally thought of as negative because it is when testing results in a narrow curriculum: teachers' teaching and the content taught mirror the test (Hamp-Lyons, 1997). In this study, as previously mentioned in the introduction, I use the term *washback* in the sense that testing influences teaching and learning at a classroom level (Alderson and Wall, 1993) and apply the term *impact* to signify the effect that testing has on society (McNamara, 2000).

The presence of washback or impact was understood by many educators and researchers for many years (Bailey, 1996), but the much discussed 'effect' was not explicitly looked at by examining what was going on in the classroom. However, in the late 80's and early 90's washback gained a strong research focus with Alderson and Wall (1993), Wall and Alderson (1993) and Messick's (1989) work and not long after Shohamy, Donitsa-Schmidt and Ferman (1996), Bailey (1996), Wantanabe (1996) and

³ In language testing, validity is a traditionally number-based concept that focuses on the relationship between the test takers' performance and the inferences drawn about their ability in relation to the test or task's criterion (Messick, 1989; McNamara, 2001). In other words, is the ability being described what the test or task intended to measure?

Cheng (1997; 1998) followed suite to develop washback and test impact into the well-researched topic of today as was apparent at the LTRC 2005 and 2006.

Alderson and Wall (1993) challenged the idea that washback was more than a straightforward one-dimensional relationship of tests affecting teaching and learning: to them it was a complex issue with a multitude of possible contributing factors. They commented that various types of test effects should be considered when implementing a new test that is designed to influence the curriculum or when examining the test's impact and thus developed 'Washback Hypotheses' which are as follows:

1. A test will influence teaching
 2. A test will influence learning
 3. A test will influence what teachers teach
 4. A test will influence how teachers teach
 5. A test will influence what learners learn
 6. A test will influence how learners learn
 7. A test will influence the rate and sequence of teaching
 8. A test will influence the rate and sequence of learning
 9. A test will influence the degree and depth of teaching
 10. A test will influence the degree and depth of learning
 11. A test will influence attitudes to the content, method, etc. of teaching and learning
 12. Tests that have important consequences will have washback
 13. Tests that do not have important consequences will have no washback
 14. Tests will have washback on all learners and teachers
 15. Tests will have washback effects for some learners, but not for others
- (Alderson and Wall, 1993, p. 120-121)

They also pointed out that a major failing of most of the washback studies up until that point were their lack of classroom observations.

In their own washback study, Wall and Alderson (1993) used classroom observations and questionnaires to examine and describe the washback that they found after a new test was introduced into the Sri Lankan school system. They observed that in the term that the test was to be given there was an “obvious exam impact on the content of teaching” (p.62) where the majority of the teachers would use material from previous papers or commercial publications instead of the textbook. The teachers that did teach from the textbook would often focus on the last three units of the book, which is what they thought the test would cover. The use of materials other than the textbook was contrary to the administration’s intent since the test was to ensure that the teachers were teaching the textbook. In other words, they wanted to promote positive washback. The authors concluded that the exam had an influence on the content taught but not the method in which it was taught. While Messick (1989) argued for a re-conceptualization of the primarily number-based idea of validity to include a more qualitative look at test scores’ use, and that the consequences of those scores be incorporated into a “unified validity” (p.92).

Furthering the discussion, (Shohamy, Donitsa-Schmidt and Ferman, 1996; Shohamy, 2001) investigated the washback of three types of tests in a centrally controlled education system and found that the status of the language played a major role in the degree of washback or even the presence of it. Cheng (1997; 1998) drew on classroom observations, interviews and questionnaires to investigate whether the intended washback of a public high-stakes exam in Hong Kong occurred. Although the new test was to promote major changes in the syllabus and teachers’ teaching practices, Cheng’s findings were similar to Alderson and Wall’s (1993) study in that the content taught seemed to

change but not the teacher's behavior. She concluded that the role of the publishers in designing materials and teacher training is a factor to consider when looking at washback. Watanabe (1996, 2004) made use of classroom observations and interviews with five teachers to explore whether positive or negative washback was present in a Japanese context. The author concluded the presence of negative and positive washback existed with some teachers, but not all. He even suggested that Alderson and Wall (1993) should modify the 'Washback Hypotheses' to include "a test produces different types of washback to different teachers" as an additional hypothesis (Watanabe, 2004, p. 138).

Bailey (1996) developed a model for washback where she made a distinction between "washback for learners" (p. 263) as opposed to "washback to the program" (p. 264). In the former, she described test-takers and in the latter she grouped together "teachers, administrators, curriculum developers, counselors, etc." (p. 264). By drawing a distinction between the two groups, Bailey stated that Alderson and Wall's (1993) Washback Hypotheses should be clustered accordingly.

In the last decade, numerous washback studies have been conducted, for example, Saville and Hawkey (2004) investigated the role IELTS preparation materials have on stakeholders. They had raters evaluate the IELTS preparation book and the course book according to pre-determined criteria. Saville and Hawkey's findings were still in the initial stages, though they indicated that some of the raters' overall comments were promising. Green and Wall (2005) provided a synthesis of issues from three investigations in a military context in Central and Eastern Europe. They came to the conclusion that key issues surrounding the main high-stakes test were related to the design process, practical problems and political issues, such as a lack of published needs

analyses, invariable test population and little understanding of what goes into to making a valid test, to name a few. In another example of a recent washback study, Luxia (2005) explored whether a high-stakes test reached its intended washback goals after it had been in place for a number of years. Only a few of the intended washback effects were found for a number of reasons, some of which were the students and teacher's beliefs, large test populations, and misuse of scores to evaluate teachers and schools.

A major theme present in many of the washback and impact studies is that tests are not the sole means for reform and countless factors could play a role in determining the existence of washback or impact, such as classroom conditions (Alderson and Wall, 1993), teacher's background (Watantabe, 2004), teacher's style and motivation to change (Alderson and Hamp-Lyons, 1996). Surprisingly, one recurrent factor in the washback studies that does appear is the teacher's perceptions of the test can play a major role (e.g, Luxia, 2004, Alderson and Wall, 1993, Shohamy, 2001). Furthermore, Wall (2005) indicated that introducing new tests or materials are only a "superficial" (p.66) change and that modifying teaching practices requires changing people's beliefs, which according to Kennedy (1988, as cited in Wall, 2005) will only occur if people "change the way they think about certain issues, which is a deeper and more complex change" (p. 66).

An important aspect of washback that relates to this study is that the presence of a high-stakes test could possibly result in behavior from students that would not normally exist, such as taking several practice tests. Moreover, preparation courses for tests could be thought of as a direct result of washback - the classes are for the purpose of trying to

pass the test. Research related to such courses or practices will be addressed in the next few paragraphs.

From a general education perspective, Smith (1990) characterized eight types of test preparation from her comparison of two elementary schools using classroom observations and interviews. The types of preparation were (1) no preparation, (2) teaching test taking skills, (3) encouragement, (4) teaching known content, (5) teaching to the test in format and content, (6) stress inoculation, (7) practicing test or parallel test items and (8) cheating. These varied types of test preparation are not surprising when considering Cheng (1997, 1998) and Wall and Alderson's (1993) findings that the type of washback varies from teacher to teacher which seem to relate to the teachers' belief of teaching and testing.

The controversy of coached versus uncoached test takers, in fact, dates back to the early twenties and thirties when it was observed that an elevated Stanford-Binet score was achievable with practice or coaching (Rugg and Colloton, 1921, as cited in Millman, Bishop, and Ebel, 1965). In a large scale study, Powers and Rock (1999) conducted a survey of 4,200 coached and uncoached test takers for the SAT. They considered the test takers' anxiety, how important it was to obtain a good score, what colleges were their first choices, average grades, etc. It was found that some of coached test-takers demonstrated a dramatic score increase (12 % of 427 improved by 100 points on the verbal portion), however, not all coached students showed an increase (36 % of 427 scores decreased or remained exactly the same) on their retests. In addition, some of the uncoached students exhibited a remarkable score increase (one test taker increased their

verbal score by 150 points). Interestingly, this one student studied the SAT preparation book and other test preparation material.

As was demonstrated in some of the washback research, specific preparation varies from student to student regardless if they are coached or uncoached. Powers and Rock's findings do demonstrate a slight increase in coached students' scores, but they are somewhat surprising considering the statement made from Kaplan, a test preparation school, explaining why the student should study with them for the SAT:

Confidence is one of the key elements to succeeding on standardized tests, and it all starts with your confidence in our ability to help you achieve your goals. That's why we back our programs with The Kaplan Higher Score Guarantee—the most comprehensive guarantee in the industry...If for any reason your score doesn't increase, you can study with us again for free or get your money back.” (Kaplan, 2006, ¶ 2)

If Powers and Rock's (1999) findings are transferable, how can Kaplan offer students their money back or a free course and not be bankrupt considering that the students' scores increasing ultimately depends on the student? That being said, Kaplan does not say anything about first time test takers or how much the students' scores have to increase, thereby safeguarding itself from possibly losing a substantial amount of money. While Powers and Rock admitted that some of the students might not have been as motivated as if they were taking a test with real stakes, they insisted that there is a discrepancy between their findings and the promises made by coaching companies, such as Kaplan and Princeton Review.

In addition to examining students' success on the SAT, Powers and Rock (1999) presented characteristics of test takers who took preparation courses vs. those who did not. Overall, they concluded that coached students were often Asian American, from well-educated parents, had high grades, more years learning foreign languages, came

from affluent families, and had prepared for the SAT in a variety of other ways. Interestingly, 32 percent of the uncoached students felt that their recent SAT scores were “pretty good estimates of my abilities” (p. 107) compared to 20 percent of the coached students, on the other hand, the coached students placed more importance on receiving good scores (52% vs. 40%). Therefore, it seems likely that the tests were perceived by the coached students as having higher stakes, which is why 13 percent of them reported feeling “extremely nervous” (p. 107) when taking their most recent SAT I compared to 8 % percent of the uncoached students.

In a washback related study, Alderson and Hamp-Lyons (1996), investigated whether ESL students in TOEFL preparation courses were exposed to unnatural teaching or inappropriate language learning strategies. The authors observed that most of the teachers taught the book with two exceptions: Teacher 1 whose preparation book was being published, had taught the preparation course for 17 years using his own materials, and an inexperienced teacher, Teacher 2, used interactive activities with students. The observation data of two teachers teaching a preparation course and a non-preparation course revealed an obvious difference between the two types of courses: test-taking was much more frequent, the teachers talked more in class, implemented less pairwork, used more metalanguage per lesson, and less laughter events per lesson in the preparation course compared to the regular course. The authors suggested that some of the differences could be due to the large class size but allude to the fact that more research is needed to better explain the differences. Although there seemed to be a disparity between the two types of courses, the two teachers also differed in how they taught the TOEFL preparation course. The authors concluded the effect of tests varies from teacher

to teacher, and that the test is not solely responsible for how much washback results: “It is the administrators...materials writers...and teachers themselves ...who caused the washback [they]... observed” (p. 295).

The concept of teaching to the test or the overuse of test practice materials might undermine the test’s construct, although Messick (1996) pointed out that it would only affect the validity of the coached scores:

A validated proficiency test can be subverted by test preparation practices or coaching emphasizing testwiseness strategies that might increase test scores without correspondingly improving the skills measured by the test. Although this would not compromise the validity of the uncoached test in general, the validity of the interpretation and use of the coached scores would be jeopardized. (p. 246)

Moreover, Messick (1982) suggested three possibilities of how test scores could be improved from coaching: 1) reflective of improved abilities, 2) enhancement of “test-taking sophistication” (p. 81) or reduction of test anxiety, 3) or test taking strategies. According to Messick, the last factor is the one that would result in test scores that are not an accurate judgment of the student’s ability. This would “not only dilute the construct validity but jeopardize its predicative validity as well” (p.81). He also felt that the other two possibilities would be positive for both the student and from “the standpoint of test validity”. Despite the fact that test takers using test taking strategies would be negative, in his opinion, it would probably be slight with a well-developed test.

The teaching of strategies in a preparation course is most likely to occur in test preparation courses, as Smith (1990) listed it as one of the types of preparation she observed. In the final section, I will provide an overview of test taking and learning strategies’ definitions and discuss some of the studies conducted on strategy use and language testing.

Learning a Second Language and Taking a Test: The Strategies Used

Strategies are intangible processes which involve inferring what students are doing or relying on students' accounts of what they are thinking. Therefore, before discussing strategy definitions further it is imperative to consider the common methods employed in strategy investigation. The most common form of strategy investigation is the use of questionnaires. A number of studies draw on existing questionnaires, such as the *Strategy Inventory for Language Learning (SILL)* developed by Oxford (1990) as seen in Cohen, Weaver and Li, (1998); Oxford (1990); Wharton (2000) and Al Naddabi (2006). Other studies have based questionnaires on tasks done by learners (see Chamot, El-Dinary, and Robbins, 1999; Kojic-Sabo and Lightbrown, 1999; Fan, 2003). A key advantage to questionnaires is the ability to administer them to large groups and the disadvantage is the uniformity of the questions (Cohen, 1998).

Another commonly used method is verbal reports where the learner describes his or her thoughts while completing a task. Chamot (2005) noted that such procedures allow insights into online processing – not features of planning or evaluating (metacognitive) (see Cohen et al., 1998; O'Malley and Chamot, 1990). This type of method brings forth a number of questions regarding its usefulness, since the learner's memory capacity will not be able to remember everything nor may the test takers be completely honest of why they choose a specific answer. Nevertheless, verbal reports are one of the most accurate predictors at our disposal (Lieberman, 1979, as cited in Cohen, 1998).

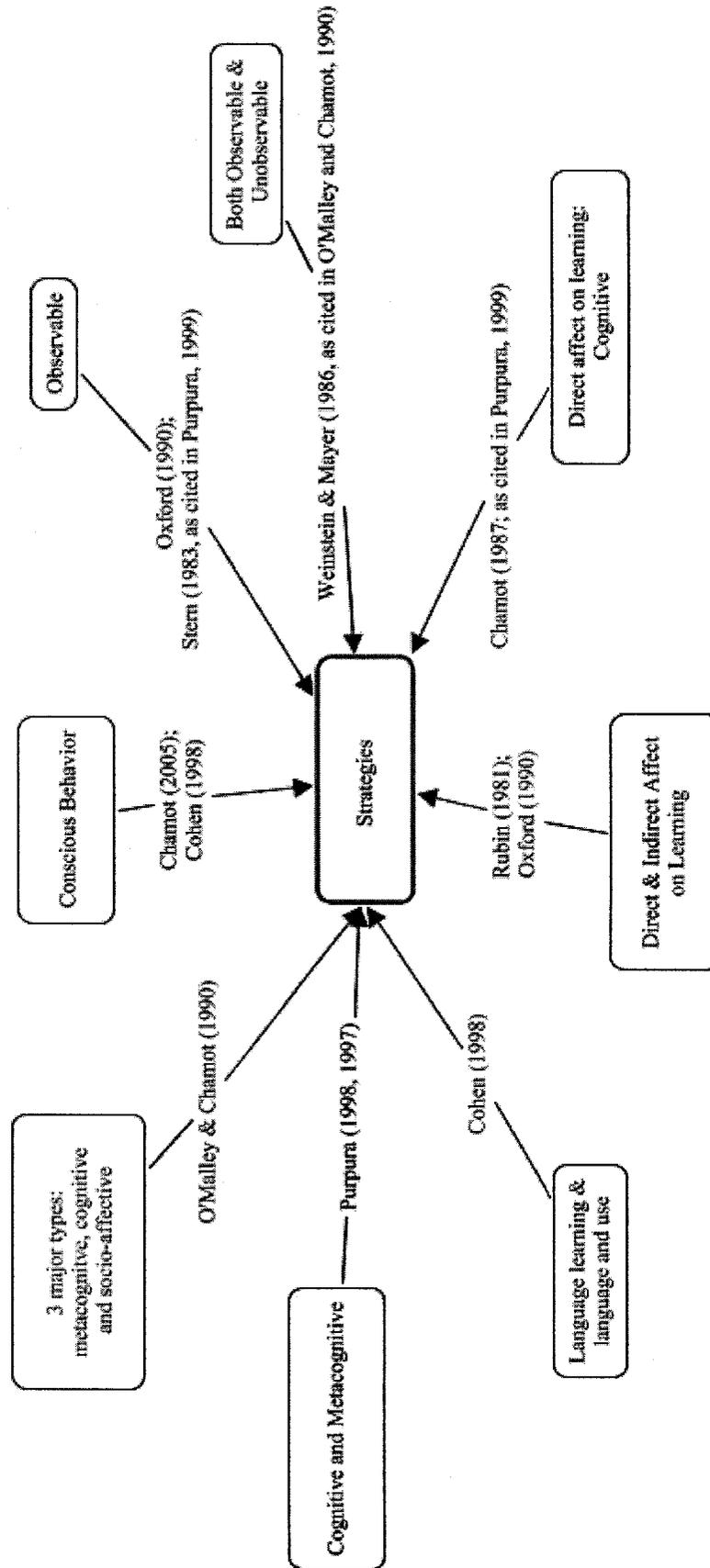
An additional method used in strategy studies are interviews and as Chamot (2005) stated, the most effective form of interviews for strategy investigation are stimulated recall, especially when they occur immediately after the learners complete a

task. In the literature, classroom observations are the one form of research tools that are underrepresented in strategy studies. Chamot (2005) commented that the lack of classroom research may very well be due to the inability to control variables, like motivation, when conducting classroom research.

Strategies are no different from any other facet of the complex field of linguistics, they are multifarious. This complexity is partly due to the nature of strategies – they are a mental process, and researchers are unable, legally and ethically, to open up the black box of the mind to see exactly what strategies are. Therefore, strategy research should encompass multiple methods to ensure the best possible understanding of strategies is obtained. In the next few paragraphs I will explore the multitude of ways strategies are described within the applied linguistics and language testing community.

To begin with, one of the researchers on the forefront of second language strategy inquiry, Rubin (1975), identified good language learners and then classified the strategies that they most commonly used. In this classification she viewed strategies as “the techniques or devices which a learner uses to acquire knowledge” (p. 43). Since her seminal paper, strategy classification and definitions have become increasingly complex. The 2001 *Oxford Dictionary of Current English* defines a strategy as “a plan designed to achieve a particular long-term aim” (p. 900). With such a broad definition it is no surprise that the word *strategy* can be defined in so many ways. Figure 1 only a small selection of the different ways in which strategies are defined in the literature. The figure was initially adapted from Purpura’s (1999) discussion of the variance in the classification, definitions, and models of strategies.

Figure 1. The Differences in Researchers' Definitions and Classification of Strategies



In developing a strategy classification system, O'Malley and Chamot view strategies as cognitive, metacognitive and social/affective grounded in Anderson's (1983, 1985, as cited in O'Malley and Chamot, 1990) three stage theory of skill acquisition: the cognitive, associative and autonomous stages. Oxford (1990), similar to O'Malley and Chamot, categorizes strategies as cognitive, metacognitive, affective or social, but she further divides them into six categories in her *Strategy Inventory for Language Learning* (SILL): cognitive, metacognitive, social, memory-related, general cognitive, and compensatory. According to Oxford, learners are real people who are intellectual, social and emotional and can not be categorized simply. From a language testing perspective, Bachman and Palmer (1996) view strategic competence as an essential component to test taking, as a set of metacognitive strategies (setting what goals are needed, assessing what is needed, planning what is needed to be done) that aids in higher processing, but is not limited to language ability.

Though the previous authors listed define strategies from a mental processing perspective and make a distinction between metacognitive and cognitive, Cohen (1998) groups test taking strategies into *language use strategies* and *test wiseness strategies*. He defines *language use strategies* as "steps or actions that learners consciously select in order to accomplish language tasks (p. 219) and *test-wiseness strategies* as "something that may be dependent on the respondent's knowledge of how to take tests" (p. 219). For the purposes of this study, I adopt Cohen's notion of test taking strategies as encompassing both language use strategies and test-wiseness strategies. The specific approach taken in the study is described in detail in the analysis section in Chapter 3.

Considering strategies from a metacognitive and cognitive perspective, Purpura (1997, 1998, 1999) presents his findings from his doctoral dissertation on learning strategies, based on a human processing model, compared with test-taker's performance on second or foreign language tests. The Structural Equation Modeling (SEM) results revealed that metacognitive strategies had a strong impact on cognitive strategies but not a direct effect on the test performance. Purpura noticed a difference between the strategies of high and low level test-takers: 3 cognitive (inferring, linking with prior knowledge, and naturalistically) and 2 metacognitive (self-evaluating and monitoring) strategies were used by more high-level learners. In addition, the self-evaluating question demonstrated a direct relationship with the grammar, sentence formation, and cloze questions, but the low level group had no such relationship although the test takers were reminded to check their work throughout the test. In addition, Wantanabe (1990, as cited in Oxford & Burry-Stock, 1995) reported students who rated their own proficiency higher used more SILL strategies than those that rated their proficiency lower.

In an early strategy study on the test taking process, Nevo (1989) had test takers respond on an item by item basis with a 15 strategy checklist after each question. Although the test takers transferred strategies from the L1 to the L2, the results indicated that strategy choice was less likely to correlate to the correct choice. That being said, Phakiti (2003) concluded that successful test takers reported using strategies more than unsuccessful test takers. Nankatani (2005) investigated the effects of training on learners' oral proficiency scores. The findings indicated that training significantly increased the learners' scores through increased strategy use, maintenance and negotiation of meaning

strategies, but also reduced the number of reduction or abandonment strategies (strategies where the learners would sometimes stop speaking in mid-sentence).

There has always been the question of what background knowledge test takers bring to a test and how that may affect their performance and strategy use. Afflerbach (1990) examined the effect that prior knowledge had on reading strategies. The results demonstrated that experienced readers with prior knowledge were able to automatically construct the main idea of the text more readily than readers without any prior knowledge. On the other hand, expert readers with limited content knowledge drew on draft and revision strategies, thus, they broke an unfamiliar text into more manageable subtasks. In their study of pre-teaching background knowledge of the test topic to high and low proficiency language learners, Fox, Pychyl and Zumbo (1997) concluded that pre-teaching material to low proficiency language learners did not affect their performance on the CAEL Assessment. High level language learners, in contrast, responded to the pre-teaching and modified their answers accordingly. The authors determined that the difference was most likely due to their high level language learners' ability to retain the macro-structure from the pre-teaching thereby freeing resources for micro-level processing by being able to attend to specific vocabulary.

Inquiry into strategy use is important for the purpose of test validation. Anderson (1989, as cited in Cohen, 1998, 1991) observed that strategies were used in different ways depending on the types of questions asked. Tsagari (1994) also found a test method effect for the strategies used: different behaviors were performed in two tests with different formats, but identical content. Consequently, it seems that different test formats

or question types elicit different behaviors from test takers other than the intended ones by the test constructor.

In summary of what I have discussed thus far, language use strategies may vary by proficiency level (Purpura, 1997; Wantanabe, 1990; as cited in Oxford and Burry-Stock, 1995), background knowledge (Afflerbach, 1990; Fox et al., 1997), cultural background and test format (Anderson, 1991; Tsagari, 1994). In addition, it was seen that training can have a positive affect on oral communication (Nakatani, 2005), and successful test takers report using strategies more than unsuccessful test takers (Phakiti, 2003). These studies were based on language use strategies which logically require some level of language ability. In contrast, test wiseness strategies may be independent of language ability. Furthermore, in the literature there is no distinction made between test-wisness strategies used by a second language learner or those used by someone in their first language, namely, test-wisness strategies are never referred to as language test-wisness strategies.

The concept of test wiseness as previously discussed relates to the test taker's knowledge of taking a test or as Millman et al. (1965) describes it as "a subject's capacity to utilize the characteristics and formats of the test and/or the test taking situation to receive a high score" (p. 707). The authors expanded on the notion of test-wisness by constructing a taxonomy of principles that were separated into two key groupings: strategies independent of the test constructor and those dependent on the test constructor. (See Methods Section) Though Millman et al. (1965) state that the test-wise principles are independent of the test taker's knowledge of the subject matter being tested, they

acknowledge that in order for test takers to effectively apply the deductive reasoning strategies the test takers must apply some partial knowledge.

Since their seminal paper, numerous studies have been conducted using the test-wisness principles (see Benson, 1985, as cited in Rogers and Bateson; Sarnacki, 1979). Rogers and Bateson (1991), in a general education study on a high-stakes high school leaving exam, adopted Millman et al.'s (1965) principles but concluded that partial knowledge and timely application of test-wisness strategies, deductive reasoning and cue using strategies, must coincide to increase the likelihood of answering an item correctly. For their study they devised a Test of Test-Wisness (TTW) to assess students' ability apply test-wisness strategies to questions. Yang (2000) administered to TOEFL test takers to determine the role of test-wise knowledge on test takers' performance. She separated them into two groups according to their test-wisness awareness (test-wise and test-naive) and had them give verbal reports of items that would be susceptible to such strategies. The researcher indicated that test takers who were more knowledgeable in test-wisness strategies approached the items in a more logical and thoughtful manner, were more academically knowledgeable, and used that knowledge to their advantage.

The partial knowledge discussed in Rogers and Bateson's (1991) and Yang's (2000) studies do not refer to language knowledge but knowledge of the subject matter of the test. However, there have been studies that have found positive correlations between the successful application of test-wisness strategies and verbal ability and only a slight correlation with intelligence (Diamond and Evans, 1972). Despite the logical assumption that test-wisness and test anxiety would be negatively correlated, Millman (1966, as cited in Rogers and Bateson, 1991) found no such relationship.

The majority of the studies that have been conducted using Millman et al.'s (1965) taxonomy of test-wiseness were on multiple choice tests (Rogers and Bateson, 1991; Yang, 2000). Indeed the examples and discussion that Millman et al. provides as an explanation of the taxonomy are focused on either multiple choice or true or false questions. Sarnacki (1979) argued that time-using and error avoidance strategies are applicable to more objective based tests as well, such as essay examinations.

Cohen (in press) mentions that there is limited research on strategy or coaching instruction for high-stakes tests. That being said, there have been a few studies that looked at strategy use for the IELTS and TOEFL. One such study is Griffiths and Jordan's (2005) study that examined preparation course students' strategy use and scores drawing on interview and questionnaire data. The authors came to the conclusion that more strategic students performed better on the IELTS than less strategic students. Although the authors investigated preparation students' strategy use, they failed to conduct any classroom research.

Finally Tian (2000) investigated Taiwanese students' reading behavior for TOEFL reading comprehension questions through a think aloud protocol (that was coded to build a taxonomy of items), having the students write down what they remembered from the passage, and interviews focused on the students' test preparation and perception of the training. Tian constructed a 42 item taxonomy from the verbal protocol that was split into three categories (technical, reasoning, and self-adjustment strategies). Overall the researcher found that low level test takers followed the coaching instruction, skip reading the passage and read the questions first, relied on "key word matching" (p. 215) to find the information and focused on "word-level reasoning processes" (p. 215) more

than high level test takers. High level scorers demonstrated more repair strategies, emphasized their need to improve their comprehension and personalize their strategies. As a result of the findings Tian recommends that instructors with an emphasis on reading comprehension should focus on comprehension and teach strategies as support for general comprehension.

As was discussed at the beginning of this chapter, this study focused on three intertwined concepts of high-stakes proficiency testing, washback and coaching instruction, and learning strategies that included language learning and test taking strategies. The literatures discussed in this chapter relate to either the guiding research question or more specifically to one of the sub-research questions (or both). The main research question for this study was: What is the role of preparation in test takers' strategy use and test performance?

The four sub-research questions were:

- 1) What strategies were suggested to the students in a preparation course?
- 2) Did the observed strategies from the preparation course appear in the pre-test interviews with the test takers from the course; and if so, which ones?
- 3) What strategies did the test takers report using from the preparation course and the pre-test interviews in the post-test interview? What was different?
- 4) Was there a difference between the strategy use of high and low level test takers who did take (or did not take) a preparation course?

As I mentioned at the beginning of this chapter, research is meaningless without relating it to the external world by describing its context. It is also just as important to describe how the study proceeded in as much detail as possible. After all, every reader brings their own interpretation and cultural background to the meaning of words and phrases (Bahktin, as cited in Lemke 1995). Therefore, in the next chapter I describe in detail the methods used in conducting this study.

CHAPTER 3: METHODOLOGY

No single truth is ever sufficient, because the world is complex. Any truth, separated from its complementary truth, is a half truth. (Myers and Spencer, 2001, p. 74, referring to Blaise Pascal's philosophy)

Through the use of qualitative and quantitative measures, I acquired a deeper and more comprehensive understanding of test takers' strategy use and the role of preparation in completing the CAEL Assessment, which would not have been possible if I had only approached the study with one ontological view. This is the case with Li, Marquart, and Zercher's study (2000, as cited in Tashakkori and Teddlie, 2003) study:

The cross-over tracks approach started with two sets of data collected for the purpose of a holistic understanding of a specific programme. The two tracks crossed each other first at the data transformation stage where some of the quantitative data was turned into qualitative data and vice versa. During data comparison, the focus of analysis shifted back and forth between the two data sets. The process served well the purposes of triangulation, complementarity, and initiation of mixed-method studies. (p. 70)

Or as Bachman (2006a) states, “[a] given research study should be based not on doctrinaire or philosophical positions, but rather on the consideration of which approach is the most appropriate and likely to yield the most insights into the research question” (p. 201).

The mixed method approach (see Fox, 2003) employed in this study allowed me to explore my research questions from the best vantage points. To explain in more detail, the overall research question was answered by drawing on qualitative (classroom observations, pre-test interviews, and post-test interviews) and quantitative (questionnaires) forms of data collection. According to Tashakkori and Teddlie (1998) this mixed-method approach demonstrates “methodological triangulation” (p. 18). Furthermore, the different forms of qualitative measures provided a means of “data

triangulation” (Tashakkori and Teddlie, 1998, p. 41) which Denzin (1978, as cited in Tashakkori and Teddlie, 1998) defined as “the use of a variety of data sources in a study” (p. 41). Cohen (1998) drew attention to the importance of multiple data sources for any form of strategy research, especially those that draw on verbal reports, to be able to draw stronger inferences of the test takers’ mental processes. In addition to data triangulation, the qualitative portion of the study was exploratory in nature because the tracking aspect of the pre- and post-test interviews in relation to the strategies discussed in the classroom observations was not found in the literature. Finally, since the study was driven by my research questions and not any particular paradigm (Tashakkori and Teddlie, 2003), I adopted a pragmatist approach (Bachman, 2006a; 2006b; Jamieson, 2006).

In Chapter One, I presented the basis for my study and identified the research questions I explored in this study. In Chapter Two, I reviewed the literature that relate to these questions, and in this chapter I describe the methods used to answer the questions. In this chapter I explore the methods used to investigate the four sub-research questions. This study was comprised of four integrated components of data collection: 1) observations of two CAEL Assessment preparation course sections; 2) two sets of interviews with the preparation course students: specifically five pre-test interviews with a retrospective verbal report component (Cohen, 1998) and three stimulated recall post-test interviews (Chamot, 2005); 3) one interview with the course instructor; 4) 270 post-test questionnaires, developed from a taxonomy of strategies suggested to the students in the preparation course, and administered to test takers after they had completed the CAEL Assessment.

Participants

Five students, Charles, Ben, Alexandra, Edward, and Donald⁴, from the CAEL Assessment preparation course were observed in a CAEL Assessment preparation course and interviewed four to five days before the CAEL Assessment on December 3, 2005⁵. Charles, Ben and Alexandra were interviewed again three to five days after the live CAEL Assessment on December 3, 2005. Table 1 indicates the participants' first language, country of origin, and level of ESL study.

Table 1 Characteristics of Test Takers Interviewed

Name	Country of Origin	First Language	Level of ESL Study
Charles	Mainland China	Mandarin	ESL - Intermediate
Ben	Turkey	Turkish	ESL – Intermediate
Alexandra	Saudi Arabia	Arabic	ESL – Intermediate
Edward	Jordan	Arabic	ESL - Low Intermediate
Donald	United Arab Emirates	Arabic	ESL - Intermediate

Sam, the course instructor, was interviewed ten days after the CAEL Assessment preparation course had finished. At the time of the study, he had been teaching for nine years: three years with the Canadian Army, three years outside of Canada (EFL), and three years at the university (ESL) where this study took place. Two months before the beginning of this study, Sam had completed his Master's degree in Applied Language Studies. This was the second time that Sam had taught the CAEL Assessment preparation course; the first time was a six week course just prior to the one in this study.

The students from the CAEL Assessment preparation course were required to bring *The CAEL Assessment Test Taker's Preparation Guide and Reading Booklet*

⁴ Fictitious names are used for the participants to ensure anonymity

⁵ This study was approved by the Carleton Ethics Committee (See Appendix II for the ethics certificate).

(Brisson and Fraser, 2003) to class. The course was offered and advertised for students who had completed ESL Intermediate or above in the ESL program at the university; however one of the students was at a low intermediate level but, because of his high speaking level, he was allowed to take the course.

On November 19, 2005, the CAEL Assessment was administered to 76 participants and 51 of these completed the post-test questionnaire. 215 test takers took the CAEL Assessment on December 3, 2005 and 120 of these completed the post-test questionnaire. A comparative summary of the CAEL test takers in this study to average CAEL test takers is given in Table 2.

Table 2 Comparison of Typical CAEL Test Takers in this Study

Characteristics	Typical CAEL Test Takers ⁶	CAEL Test Takers in this study (n=170)
Male/Females (%)	59.8/40.2	52.7/47.3
Mean/Median Age	26.6/25	26.1/20
<u>Most Common First Languages: (%)</u>		
Arabic	19	14.3
Cantonese	12.9	1.8 ⁷
Mandarin	11.4	3.1
Korean	5.6	1.3
<u>Area of Intended Study:</u>		
Engineering	30	21.7
Arts and Social Sciences	25	25.2
Sciences, Mathematics and Statistics	8.2	7.9
Computer Science	21.4	8.6
Business	7.5	35.7

⁶ The characteristics reported here are from Fox (2002).

⁷ The low numbers can be explained by the fact that 46.5% of the test takers responded that their first language was Chinese, but it was unknown whether they meant Mandarin or Cantonese.

There were four major differences between the test taker population that was sampled in this study using the questionnaire in comparison to the typical test taker characteristics (Fox, 2002): (1) a higher percentage of females, (2) the female test takers were younger by five years, (3) fewer Arabic and Korean speakers, (4) a higher number of Chinese speakers, and (5) fewer test takers who intended to study Engineering, and an increased number of students who planned on studying Business. Therefore, the test takers in this study may not have been representative of the typical CAEL test taker population, however; the data collected for the *Test Score and User's Guide* was from 2002 and may be out of date. Another possible explanation is that the data was not representative of the overall test taker population, but it may have been representative of students taking the test at that time of year. The questionnaires were administered at the end of the Fall semester, so the test takers may have taken the test to satisfy the language requirements for full admission or concurrent study that would have begun the following January. Therefore, these test takers may have had different motivations for taking the test from the participants taking the test during the semester.

Instruments

Canadian Academic English Language (CAEL) Assessment

A full description of the CAEL Assessment is given in Chapter 2 under the subheading *High-stakes tests and CAEL Assessment related research* (See pages 7-9).

Post-test Questionnaire

The questionnaire was divided into four sections. (See Appendix III for the post-test questionnaire) The first section allowed for the placement of the test takers' scores on the CAEL Assessment. The second section had questions related to the test takers'

background information. The third section titled *What did you do to prepare for the CAEL Assessment* covers the preparation or experience the test takers had prior to taking the test. The questions required yes or no responses. The questions developed for the final section were from the taxonomy of strategies observed in the preparation course, except for the first two questions. The section was titled *What did you do while completing the CAEL Assessment?* The purpose of this section was to investigate what strategies the students employed during the CAEL Assessment.

There were no questions about the OLT on the questionnaire, because the OLT was administered on a separate day from the main test date when the questionnaire was given.

Semi-Structured Pre-test Interview – Students

There were two phases for this set of interviews. In the beginning of the interview, I asked the test takers/students⁸ background questions, for example, *where are you from?* and *how long have you studied English?* The nature of the semi-structured interview allowed me to ask additional questions that developed throughout the course of the interviews. (See Appendix IV for a complete list of the questions used in this portion of the interview)

Verbal Report

The second half of the semi-structured interview made use of verbal reports (Cohen, 1998) and focused on the strategies the students reported using while they completed the *Brain Chemistry* practice CAEL Assessment (Fraser and Brisson, 2003).

⁸ The term *students/test takers* is used to describe the participants I interviewed since they were students in the preparation course, but were also CAEL test takers.

Practice Test – Brain Chemistry

The practice test, *Brain Chemistry*, which is the last practice test in *The Test taker's guide to the CAEL Assessment* (Fraser and Brisson, 2003), was comprised of 2 readings (reading 1 and 2), 1 listening, and 1 writing section(s). Due to time constraints and the students' fatigue, I used a shortened version of the practice test: reading 2⁹ (10 questions), listening (10 questions), brainstorming portion of the writing, and three questions¹⁰ from the oral. (See Appendix V for the practice test used.)

Semi-Structured Interview – Instructor

The purpose of the interview with the course instructor was to better understand his approach to the preparation course. The questions asked in the interview covered the following topics: how long have you been teaching, what kind of classes have you taught, have you taught the CAEL Assessment preparation course before, do you think teaching strategies to the students is useful, etc. (See Appendix VI For the complete list of the questions)

Post-Test Interview

With the post-test interview, I investigated the types of strategies the students reported using during the live CAEL Assessment. To elicit the best responses, I used a stimulated recall method (see Cohen, 1998, Chamot, 2005) by asking them to remember specific times throughout the day. In beginning of the interview, I asked questions such as: where were you sitting, what was the topic of the test, what was the order of the readings and listening, etc. Then, I asked if they could remember answering the questions

⁹ Reading 2 was chosen because of the subtitles in the text. Since one of the strategies suggested to the students in the preparation course involved subtitles, I wanted to include a reading that had subtitles to see if the test takers drew on subtitles to predict the reading as was suggested to them.

¹⁰ The three questions for the oral were chosen because those were the three questions that were discussed in the preparation course.

and, if so, how they answered them. (See Appendix VII for the complete list of questions)

Procedure

This study consisted of five stages: (1) classroom observations, (2) a pre-test interview with students/test takers, (3) a post-test questionnaire given to test takers, (4) a semi-structured interview with the course instructor, and (5) a post-test interviews with students/test takers. Data collection began on November 8, 2005 and was concluded on December 13, 2005.

Observations of the CAEL Assessment Preparation Course

Twelve classes, six for each section (A and B), were observed between the dates of November 8, 2005 and November 30, 2005. For reasons beyond my control, I was unable to observe the first two lessons and two in the middle of the observations (one in each section). I asked the instructor to let me know what was covered for the days I had missed. For each lesson I collected detailed field notes, kept copies of all handouts given and wrote in my notes which handout was given on which day. Afterwards, I used the field notes to write a narrative of each lesson with a summary of the strategies discussed in class at the end of the notes. (See Appendix VIII for an example of a written narrative of one the days observed)

Pre-test Interview

I interviewed five students a week before they took the CAEL Assessment to see what strategies they reported using while completing a CAEL Assessment practice test or why they answered a question in a particular way. Phase 1 of the interview focused on questions regarding the participants' background information. In Phase 2, I gave the

participants a practice CAEL Assessment and had them do a retrospective verbal protocol. Procedures regarding the verbal reports are described in more detail below. This set of interviews ranged from 60 to 75 minutes.

Retrospective Verbal Report

Before starting the verbal report section of the interview, I told the students that they were going to complete the various sections of the CAEL Assessment and afterwards they would tell me what they had done while completing the sections. If I observed some strategy use that they did not mention in their accounts, I would ask them questions about it. For example, if one of the students underlined a word or re-read a question a number of times, I inquired what it was that they were doing.

The CAEL Assessment practice test components were given in the following order: listening, reading, writing and oral. All the interviews were recorded using a digital mp3 player. By including an interview prior to the students' live CAEL Assessment, I acknowledge that there was most likely a researcher effect on the students' performance for the live CAEL Assessment.

Post-test Questionnaire

A questionnaire was administered to all CAEL Assessment test takers after they finished the test. The test administrators were given a script that was read to the test takers. The purpose of the script was to ensure consistent administration of the questionnaires. (See Appendix IX for the script)

The students were also given an informed consent form that required their permission for their scores to be assigned to the questionnaire after the results had been released. On November 19, 2005, there were two sets of questionnaires distributed to the

test takers in alternating stacks of ten. For the first set of questionnaires, the informed consent form was on the bottom and the questionnaire was on the top and for the second set it was the reverse¹¹. There was no significant difference whether the consent form was on the top or the bottom for the test takers' responses; therefore, on December 3, 2005, the consent form was placed on top of the questionnaires. I believed that placing the questionnaire on top would ensure the participants were well informed about the questionnaire's purpose in advance of completing the questionnaire.

Post-test Interview

I began the interview by asking three of the test takers to remember the time when they took the written portion of the CAEL Assessment and the OLT as discussed above. After they discussed the particular day, the topic of the test and its layout, I asked them more specific questions. All the interviews were recorded using a digital mp3 player and ranged from forty-five to sixty minutes in length.

Analysis

The field notes were written up with a summary of the strategies at the end of the narratives. Midway through the observations, the summaries were compiled into a taxonomy of strategies which formed the basis for the questionnaire development and upon completion of the observations, the taxonomy was finalized. This final taxonomy provided the basis for the analytical deduction (Becker, 1958, 1998) conducted over the three forms of qualitative measures used in this study. According to Becker the process of analytical deduction begins with the researcher observing the situation and organizing the problems or concepts in a systematic manner and then suggesting a theory based on those

¹¹ In the first administration of the questionnaire, it was unknown whether the test takers would respond differently if they were presented with the consent form before or after the questionnaire.

observations. This first theoretical model forms the foundation to explain the phenomenon. Once the initial model is constructed, the researcher must apply the theory to new situations where problems or concepts may exist that were not present in the initial setting and he or she will have to decide whether or not to add the new components to the theory or to revise it.

Therefore, this type of theory development is one where the researcher continually refines their model based on multiple forms of sequential evidence and takes into account the indicators that are contrary to those originally observed, which may force modifications on the newly established theory. The final stage of the process consists of rechecking the model for accuracy against transcriptions or other forms of existing evidence. Also Becker (1998) suggests that the most beneficial type of data collection for this style of model development may be one of an unthinking nature, where the researcher does not consider the concepts or hypotheses that guide the study.

In relation to my study, the classroom observations were the initial stage in Becker's (1998) description of analytic deduction. The pre- and post-test interviews provided additional forms of evidence for confirming or adjusting the taxonomy. Moreover, the process of developing the taxonomy and applying it to other forms of data collection presented a method for developing a classification system that was situated in this particular context and related to the CAEL Assessment. In doing so, the test takers' strategy use could be tracked through a realistic lens that is based on the setting being studied. Becker's (1958) notion of the researcher in this process is that of a participant observer where the researcher "gathers data by participating in the daily life of the group or organization he studies...[and] enters into conversation with some or all of the

participants in these situations and discovers their interpretations of the events he has observed” (p. 652). However, in my study I did not view myself as a participant observer for the data collection. In fact, I made a conscious effort not to guide the test-takers’ responses during the pre- and post-test interviews by asking open-ended questions and maintained a certain distance to the test takers guided by a ‘professional researcher’ persona – rather than a trusted friend.

After the taxonomy was compiled according to skill, I then categorized all language use strategies according O’Malley and Chamot’s (1990) framework. (For a complete list of the descriptions used to categorize language use strategies see Appendix X). In addition, all of the test-wiseness strategies were categorized according to Millman et al.’s (1965) framework. (For a complete list of the descriptions used to categorize test-wiseness strategies see Appendix XI) The digital recordings with the instructor and the pre-test interviews with the test takers were transcribed using the coding scheme made available in Lazaraton and Taylor (in press). (See Appendix XII for an example of the transcription coding scheme)

The pre-test interview transcriptions were analyzed through the ongoing process of analytic deduction (Becker, 1998) based on the pre-defined strategies from the previously compiled taxonomy of strategies. Then, the summary of the strategies used by the test takers in the pre-test interviews were summarized in a chart format, which formed the basis for coding the post-test interviews. Half of Charles and the first fifteen minutes of Alexandra’s interviews were coded by the researcher and another rater to ensure reliability using the chart in addition to those strategies that were not on the chart. The

inter-rater reliability was calculated using Pearson's Correlation coefficient ($r = .98$, $p < .01$).

The post-test questionnaire data was grouped into high (40 and above) and low (30 and below) levels according to the total bandscores on the CAEL. These levels were then further grouped by those that had taken a CAEL Assessment preparation course and those that had not. The two groups' responses to the questions on the questionnaire were depicted in a chart and then examined for key differences and similarities. Chi-square analysis was used to further examine the differences between high and low test takers and their strategy use as identified in the chart (see above). The data were also analyzed using Spearman's correlation coefficient, using SPSS, 11th edition, to observe any meaningful or significant relationships between the strategies used and the test scores.

For all tests of significance in this study, the alpha level was set at ($p < .05$) and when examining meaningful differences in the percentages the cutoff was set at 10%. Thus, in crosstab analysis, represented in a contingency table which Field (2006) defines "as a table representing the cross-classification of two or more categorical variables. The levels of each variable are arranged in a grid, and the number of observations falling into each category is noted in the cells of the table" (p. 726). If one group of cells differs from the other by more than 10% this is viewed as a meaningful trend.

CHAPTER 4: RESULTS AND DISCUSSION

The organization of this chapter follows the same sequence in which the data was collected: classroom observations, pre-test interviews, post-test interviews, and finally, post-test questionnaires. Each section is numbered 1 through 4 to provide a clear distinction between paragraphs.

1. Preparation Course Observations

In this first section of the results and discussion chapter, I describe the preparation course in detail and the instructor's approach to the course and what he perceived strategies to be. Then I outline the strategies discussed in the course according to the skill tested and then by their classification in the language use and test-wiseness categories.

The two sections (A and B) of the course met twice a week for two hours during a 4 1/2 week period, with nine classes per section¹². In my description of the individual classes, I refer to the class by the section letter (A or B) and number (1-9), for example, A5 refers to section A, class number five. According the instructor, he planned the first class to be a general introduction of answering the first two questions on the oral component of the CAEL Assessment and did a needs assessment for the students that he had developed himself, the second class they started the CAEL diagnostic¹³. The remainder of the classes for both sections was to have the following sequence: 1) two readings, 2) one listening, 3) one speaking, and finally, 4) three writing; however, one of the students in Section A requested the order of the writing and speaking be reversed. For

¹² There were 18 classes total, but only 12 of those were observed. The missed classes: 4 [A 1,2 & B 1,2] at the beginning due to gaining permission and the timing of observing the course, and 2 [A 7 & B7] in the middle due to extraneous circumstances.

¹³ The CAEL diagnostic is a diagnostic test that provides the opportunity for test takers to become aware of their strengths and weaknesses are in relation to the skills and strategies needed to be successful on the CAEL Assessment (CAEL Assessment Office, 2005). The skills and strategies are designated with a letter and number that corresponds to the lists of skills and strategies in *The CAEL Assessment Test-taker's Preparation guide* (Fraser & Brisson, 2003).

section A only fifteen minutes was focused on the oral component due to extra time being spent on the writing section of the CAEL Assessment.

In the interview with the instructor he said that he felt the purpose of the CAEL preparation course, idealistically, was “not to just help students prepare for this particular test but help them prepare for university, the CAEL is like a mini lecture”. Though he admitted that “realistically the CAEL preparation course is being offered because the students think that they want it...[and] the motivation to create this course was to give students what they wanted, which was a course that focused on CAEL. So they can get past the CAEL, so they can get studying in regular university”.

Overall Description of the Course

The instructor incorporated the idealistic and realistic ideologies of the CAEL preparation course into his teaching: when there were elements that students were unaware of he would relate them to something that they would see in their future studies as mainstream students, but he also talked about what the markers would look for on the CAEL Assessment (e.g. what kind of answers would be acceptable on the test and what answers would only receive half points). The class sizes were quite small (three students per section¹⁴), which is contrary to Alderson and Hamp-Lyons’ (1996) findings that preparation class sizes were larger than that of normal class sizes (about 30 compared to 15). However, the instructor indicated that previous sections that he taught were a bit larger than the ones I observed (about five or seven), but still nowhere near the 30 student mark. These small class sizes may be a result of the course only being offered to students in the ESL program and those that have completed the intermediate ESL level course.

¹⁴ Although there were three students registered for Section A, I only observed one class where all three students were present. For the remaining five classes one of the students stopped coming to class and as a result was not interviewed.

This small class size allowed for a great deal of discussion between the instructor, myself and the students. The instructor would ask my opinion of how I would answer certain CAEL Assessment questions during the discussions. For example, in one of the classes on writing, when Sam explained his writing process he asked how I would approach an essay. I believe in this particular situation he asked my opinion to demonstrate that everyone is different in how they write an essay but, most importantly, to demonstrate that those in university brainstorm to some extent when writing. On the whole, however, I believe that Sam included me in the discussions to make the students feel comfortable with my presence and to bring another native English speakers' perspective on academic-like situations into the discussions.

As mentioned above, the classes included many discussions that were about the answers that the students provided for the practice test or their practice essays, which they would do at home and then give to the instructor and he and, possibly, a CAEL marker looked at them. Discussions concerning the students' writing would often occur at the beginning of the classes even if it was a non-writing class for five or ten minutes on a one-on-one basis, while the other students were arriving. During this time the instructor would identify major issues with the students' writing and suggest ways for them to overcome their problems.

Observations of Strategies Suggested to the Students

The instructor said in the interview that he didn't focus on skills the same way that the preparation guide does because the text was too difficult for the students to understand and skills are something that are developed over time, which was not possible in the course due to time constraints. He provided the analogy in class and again in the

interview of driving a car and how it relates to the use and development of strategies and skills: when you begin driving a car, you do many things that you are aware of and after awhile you do not notice the little things anymore. Sam felt this was “taking a strategy to a skill where you have operationalized it... to where you don’t think about it anymore”. Therefore, it seems that he is referring to the idea that strategies were consciousness based whereas skills were more automatic.

Various strategies were suggested to the students throughout the 12 classes and the mode of instruction was explicit – the students were told the purpose of the strategy. In Table 3 I present the strategies discussed in class and categorize them into the four language skills: reading, writing, listening, and speaking, which is the way they were presented to the students. (See Appendix VIII for an example of my field notes of one of the classes observed).

Table 3 Strategies Discussed in the Preparation Course Grouped According to Skill Tested

<i>Reading</i>
<p><u>Pre-reading Strategies:</u></p> <ul style="list-style-type: none"> ◦ Predict what is in the text by looking at the titles and subtitles ◦ Read all of the questions before reading the text ◦ Underline keywords ◦ Examine whether the flow charts require comparison or categorical type of answers <p><u>Reading Strategies:</u></p> <ul style="list-style-type: none"> ◦ Write and underline in the text ◦ Infer the meaning of a word from the context ◦ Skim and scan for answers ◦ Write answers in point form (CAEL does not require complete sentences for the reading or listening sections) <p><u>While Answering the Questions:</u></p> <ul style="list-style-type: none"> ◦ Know how much information to include by looking at the point value of the question

Listening

Pre-Listening Strategies:

- Read the questions first:
 - to look for keywords (to be able to recognize when information is provided for the answers while listening)
 - to look at the transitions among the questions (to get a sense of the sequence of the lecture)
 - to look at the point value to predict the focus of the lecture
 - to be aware of the direction (e.g. to know how many answers to check for a multiple choice question)
- Examine whether the flow charts require comparison or categorical type of answers

Listening Strategies:

- Listen for changes in pitch to identify important content
- Listen for transitions to understand the lecture
- Try to recognize the previously identified keywords in the lecture
- Write answers in point form

Post-Listening Strategies:

- Go over answers at the end:
 - to make sure the answers are legible
 - to proofread your answers

Writing

Pre-Writing Strategies:

- Critically analyze the writing prompt (is it argumentative, descriptive, etc.)
- Focus the argument first before brainstorming
- While completing the reading and listening sections, decide which direction to take for writing the essay
- Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay

Writing Strategies:

- Avoid using personal pronouns
- Keep what you cite simple
- Paraphrase your information by reading what you want to write and look away from the text and then write the information
- Integrate reading questions and answers into the essay
- Integrate listening questions and answers into the essay
- Include about three quotes in the essay
- Include specific information that answers the questions of who, what and how
- Integrate listening questions and answers

<p><u>Overall Strategies:</u></p> <ul style="list-style-type: none"> ◦ Use the time effectively (take 10 minutes to brainstorm and organize your thoughts and 20 minutes to write)
<p><i>Oral</i></p>
<p><u>First Question:</u> (The possible questions are known beforehand)</p> <ul style="list-style-type: none"> ◦ Memorize and practice your answers for the two possible questions ◦ If you don't know what to say, then make up the information ◦ Plan out what you are going to say <p><u>Read Aloud Questions:</u></p> <ul style="list-style-type: none"> ◦ Pay attention to where you place the stress on the words (it is not important to know the meaning of the words) <p><u>Overall Strategies:</u></p> <ul style="list-style-type: none"> ◦ Keep track of the time (the organization of your thoughts are important)

It is apparent from Table 3 that the number of strategies suggested for writing (13) is more than the suggested listening (11) and reading strategies (9), but they are not that different (9 or 11 compared to 13) while the number of strategies suggested for speaking is considerably lower (5). I acknowledge that these numbers are only an approximation due to the missed observations days and the nature of defining and determining what constitutes a 'strategy'. Therefore, I think it is important to define again what I mean by the term strategy, which I am defining here as the conscious effort made by a language learner or test taker to obtain a particular language goal (Cohen, 1998). In constructing the previous list of strategies I included tactics that in some situations may in fact not be considered a strategy because the classification of strategies is ultimately determined by their use. In other words, if the test taker makes a conscious effort to employ a tactic to help them reach a language goal, then that is a strategy, for the purposes of this research.

It is not surprising that the majority of the suggested strategies were writing strategies, considering the importance placed on the writing component in the CAEL

Assessment (is more heavily weighted) and is reflective of the importance of written expression according to the CAEL developers (Fox, 2002) of what the students will encounter in their mainstream classes. The listening and reading strategies may be comparable in number for a number of reasons. First of all, the essay is developed from the readings and listening, so it stands to reason that the test taker must have some understanding of both skills to do well on the essay. Moreover, the director of the CAEL had told the test takers that the writing is a secondary measure of the students' reading and listening ability. Second, reading and listening strategies are the most researched and therefore may be the easiest to teach. It may be that the limited number of strategies focused for speaking is due to the fact that speaking strategies are difficult to teach. Or the oral component of the CAEL Assessment is not perceived to be as important this seems plausible since the time spent on speaking was the lowest of the four skills, in fact, extra time was used for the writing instead of the speaking.

Furthermore, the writing has a weighted score because the CAEL Assessment developers concluded that for students to be successful in their mainstream studies they needed to be competent writers and not necessarily proficient speakers (Fox, 2002); therefore, at that time the test specification and weights reflected the students' needs. Since then, however, many changes have occurred within the university regarding the programs offered and course requirements for undergraduate students. In 2005 the Arts One program was implemented at the university where this study took place with a trial of 100 students and then in fall 2006 the number of students was increased to 600 students. The program provides organized sets of classes for Bachelor of Arts students based on a central theme so the students form study groups and friendships that may ease

them into their studies more readily than if they were in five classes with different classmates. The philosophy of this program is focused on communication with peers and forming friendships that build a strong foundation for their time at university. Therefore, a student in this program may have to communicate more than if they were in the regular Bachelor of Arts Program. In addition to the Arts One Program, first year seminar classes were introduced as a mandatory requirement for Bachelor of Arts students in 1999. The class sizes are small with no more than 30 students, so the students in their first year are in a non-threatening environment that encourages them to ask questions and speak more than they would in a large traditional lecture. Hence, the needs of students who do not speak English as a first language may have changed since they were assessed in the late 90's and then related to the current CAEL test specifications.

In order to discuss how the strategies relate to language learning and the test taking, I have separated the various strategies into Cohen's (1998) language use and test-wisness strategies groupings as seen in Table 4. The strategies are further divided into O'Malley and Chamot's (1990) categories for language learning and Millman et al.'s (1965) categories of test-wisness, respectively. By separating the suggested strategies into language use and test-wisness the course's emphasis on language learning and test-wisness provides a different perspective on the strategies suggested to the students. The strategies are followed by square brackets that surround the letter of the language skill that they refer to (e.g. [L] or [W]). A number of the strategies appear in both categories of language use and test wiseness which is identified by an asterisk before the skill identification symbol (e.g. *[L]).

Table 4 Test Taking Strategies Discussed in Class: Language Use versus Test-Wisness Strategies

<i>Language Use Strategies</i>
<p><u>Planning:</u></p> <ul style="list-style-type: none">◦ Focus the argument first before brainstorming* [W]◦ Use the time effectively (take 10 minutes to brainstorm and organize your thoughts and 20 minutes to write)* [W]◦ While completing the reading and listening sections, decide which direction to take for writing the essay* [W]◦ Plan out what you are going to say* [S] <p><u>Problem Identification:</u></p> <ul style="list-style-type: none">◦ Critically analyze the writing prompt (is it argumentative, descriptive, etc.)* [W] <p><u>Inferencing:</u></p> <ul style="list-style-type: none">◦ Predict what is in the text by looking at the titles and subtitles* [R]◦ Infer the meaning of a word from the context [R]◦ Examine whether the flow charts require comparison or categorical type of answers* [R/L]◦ To look at the point value to predict the focus of the lecture [L]* <p><u>Self-Monitoring/ Production:</u></p> <ul style="list-style-type: none">◦ Go over answers at the end:* [R]<ul style="list-style-type: none">- to make sure the answers are legible- to proofread your answers◦ Keep track of the time (the organization of your thoughts are important)* [S]◦ Pay attention to where you place the stress on the words (it is not important to know the meaning of the words) [S]◦ Avoid using personal pronouns* [W] <p><u>Selective Attention:</u></p> <ul style="list-style-type: none">◦ Underline keywords* [R]◦ Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay* [W]◦ Skim and scan for answers* [R]◦ Listen for changes in pitch to identify important content [L]◦ Include specific information that answers the questions of who, what and how* [W]◦ To look for keywords (to be able to recognize when information is provided for the answers while listening)* [L]* <p><u>Directed Attention:</u></p> <ul style="list-style-type: none">◦ Keep what you cite simple* [W] <p><u>Selective & Directed Attention:</u></p> <ul style="list-style-type: none">◦ To be aware of the directions (e.g. to know how many answers to check for a multiple choice question)* [L]*

- Read all of the questions before reading the text* [R]
- Read the questions first* [L]¹⁵
- To look at the transitions among the questions (to get a sense of the sequence of the lecture) [L]*
- Include about three quotes in the essay [W]

Resourcing:

- Integrate reading questions and answers into the essay* [W]
- Integrate the listening questions and answers into the essay* [W]

Note Taking:

- Write and underline in the text [R]
- Paraphrase your information by reading what you want to write and look away from the text and then write the information* [W]

Creative Elaboration:

- If you don't know what to say, then make up the information* [S]

Repetition:

- Memorize and practice your answers to the two possible questions* [S]

Test-Wiseness Strategies

Time-using Strategy:

- Use the time effectively (take 10 minutes to brainstorm and organize your thoughts and 20 minutes to write)* [W]
- Write answers in point form (CAEL does not require complete sentences for the reading or listening sections) [R/L]
- Integrate the reading questions and answers into the essay* [W]
- Read questions before reading the text* [R]
- While completing the reading and listening sections, decide which direction to take for writing the essay* [W]
- Focus the argument first before brainstorming* [W]
- Predict what is in the text by looking at the titles and subtitles* [R]
- Underline Keywords* [R]
- Keep what you cite simple* [W]
- Integrate reading questions and answers into the essay* [W]
- Skim and scan for answers* [R]
- Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay* [W]

¹⁵ The listening strategy *to read the questions first* has numerous sub-strategies which are designated with the symbol [L]*

Error-Avoidance Strategy:

- To be aware of the directions (e.g. to know how many answers to check for a multiple choice question) [L]*
- Know how much information to include by looking at the point value of the question [R]
- Examine whether the flow charts require comparison or categorical type of answers *[R/L]
- Critically analyze the writing prompt (is it argumentative, descriptive, etc.)* [W]
- Go over answers at the end:* [R]
 - to make sure the answers are legible
 - to proofread your answers
- Plan out what you are going to say* [S]
- Paraphrase your information by reading what you want to write and look away from the text and then write the information* [W]
- Keep track of the time (the organization of your thoughts are important)* [S]

Guessing:

- If you don't know what to say, then make up the information* [S]
- Search for keywords* [R/L]

Deductive reasoning Strategy:

- Integrate Listening questions and answers into the essay* [W]

Intent Consideration Strategy:

- Include specific information that answers the questions of who, what and how* [W]
- Avoid using personal pronouns* [W]

Analysis of Table 4 above demonstrated that there were more language use strategies than test-wisness strategies suggested to the students. A summary of the number of language use and test-wisness is provided in Table 5.

Table 5 Summary of Language Use and Test-Wisness Counts

	<u>Language Use</u>	<u>Test-Wisness</u>
Number of strategies that appear only in this category	5	2
Number of strategies that appear in both categories	28	25
Total	33	27

The strong emphasis on language use strategies is most likely reflective of the nature of the CAEL: its major focus is on production either through open-ended questions, the essay, or speaking tasks which is evident with the number of self-monitoring production strategies.

Finally, in an examination of the strategies that were both language use and test-wisness majority of the selective and directed attention (5 out of 9) and planning strategies (3 out of 4) were time using strategies, whereas the self-monitoring/production strategies were error avoidance strategies. Within the two categories, majority of the language use strategies were focused on self-monitoring (production), selective and directed attention and inferencing. While the two most prominent test-wisness categories were time-using and error-avoidance which according to Messick (1982) would increase the test takers sophistication of taking a test.

2. Pre-Test Interviews

Below, I describe the participants in more detail followed by the students' verbal reports of the CAEL Assessment in the following order: listening, reading, writing and speaking.

Charles

The first time Charles took the CAEL assessment was just over three months before the first interview. His best friend in Canada was from a middle eastern country, so they spoke English together. The CAEL Assessment preparation instructor, Sam, mentioned that Charles was probably the one student out of the five who benefited the most from the preparation course. During the interview Charles said that he liked doing

practice tests. He also mentioned that there was a CAEL preparation course offered off campus, but he said that he didn't think it would work.

[The course] is opened by the Chinese to help students pass the CAEL.
But it doesn't work, although lots of students believe it.

Alexandra

Alexandra had studied English in her home country for nine years, however, she said that her experience with English came mainly from traveling with her family to the United States, France and England. Unlike the other participants, she had never taken the CAEL Assessment. Her first language was Arabic. She wanted to enter the master's program in Computer Engineering at the university where the study took place, and she had a post-graduate degree from her home country in Fine Arts. She hoped to get a 40 or 60 on the CAEL although she indicated that it did not matter because it was her first time.

Alexandra mentioned that part of her motivation to take the CAEL Assessment was because it would prepare her for university:

I think it is more academic (.) to prepare me for (.1) university because final (.2) uh ^you know I have my brother he is in ...[another] University and he get the TOEFL test but also he said that until now I have some problem for writing essay and because he didn't have any %you know%. And he said the CAEL test is better because um in [this university] they will teach you about how you write the essay ah assignment. And this is what you need it in university. So I should now take the CAEL test.

Alexandra also felt that if she took the CAEL Assessment and received the appropriate bandscore then she would be good in English even if she did not complete her studies in Canada because she viewed the test as a certificate:

If I complete in [this university] or not and I come back to my country I have a ^CAEL test, so that means I am good in English. I have English test you know and (.2) so I think it is a certificate for me.

She felt that one of her problems with the CAEL was that she was slow to answer questions and read the text:

The time. I am so slowly because I want to write down the main you know. I want to this is this is what he want the answers. So I just read carefully. ^Anyway I like to read not just for one time but for many time. And this is problem because I have another questions to read and no time, so I try to read it one time and write all or (.) that at the last I review and again try to read it and answer. But it is still difficult because of time.

Though she indicated that she was a slow reader and would read things many times, which is an excellent learning strategy but not a effective test-wiseness strategy.

Ben

Ben was from Turkey and had been studying English since he was in preschool. He came to Canada three months before my first interview with him. He was also best friends with someone who was from a country where they did not share the same language so they spoke English together. Ben did not receive scores from his first CAEL Assessment because he missed the bus for the Oral Language Test and as a result did not obtain any scores for the written portion of the test.

Donald

Donald was from the United Arab Emirates. He had studied English for 14 years in his home country and in Great Britain for 6 months. He took the CAEL for the first time 9 months ago. He said that he felt reading and listening were his strongest skills and had hoped to get a 40 on the CAEL Assessment. He had planned to take the CAEL Assessment with the other students, but unfortunately he did not bring all of the required documentation to sit for the test.

Edward

Edward was from Jordan and had studied English for 12 years. He came to Canada directly after high school. He mentioned that his roommate was from India and spoke English to him. At the time of the interview, Edward was in the high beginner ESL program and completed the required homework, but also did extra homework because he wanted to improve his writing and reading skills. Edward would like to get a 60 on the CAEL because he wanted to gain admission directly into the year-long ESP course and had signed up for the preparation course the following semester. So he did not take the CAEL Assessment when the other test takers did because he felt he was not ready to take it and planned to take it in six to eight months time. He also mentioned that he saw the CAEL Assessment as a test that measured both English and academic skills which was different from other tests such as the TOEFL or IELTS:

Other tests like TOEFL or IELTS just measure about English not your academic skills so that is why the [CAEL] topic that we had was academically.

In Tables 6 and 7 shown below, I present the students' bandscores from their first CAEL Assessment and the scores they received on the second CAEL when they completed the CAEL Assessment preparation course.

Table 6 Bandscores for Test Takers' First CAEL Assessment

	Total Bandscore	Reading	Listening	Writing	Speaking
Charles	20	20	20	10	20
Alexandra	N/A	N/A	N/A	N/A	N/A
Donald	20	N/A	N/A	N/A	N/A ¹⁶

¹⁶ Donald was unable to remember his specific bandscores for his first CAEL Assessment, but he did remember that his writing was the lowest score.

	Total Bandscore	Reading	Listening	Writing	Speaking
Edward	20	20	30	10	30
Ben	N/A	N/A	N/A	N/A	N/A

Table 7 Test Takers' Bandscores for the CAEL Assessment after the Preparation Course

	Total Bandscore	Reading	Listening	Writing	Speaking
Charles	20	40	40	20	40
Alexandra	20	20	30	20	30
Ben	20	20	30	20	30

In considering the tables above, it seems that Charles made considerable improvement in his reading, listening and speaking scores from a 20 bandscore to a 40, he was also able to improve upon his writing by one band. Despite this considerable improvement he was unable to improve his overall score since he did not receive a writing score of 30 or higher. It was Alexandra's first time taking the CAEL and Ben did not have his scores from his first CAEL Assessment, hence, there was no concrete basis for comparison to discuss the test takers' progress or improvement. However, both test takers were placed in the same course as Charles at the beginning of the semester.

Verbal Reports

In this section I present the results from the verbal reports according to skill and grouped by strategy with a brief description of the test takers' account. Since I cannot present all of the test takers' strategies due to space limitations I have provided a table at the end of this section summarizing the students' strategy use. In addition, I only report on strategies that the test takers themselves have reported using. I acknowledge that as a result I will not be recording all of the strategies the students make use of.

Listening: Pre-Listening - Read the Questions

All five test takers read the listening questions during the pre-listening time, but when they talked about what they were doing they gave different reasons or objectives for reading the questions first. The reasons were: to underline keywords, to know how to follow directions, to underline the main point(s), to guess the answers to the questions.

Underline Keywords

During the pre-listening time, Charles circled and underlined various words throughout the questions. When he described what he was doing after the listening he responded,

Well, (.2) first I uh I should just circle the keywords. I will [then] remember the sequencing of each question. Maybe because it is easy to know what they will talk about in the listening

When Ben was asked what he was doing during the pre-listening time, he responded that he was underlining keywords so he could listen for them, and to understand the organization of the lecture:

I underline the keywords which I don't know and I want to hear this word in my mind. To organize these speakers.

To be Aware of the Directions

Alexandra indicated that one of the reasons she was reading the questions was to know how many answers were required to answer the question:

I read the question. I underline the the %main topic% and um I put the circle between if he tell me he needs three or one or two answers.

In one of the classes observed, Alexandra got a question wrong in the reading section of the CAEL Assessment diagnostic test, which was given to the students at the beginning of preparation course, because she didn't read the whole question. Although that was the

reading section and this is the listening section, it is possible that she is drawing on previous experience for what she focuses on when reading the listening questions.

To Underline the Main Point

Edward said that one of the reasons he read the questions was so he could underline the most important point or purpose of the question: "I will underline the most important point (.2) what you call the purpose of the questions".

To Guess the Answers to the Questions

Edward and Donald both said that they would guess the answers to the listening questions in the pre-listening section:

Edward: First to read the questions and to guess the answers, which is going to be useful or helpful for when I listen

and,

Donald: I am going to try to read the questions. After that I am going to answer all of the questions. First, I guess in mind what the answer will be. After that I am going to listen for it.

During the Listening

Try to Recognize the Previously Identified Keywords in the Lecture

When Charles was asked if he heard the keywords that he had underlined in the pre-listening he said that he did: "Yeah, I just hear the main function in the listening, so the answer would be later". An interesting aspect of Charles' comment here is that he says that he is listening for the keywords to hear the main function of the listening. However, the strategy suggested in class was to listen for the keywords so the test takers can know where you are in the lecture.

Ben said that for the first question he was trying to listen for keywords, specifically, *Brain and Behavior*, however, he didn't hear that phrase but he did hear *nervous system*.

I expected brain and behavior but it was nervous system. I catch it and I talked about nervous system.

The question asked for what was the topic of the lecture, and in the pre-listening section Ben had put a small check beside the answer *The brain and behavior*, which could explain why he was trying to listen for the keywords *Brain and Behavior*. In another instance, Ben mentioned that for question 6.1 "What are the three main functions of living systems" he was listening for the words *living systems* because he expected to hear what the three main functions were immediately after:

And the other questions living systems and streaming functions. In English writing essays it always says to explain and I waiting to hear living systems. After I hear living systems come I know what living systems and three main functions [are].

Look at the Point Value of the Question

Edward said that for final question, question 7, in the listening section he did not notice the point value for the *notes* question and felt that he did not do well on the question because it was worth six points:

I didn't do well on the notes. I know it is six points but that is my problem, I didn't realize that six points that I have to do well in it. I was focusing on the last question here ((pointing to question 6)) and I lost my balance here ((pointing to question 7)).

Ben also mentioned the taking notes question and that it was worth six points, but he felt he was only able to write four points:

And take notes. It's a uh [the instructor] told me that you have to write six points. Six main parts of the writing. I think I did four.

Post-Listening – To Go Over Answers

Make Answers More Comprehensible

During the time given for the post-listening, Charles went through his answers and would erase words. Charles indicated that he used the post-listening time to add information: “just add some information which I had lost”. When asked about what he was doing when he erased words, Charles responded that,

Yeah, uh, I will because when I read the listening I read some keywords and observe the listening of how to reorganize the sentence (.) reorganize the words to make a sentence.

Edward said that during the post-listening section he would write his answers clearer:

The listening end. Like the third question was wrong, so I just write something clear so that the marker can read it. That’s all.

To Change the Spelling of Words

Charles also said that while he was erasing words during the post-listening he changed the spelling of some words: “Uh change change (.5) some spelling and I think that's it”.

Reading

In the reading portion of the retrospective verbal protocols, test takers reported a variety of strategies: read all of the questions first, skim and scan for answers, underline words or look for keywords, use cues in the text or questions, star the readings, limit the time for each question.

Read all of the questions first

During the reading portion Edward read all of the questions first before answering any of them. He said that reading the questions first helped him generate a general

overview of the questions and helped him focus for when he would skim and scan for answers:

[I read the questions first] to make a general idea about how the questions are going to look. And to make me more specific when I skim, make me more easier when I scan the answers. But I wasn't lucky in the scanning.

Skim and Scan for Answers

Edward felt that he was not able to skim or scan for answers that well because of his limited vocabulary and low English proficiency:

I think if my English was a little bit better I could do much better. Because if my English is going to be better I could be able more to skim the reading in like four minutes or five minutes... So, I first read the questions so that when I skim it make me very easy to answer the question again. So, I was very slow because there is a lot of new vocabulary you know.

Underline words/Look for keywords

While Ben completed the listening he had underlined some words twice and circled others. When asked why he double underlined and circled words he responded that it was because he did not know the meaning of the word:

Christine: Why were you double underling here?

Ben: I don't know what was the meaning of the word. (.) Very biological. I think it comes from latin.

and,

Christine: You circled trait here?

Ben: I don't know what this trait is. But it will be important. Too many traits...%I underlined something%. I don't know why I couldn't find it.

Charles had some difficulty with the same question and resorted to identifying the keyword trait to answer the question.

In this question I don't understand clearly. But uh (.) I just think (.) because I find the word trait is repeated here, so I just choose that answer here ((pointing to the last sentence of paragraph 6 in Reading 2)).

Ben said that for some of the questions he underlined words but was unable to find them in the reading:

I underlined the main or specific words which I need, but the two meaning inverse in there in which I underlined before. And I don't know it was hard for me. I thought it empty because I couldn't find mechanism or brain.

The question that he was referring to was, "What two technologies can now allow us to understand how mechanisms in the brain determine differences in personality" (Reading 2, p. 47). And the words he was searching for was *mechanisms* and *brain*. The fifth paragraph of the reading text where the answer was located primarily was about *new tools* and how scientists can now "visualize how the brain works" (Reading, 2 p. 47). Therefore, it seems that Ben was unable to find the answer because the keyword *mechanism* was not in the text.

Edward also had trouble with the same question when he was searching for the keyword *technologies*, in fact, when he was talking about looking for the keyword he was pointing to paragraph 5: "I didn't find the technologies here (fifth paragraph) I didn't find any word technologies". Although *technologies* and *new tools* were used synonymously in the text, Edward was unable to make the connection between the two.

As with Ben and Edward, Alexandra had trouble with the question. She left the answer blank, but she started to make the connection during the verbal report:

Alexandra: The fourth one is technologies. .hhh You know first thing is that I think it is here ((pointing to the paragraph)) (<biology of temperature >) this is the first one and then new one is but (.2) actually I don't find it.

Christine: So (.2) what are you looking for in ^here?

Alexandra: Technology. The new=tech the tool. Technology. And I read this (emerging technos) so actually I don't know. Here I just write I uh just read (.2) this teacher you know. It's too hard. I just (.) ((reading from the text)) %the ways in which brain function we% ((stop reading)). Yeah the first one I think ((reading from the text)) %this a kind of technology% ((stop reading))) and here (reading from the text)) %new tools for brain chemicals%. ((stop reading)) I just read new, so I think maybe it is this one and all and the other thing it's here. The new one. I just guess, but I don't know exactly what is the word. I just think about it.

Although Alexandra started to make the connection between *technologies* and *new tools*, it appeared that she was searching for a specific word as seen with her last statement "I don't know exactly what is the word".

Identify the Answer through a process of deduction

Charles was the only test taker not to use a keyword strategy to come up with the answer for question number four. He determined the answer by relating technology to machinery and deducing that new tools was referring to technology:

Technology (.) I think it is some machine, so I read here I read new tools. I think it will be the technology. And there it is a repeat of the technology it will be the answer.

In question number five, Alexandra found her answer by looking for the words *useful* and *biological*:

((reading question 4) Why it's useful to understanding ((stop reading the question)) I here ((pointing to paragraph 7 in reading 2)) understanding you know. So I just look for what=it why it's useful. The biological understand. Here is useful ((underlining the word useful in the text)) that is useful. What is it is before you know. Here the ((reading from the text)) the biological basis for describe so we=we this is the question you know.

Underline keywords/Use Cues in the Text

For question number three, Alexandra stated that she wanted to find the word factor:

A factor (.2) I just read it. I will look for this point you know. There is not the same=similar word. Not the same word, so I read again and what is factor and I read here there is something factor and there is and: this is name of factor and so (.2) by is the first after that and this is the second one.

The text that Alexandra was referring to is: “our personal appearances are shaped by genetic and biological factors” (Reading 2, p. 47). It seems that Alexandra used the keyword in the question and the words *by* and *and* to help her answer the question. When asked about why she choose she responded accordingly,

Christine: So, why did you choose these two answers?

Alexandra: Um because biological factors this is the name of the first one. This is the second one in paragraph, but I just oh this is factors so and this=there is (.2) word it is and and so it is two things. You know it explains there is and it join two things. I write this one down here ((point to the answer)).

Charles found the same answer using the cue *early experiences* in the question:

I guess the answer from here. Because it is talking about the early experiences it is definitely not here because it is new research, so I can not. (.) Here they talk about the life experiences and the biological factor and its because its new research I uh that I don't think so it will be the answer. Because here (.2) ((reading from text)) many of us take for granted ((stop reading)) I think it is happening before, so I think it is early experiences by the genetic and biological factors.

Though Alexandra and Charles both used the keyword strategy to find their answer, they also drew on features in the text to confirm the answer. Hence, they drew on the keyword strategy but also their knowledge of the surrounding words and the function of them in the sentence or their meanings.

Eliminate Choices for Multiple Choice Style Questions

Charles stated that he would eliminate choices given for multiple choice questions as with question 2 of the reading: “the first thing I will cancel (.) cancel some choice I think it is not totally magazine. A magazine is not always so academic”.

Answer the Main Idea Question Last

Charles said that he answered the main idea question last because he could then properly answer the question,

The first question talk (.) about the main idea of the article. When I encounter this question I always do it in the last. (.2) Because I have to read it (.) read all the topics (.) all the articles so I can guess the answer.

Look for the Questions in Sequence

At the beginning of the reading section of the verbal report, Alexandra expressed concern over the order of the questions. To her the questions were too far apart in the text:

Alexandra: I think (.2) there is some order is %you know% the order

Christine: The order isn't in sequence?

Alexandra: No, it's so far. It's in sequence, but it is so far between this question and I mistake technology.

She also mentioned that she expects the questions to be located in a certain part of the text because of where they are in the questions:

If this in the middle. We should be in the middle because one by one. ...if its in the middle it should be in the middle, so I can know this is.

Looking at the headings in the Charts

Alexandra, Charles and Donald were the only test takers to have enough time to complete the chart. Alexandra expressed her dislike for the chart during the verbal report:

“The charts. hah. I don’t like it anyway you know. It’s difficult because you know”. And when asked why she does not like charts she responded that since charts are not questions they have less context she has to read more details:

Because you know I know the chart it’s a good plan for you for if there is a medicine and the problem for this medicine and the reason for this medicine. (.) But you know sometimes it I can’t (>because its difficult word use all its new, so I don’t know if this means good or bad<). It isn’t a question. If why because of these things.... So that’s why when I saw it I hate it hah because I have to read more specifically.

She was looking at the main heading and the cue words in the chart to help her answer the question:

Alexandra: I read first these things the (.2) the main ((pointing to the headings)) you know

Christine: The headings?

Alexandra: Yeah the headings. Trait, and I read this talk about problem and this something I don’t know what, but I try to look for this one. This word (point to the word impulsivity).

Christine: Impulsivity?

Alexandra: Impulsivity. I want to see where is the problem. I read here problem (point to paragraph 9) in brain chemistry. Here problem and here impulsivity ((point to paragraph 10)) from so it is some word at the back, so I have just to look for this...so I just write (.2) there is a problem and may be irritable and maybe get into fight. I think this is kind of problem. I need one. I try to look for these things, but actually maybe I don’t even try at all. But I didn’t find it and the time over.

Look for Examples

Charles said that since the charts are always comparing something he needed to look for the examples that are used to compare the categories:

In the charts it is always compare with each other, so it talk about the problems. Firstly, I will look for the example that’s why it is talking about system. Some system I think so when I read here is the message system. I

think it is compare with the (.) with the next one. So I write it I write it down.

Manage Time for Questions

Edward took off his watch to time himself for each question. He would look for an answer for a while and then look at his watch. If he was taking too long it appeared as if he would move on to the next question. When asked why he took his watch off he responded, “I have to watch the time so I decide I should take five minutes for this question so I just do like that”. This strategy of timed questions was not observed in the CAEL preparation course for the reading section. Even though a similar strategy was observed for the OLT, I believe this was Edward’s personal strategy.

Writing

For the writing portion, I had the test takers go through the brainstorming process for writing the essay. They reported primarily on what and how they were brainstorming to write the essay. All of the strategies describe below were components of the test takers’ brainstorming. The writing prompt for the essay was “How does an understanding of brain chemistry help doctors work with people who have personality disorders?”

To use the Listening and Reading

In the notes section of the writing, Edward put a circle around the words *understand*, *help doctors*, and *work* with the writing prompt in the center of the page, which he described as the beginning of his brainstorming:

I just brainstorming. I just put the question here and put in what is the main point to answer the question...so I can relate it to the reading or the listening.

Charles mentioned that one of the things he would do while writing his essay is relate it to both the reading and the listening.

I will try to connect the information from the listening and the reading

I will (.2) then it [the essay] can be as two claims and the I am searching for some some examples from the reading and the listening and connect with the claim (.2) that's it

While Alexandra was explaining how she would focus her argument or critically analyze the prompt she mentioned that she would use two questions from the listening and reading and a part of reading 2 to write her essay:

So I need this question ((pointing to question 4 of the listening) to explain what is inside the brain. And what the job of these section also that one is how from listening to describe the brain. And from the help doctor here to understand some biological, so from here ((point to question 5 in the reading)) I can get...maybe I use a quotation where doctor says help help because here ((pointing to the reading 2)) I think there are many.

Star the Readings

During the reading, Edward put stars in the margins of his reading booklet. He said that he did this so he could go back to the starred part when writing the essay:

When I was reading I realize, I was going to write about the reading so that's why I just make some stars and points that's going to be useful in my writing. So I can get back again and take some specific information and make it in brackets ... and make my writing that much better.

Though this strategy was noted in the reading section of the verbal report, I have placed it in the writing section since it is a strategy that aids in writing the essay.

Focus the Argument

Charles mentioned that in his essay he would have focused on life experience, genetic and biological factors:

I will focus on the different aspects of how (.2) no I will try to focus on the different aspects from the life experience factors and from the genetic and biological factors.

Focus Argument/Use space as a signal of how much to write

It was also very important for Charles that he only make two ‘claims’ as he put them because of the amount of space provided:

That's where I can not now I'm not so now the writing space is not so fast, so I can not use no more than two claims.

This comment suggested that, Charles knows how much room it will take to write about two points. Therefore, he is aware of the test format and layout to predict how much time it will take him to write the essay.

Critically Analyzing the Prompt/Include Specific Information that Answers the Questions of Who, What and How

Ben analyzed the prompt by focusing on the word *how* since he stressed that the writing prompt was asking for facts in the essay because it was a *how* question:

This question how, it is not argue. It is correct, or is it wrong, or it is not arguing. It means how. It's (.) I use too many things. It means we should find it the reading, listening all things how. We don't (.) can not use our ideas because it is brain chemistry. I didn't answer before for or I didn't answer for who, what.

Alexandra said that she should first read the question and figure out what the question is asking for by focusing on the words *how*, *help* and *doctors*:

I read the question and I tried to know what he want for this question. (.2) ((read the prompt)) how does an understanding of brain chemistry help doctors ((stopped reading the prompt)) so he want me to me to try to know ^how (.2) the things like the brain has these things to describe or explain what is it. hhh What is brain chemistry maybe what is the job. And the doctor how he and how he take it from these things. How it will help him to and I read the listening. I just look for the (.2) I put some words like how and the help this mean doctors, so I just look at the questions and I just read the topic what the answer is. What is the major nervous and that. I want to explain what is the brain and what is have. The cells and the soma and it have these things.

While she was reading the prompt, Alexandra underlined the whole question but she also underlined the word(s) *how* and *help doctors* again and put a circle around *help*.

Speaking: The first question

Be aware of the time

For the first question, Edward discussed his plans for university and why he was currently studying English. But when asked what he was focusing on while answering he said that he felt focused too much on why he was studying English:

[The question] was very specific. Like what you want to study at university and why. And I way I want to study electrical engineering ... I didn't really have a chance to speak more about that maybe because I waste much time why I am not studying English right now. So I didn't know why I did that.

Have the prompt memorized

Ben said that he would have the prompt memorized because the questions are always the same:

[The instructor] said that before entering the CAEL you should create little bubbles about. Because they are always the same. And we should always think about what we are going to say.

Answer the 'what' and 'why' of the question

Alexandra said that she had to answer the questions of what and why to properly answer the first question:

She ask me what you want to study at university and why, so yeah so I have to give him the name of my major and why I choose it.

Consideration of What is being Measured

She also indicated that she felt the question was measuring her speaking ability and pronunciation:

I think this task to see how I speak English or my pronunciation. Is it good or not. And (.2) is it easy to say English because some people have some problems with but I think they want me to speak quickly and fast and a lot of words, so maybe that's it.

Third Question

In the third question Alexandra said that she would underline long or difficult words so that she would know how to pronounce them when it came to reading the paragraph out loud:

Alexandra: You know I just look for difficult words or a long word. I tried to put some comma between you know like similarities. (.2) how I say sim l-r lar ti=ties like this you know and through and I think it is one word and it have many words to stop [(and)]

Christine: [(syll)]ables?

Alexandra: yeah so I just try to at first because she said that look over the text so I just so I will underline next time. I will underline the difficult words and put some commas (>and then I read it<). Maybe some words generalized like this you know actually I don't know if it's like this I just read it. Maybe some words is silent, but I don't know which how I use it.

A summary of the test takers' strategy accounts obtained from the verbal reports is provided below in Table 8. The number of strategies accounted for are summarized by test taker at the bottom of the table and by strategy along the right-hand side of the table and the asterisk next to the strategy indicates that it was suggested to the students in the preparation course.

Table 8 Summary of Strategy Use by Test Takers

Strategy	Charles	Alexandra	Ben	Edward	Donald	Total
	<i>Pre-Listening</i>					
Underline Keywords**	√	√	√			3
Underline Main Point*				√		1
To be aware of directions*		√				1

* Indicates the strategies discussed in the preparation course

Strategy	Charles	Alexandra	Ben	Edward	Donald	Total
To Guess the Answers to the Questions				√	√	2
<i>During the Listening</i>						
Recognize Previously Identified Keywords*	√		√	√		3
Look at the Point Value of the Question*		√	√	√		3
<i>Post-Listening</i>						
Make Answers More Comprehensible*	√			√		2
Change the Spelling*	√	√			√	3
<i>Reading</i>						
Read all of the Questions 1 st *				√	√	2
Skim and Scan*				√		1
Look for Keywords*	√√	√√	√√	√	√√	9
Use Cues in the text		√			√	2
Eliminate Choices	√					1
Answer the Main Idea Question Last	√					1
Look for Questions in Sequence		√				1
Examine Headings in Charts*		√				1
Look for Examples	√					1
Look at the Point Value*	√					1
Manage Time for Questions				√		1
<i>Writing</i>						
Use the Listening and Reading*	√			√		2
Time Spent Planning*	√					1
Focus the Argument*	√	√			√	2
Star the Readings*	√			√		2
Critically Analyze the Prompt*		√	√			2
Use space as indicator of how much to write	√					1

Strategy	Charles	Alexandra	Ben	Edward	Donald	Total
	<i>Speaking</i>					
Be aware of the time*				√		1
Have the Prompt Memorized*			√			1
Answer the what the why of the question*	√	√				2
Underline long and difficult words		√				1
Organize thoughts before answering*	√					1
Total	17	13	7	12	8	49

Overall Charles was the most *strategic* test taker. Alexandra and Edward were similar in how strategic they were, but not far behind Charles. While Donald and Ben were the least strategic of the five test takers interviewed. Charles' highly strategic behavior could be explained by his serious approach to the verbal report. While completing the practice tests he was an extremely fast writer and would flip pages quickly, it seemed as if he were pressed for time and stressed about completing each section.

While Donald's behavior could be considered the opposite, but this may be due to the fact that he was not as descriptive in his verbal reports as the other test takers. In addition, he mentioned that he had trouble listening to the lecture because he felt it wasn't clear - while he was listening to the lecture there was someone talking outside of the room which may have interfered with that portion of the verbal report. Ben expressed a dislike for the practice test topic which could explain why he chose not to review his listening answers during the post-listening because as soon as he finished the listening section he said "I hate biology. This looks like a pain. I hope the next CAEL looks like

math or actual information”. I asked him if he wanted to review his answers and he said no. Nonetheless, a strategic and/or a highly motivated test taker might not view the topic as an issue and would have completed the test regardless. Since the goal of the test is to pass, a wise test taker would complete whatever has to be completed and not consider any other factors that may impede the process. In consideration of factors that could contributed to Ben and Donald’s less strategic behavior they are still the least strategic test takers because of the large gap between them and Charles.

All of the strategies used by the test takers, except for the “use examples” strategy by Charles, are test-wiseness strategies. Charles and Alexandra sometimes drew on more than one strategy or would look for language cues to support their answers. Charles used only the keyword strategy if he was not able to pick up on any other language cues, and he described it as guessing. While answering some questions Alexandra also appeared to draw on language cues in the text to and keyword strategy to answer some of the reading questions. And Edward adopted his own strategy of timing himself, but he recognized that he did not have the language required to apply the strategies effectively.

While Ben used only the keyword strategy and did not draw on other language knowledge to aid him in his decisions, which according to Millman et. al.’s (1965) notion of guessing would be “blind guessing”. He suggested that test takers should always be encouraged to make informed guesses rather than blind, but acknowledged that test takers make blind guesses when they are pressed for time or lack motivation. Furthermore, Yang (2000) found that test-wise test takers on the TOEFL appeared less random and logical in their choice of answer a description that characterized Charles, Alexandra and Edward’s behavior.

Furthermore, Ben and Donald both demonstrated un-test-wise behavior prior to the pre-test interviews. By missing the bus, Ben exhibited the opposite of what a good test-wise test taker would do: try to arrive early and be aware of the bus times on the weekend. While Donald should have been aware of the documentation because a major component of taking tests is the knowledge of what is required beforehand to be able to sit for the test, especially for such a high-stakes large scale test such as the CAEL Assessment. Though the preparation instructor had told the students bring the proper identification for the test. May be this point was not stressed enough for Donald to realize how imperative it was that he bring the appropriate documentation to ensure that he can sit for the test. The type of errors in taking a test that Ben and Donald did are not uncommon but nor are they related to language proficiency. They are the few things that Ben and Donald could have done to ensure that could have at least got a score.

Eight of the strategies reported by the test takers were not presented to them in the preparation course. I provide a summary of the strategies used by the test takers that were presented to them in the preparation course in Table 9. The number of strategies discussed in the preparation course appear in brackets to the right of each underlined sub-heading.

Table 9 Test Taking Strategies reported in the Pre-Test Interviews: Language Use versus Test-Wiseness Strategies

<i>Language Use Strategies</i>
<p><u>Planning</u>:(4)</p> <ul style="list-style-type: none"> ◦ Focus the argument first before brainstorming [W] ◦ Use the time effectively [W] ◦ Plan out what you are going to say [S]

Problem Identification: (1)

- Critically analyze the writing prompt (is it argumentative, descriptive, etc.)* [W]

Inferencing: (4)

- Infer the meaning of a word from the context [R]
- Examine whether the flow charts require comparison or categorical type of answers [L]

Self-Monitoring/ Production: (4)

- Go over answers at the end: [R]
 - to make sure the answers are legible
 - to proofread your answers
- Keep track of the time (the organization of your thoughts are important) [S]
- Pay attention to where you place the stress on the words/underline long or difficult words [S]

Selective Attention: (7)

- Underline keywords [R]
- Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay [W]
- Skim and scan for answers [R]
- Include specific information that answers the questions of who, what and how [W]
- To look for keywords (to be able to recognize when information is provided for the answers while listening) [L]

Selective & Directed Attention: (5)

- To be aware of the directions (e.g. to know how many answers to check for a multiple choice question) [R]
- Read questions before reading the text [R]
- Read the questions first [L]¹⁷
- To look at the transitions among the questions (to get a sense of the sequence of the lecture) [L]
- Use the space as a prediction to know how to write [W]

Note Taking: (2)

- Write and underline in the text [R]

Repetition: (1)

- Memorize and practice your answers to the two possible questions [S]

¹⁷ The listening strategy *to read the questions first* has numerous sub-strategies which are designated with the symbol [L]*

Test-Wiseness Strategies

Time-using Strategy: (12)

- Use the time effectively [W]
- Write answers in point form (CAEL does not require complete sentences for the reading or listening sections) [R/L]
- Read all questions before reading the text [R]
- Focus the argument first before brainstorming [W]
- Underline Keywords [R]
- Skim and scan for answers [R]
- Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay [W]

Error-Avoidance Strategy: (8)

- To be aware of the directions (e.g. to know how many answers to check for a multiple choice question) [L]
- Know how much information to include by looking at the point value of the question [R]
- Examine whether the flow charts require comparison or categorical type of answers [R/L]
- Critically analyze the writing prompt (is it argumentative, descriptive, etc.) [W]
- Go over answers at the end: [R]
 - to make sure the answers are legible
 - to proofread your answers
- Plan out what you are going to say [S]
- Keep track of the time (the organization of your thoughts are important) [S]

Guessing: (2)

- Search for keywords [R/L]

Intent Consideration Strategy: (2)

- Include specific information that answers the questions of who, what and how* [W]

Interestingly, the number of selective and directive attention and resourcing strategies are high, which may demonstrate that the test takers desired focus when they were completing the practice CAEL Assessment. When examining the test-wiseness strategies, all but one of the error avoidance strategies were used by one of the test takers. Furthermore, 75% of the test-wiseness strategies were drawn on while 65% of the

language use strategies were used by the test takers. It should be noted that there are many writing strategies not recorded this is because the test takers did not write the essay during the verbal reports and so it is obvious that during if they had written they essay there would have been more strategies reported for that portion of the CAEL Assessment.

In this section, I examined the strategies that 5 test takers employed while completing a practice test. Overall, the majority of the strategies used were ones that were previously discussed in the preparation course. In the next section, I will investigate the strategies that test takers reported using during the live administration of the CAEL Assessment.

3. Post-Test Interviews

The strategies the test takers reported using during the live test administration are summarized in Table 7. In addition to myself another rater coded three sections of Charles' interview (speaking, reading and listening) and one of Alexandra's interviews (speaking) for strategies. The inter-rater reliability coefficient was calculated using Pearson's Correlation coefficient ($r = .98, p < .01$).

Table 10 Strategies Discussed in the Post-Test Interviews

Strategy	Charles	Alexandra	Ben	Total
	<i>Pre-Listening</i>			
Underline keywords*	√	√		2
Underline difficult/unknown words		√	√	2
Underline main point*	√			1
Underline the point value*		√		1
	<i>During the Listening</i>			
Recognize previously identified keywords*		√		1
Use own information to answer questions		√		1
	<i>Post-Listening</i>			
Use the reading to answer the listening			√	1

Strategy	Charles	Alexandra	Ben	Total
Change the spelling*	√			1
Draw a line to the correct answers		√ ¹⁸	√	2
<i>Reading</i>				
Read the question and then look for the answer			√	1
Skim and scan*	√		√	2
Look for keywords*	√	√		2
Look for questions in sequence		√√		2
Manage time for questions		√ ¹⁹		1
Use The 1 st Reading To Answer Questions In The 2 nd Reading	√		√	2
Use time in the writing to answer the reading		√		1
Use subtitles to answer questions	√ ²⁰	√		2
<i>Writing</i>				
Use the listening and reading*	√	√	√	3
Focus the argument*	√	√	√	3
Critically analyze the prompt*		√	√	2
Star the readings*		√√	√√ ²¹	3
Leave space to write introduction	√			1
When 10 mins are left, leave space and write the conclusion		√		1
Place arrows to show the order of the sentences		√		1
<i>Speaking</i>				
Have the answer memorized*	√	√	√	3
Used the time to circle dates, books and professor's name	√			1
Underline long and difficult words	√	√	√	3
Plan out what you are going to say*	√		√	2
Pronounce words in your head	√			1

* Indicates strategies that were discussed in the classroom observations

** Some strategies have two check marks for the same person. This indicates that the strategy was mentioned twice.

¹⁸ Alexandra underlined the correct answers

¹⁹ Alexandra did not use a watch like Edward in the pre-listening, but said that if she was taking too long on a question she would move on to the next question.

²⁰ Charles used the subtitles to answer the questions in the charts

²¹ Ben placed a W next to the section of the reading and some reading and listening questions he planned to use in his essay

Strategy	Charles	Alexandra	Ben	Total
Write out 1 or 2 keywords before speaking		√√		2
Read newspaper out loud at home before the test		√		1
Total	14	24	15	51

Overall, the total number of strategies recalled in the post-test interview is higher than the number of strategies discussed in the pre-test interviews, even though there were fewer test takers interviewed. This may be due to the fact that the test takers were discussing what they did for the live administration, and therefore, they were more strategic, because it was a live test environment environment. Or, there could have been more accounts of strategic references because they were discussing the full-length test and not a reduced practice test.

During the post-test interview Ben was more descriptive and seemed more content about his performance on the live administration than he was on the practice test. This may be for a number of reasons: (1) Ben expressed dislike for the practice test topic and he did not express the same degree of dislike for the live administration, but he did not show any partiality towards the topic either, (2) he could have taken the live administration more seriously than the practice test, or (3) Ben may have felt better writing a complete test and commenting retrospectively after having a couple of days of reflection rather than commenting intermittently on an incomplete test. An examination of the differences and similarities between the strategies that Ben discussed reveals that he reported using 11 strategies that were different from the ones he reported in the pre-test interview. On the whole, he modified his test taking behavior the most out of the three test takers because he only recalled using four of the same strategies from the pre-

test interview. The strategies are listed below and are separated into Millman et al.'s (1965) test-wisness strategies. They are identified by the letter of the skill that was tested in square brackets at the end of each strategy.

Time Using Strategies

- Star the reading and listening [W]
- Skim and scan [R]
- Focus the argument [W]

Error Avoidance Strategies

- Draw a line to the correct answers [L]
- Underline difficult/unknown words [L]
- Plan out what you are going to say [S]
- Critically analyze the prompt [W]
- Underline long and difficult words [S]

Deductive Reasoning Strategy

- Use the reading to answer the listening [L]
- Use the 1st reading to answer questions in the 2nd reading [R]
- Use the listening and reading questions to write the essay [W]

The error avoidance strategy of underlining the difficult or unknown words could be considered a non-test-wisness strategy depending on what Ben did with the words during the listening. If he underlined the unknown words and was listening for only those words he may have mistaken what was said and would be unable to write the answer down, as he almost did in the pre-test interview. In addition, Ben made use of two deductive reasoning strategies that were not discussed in the classroom observations, so it seems that the strategies were his own. By using these two strategies he could be demonstrating a high level of test taker sophistication.

Despite the fact the number of strategy accounts for Charles was only two less than was coded for the pre-test interviews he seemed less descriptive overall, especially

for the reading section. When he was asked what he did during the different sections he could reiterate his answer, but he seemed to have difficulty recalling how he chose to answer the particular question. This may be attributable to the different interview formats: in the pre-test interview there was the practice test to refer to, so he could show me what he was doing or point to sections, while the post-interviews were centered on the test takers' recall. He also said that he felt that it was hard to hear the recording in the live administration.

Charles recalled using six different strategies in the post-test interview than the ones in the pre-test interviews. The strategies are listed below and are separated into Millman et al.'s (1965) test-wiseness strategies. The strategies are also identified by the skill tested with square brackets at the end of the strategy and the letter of the skill in the middle.

Time-Using Strategy

- Skim and scam [R]
- Leave space to write the introduction [W]
- Use the time to circle dates, books and professor's name [S]

Error Avoidance Strategy

- Pronounce the words out loud in your head [S]
- Underline long or difficult words [S]

Deductive Reasoning Strategy

- Use the 1st reading to answer questions in the 2nd reading [R]
- Use subtitles to answer questions [R]

The primary strategies that Charles seemed to report drawing on were time-using strategies and deductive reasoning strategies. Furthermore, the strategies that are different from the ones discussed in the preparation course are: to leave space and write the introduction, to use the 1st reading to answer questions in the 2nd reading, and to use the

subtitles to answer questions. The last strategy could also be considered a time-using strategy. Therefore, it is reasonable to think that due to the pressures of being in the high-stakes timed event Charles modified his practice test taking behavior to draw on other resources in order to answer as many questions as possible within in the designated time period. In addition, the speaking strategy of circling the professor's name and dates that he adopted could have been part of his typical practice test-taking behavior when completing that task because it was not part of the pre-test interview. Furthermore, it is unknown whether the writing strategy was his typical practice since he did not write the essay component in the pre-test interview. Though Charles was less descriptive he was the test taker who seemed to modify his behavior the least since he described using 6 different strategies out of 15.

Alexandra had the most strategy accounts out of the three test takers interviewed for this section. During the pre-test interview she seemed most concerned with the fact that the reading answers were in sequence and mentioned that she would re-read paragraphs if she was reading too far ahead from where she found the previous answer. In addition, all three test takers were more descriptive about the speaking section. This may be due to the fact that it was the last section that the test takers completed and it was the first part of the interview, so the test takers might have been able to recall with relative ease that portion of the test. Alexandra recalled using 14 different strategies from those observed in the post-test interviews:

Time-Using Strategy

- Manage time for questions [R]
- When 10 mins are left, leave space and write the conclusion [W]

Error-Avoidance Strategy

- Place arrows to show the order of the sentences [W]
- Underline long and difficult words [S]
- Draw a line to the correct answers [L]
- Read the newspaper out loud at home before the test [S]
- Underline difficult or unknown words [L]

Guessing Strategy

- Recognize previously identified keywords [L]

Deductive Reasoning Strategy

- Use own information to answer questions [L]
- Use subtitles to answer questions [R]
- Use the listening and reading [W]

Non-Test-Wiseness Strategies

- Use time in the writing to answer the reading [R]

An examination of the strategies that Alexandra recalled using in the post-test interviews reveals two contradictory strategies. She reflected using the time-using strategy of managing her time, but then also said she used the writing time to go back and answer the reading questions, which demonstrates an unwise test taking strategy. Since the writing portion of the CAEL Assessment is weighted the heaviest, all of the time provided should be focused on the essay and not on other sections of the test. Furthermore, if the test takers do not receive a bandscore of 30 or higher on the writing they will not be eligible to receive a total bandscore of 30 or higher in the overall test. Given that Alexandra wanted to get a bandscore of 40 or higher to begin her concurrent studies she should have realized the importance of the writing section. On the other hand, she could have overestimated her ability of writing the essay in the time allotted considering that it was her first time taking the CAEL Assessment, which is most likely the case since she indicated she was unable to complete her essay. Alexandra also drew

on error avoidance strategies and deductive reasoning strategies that were not discussed in the preparation course which could also be a result of being pressed for time and demonstrating the test-wiseness of not leaving questions empty.

An analysis of the post-test interview strategies as language use strategies reveals that the test takers recalled using more resourcing strategies than previously observed. The strategies are summarized in Table 11.

Table 11 Language Use Strategies Recalled During the Post-Test Interviews

<i>Language Use Strategies</i>
<p><u>Planning:</u></p> <ul style="list-style-type: none"> ◦ Focus the argument first before brainstorming [W] ◦ Use the time in the writing to answer the reading [W] ◦ Leave space to write the introduction [W] ◦ Plan out what you are going to say [S] ◦ When 10 mins are left, leave space and write the conclusion [W] <p><u>Problem Identification:</u></p> <ul style="list-style-type: none"> ◦ Critically analyze the writing prompt (is it argumentative, descriptive, etc.)* [W] ◦ Underline difficult/unknown words [L] ◦ Underline long or difficult words [S] <p><u>Self-Monitoring/ Production:</u></p> <ul style="list-style-type: none"> ◦ Go over answers at the end: [L] <ul style="list-style-type: none"> - To check spelling ◦ Pronounce words in head out loud [S] ◦ Read newspaper out loud at home before the test [S] ◦ Manage time for questions <p><u>Selective Attention:</u></p> <ul style="list-style-type: none"> ◦ Underline keywords [R] ◦ Place a star (asterisk) in the reading text, reading questions or listening questions to go back to when writing the essay [W] ◦ Skim and scan for answers [R] ◦ Recognize previously identified keywords [L] ◦ To look for questions in sequence [R]

Selective & Directed Attention:

- Be aware of the directions (e.g. to know how many answers to check for a multiple choice question) [L]
- Read questions before reading the text [R]
- Use the time to circle dates, books and professor's name [S]

Resourcing:

- Integrate reading questions and answers into the essay [W]
- Integrate the listening questions and answers into the essay [W]
- Use subtitles to answer questions [R]
- Use the 1st reading to answer questions in the 2nd reading [R]
- Use own information to answer questions [L]

Note Taking:

- Write out 1 or 2 keywords before speaking [S]

Creative Elaboration:

- If you don't know what to say, then make up the information [S]

Repetition:

- Memorize and practice your answers to the two possible questions [S]

In investigating the strategies recollected by the test takers in the pre-test interview, it is evident they discussed numerous language use strategies. In fact, only two of the strategies discussed in the interviews are not language use strategies according to O'Malley and Chamot's (1990) classification. They are: draw a line to the correct answers and place arrows to show the order of the sentences. That being said, non-language use strategies are excellent strategies from a test-wiseness perspective and should not be discouraged. Overall, there is an increase in resourcing strategies which are also deductive reasoning test-wiseness strategies and this is most likely the result of the test takers resourcing in a high-stakes, timed event. Furthermore, Alexandra's non-test-wiseness strategy is a planning strategy according to O'Malley and Chamot's framework, so her non-test-wiseness strategy is a language use strategy.

Finally, it appears that different interview formats might be better for different test takers as seemed to be the case with Ben and Charles. The high stakes environment elicited different strategies than that of the practice environment. In addition, Ben exhibited the greatest degree of modification in test taking behavior, while Charles recalled fewer strategies but apparently modified his behavior the least of all. Alexandra demonstrated the most strategy accounts, however, she also reported using a non-test-wiseness strategy that most likely contributed to the fact that she was unable to finish her essay.

4. Post-Test Questionnaires

In the final section of my data analysis and the final section of this chapter, I present the key findings of the questionnaire data and discuss the trends for high and low level test takers who took and did not take a CAEL Assessment preparation course. Furthermore, the two types of questions (Likert and yes/no) are discussed separately to provide a detailed analysis of the data.

Overall there was not a significant relationship between test-takers' total bandscores (i.e. separated into high (40-90) and low (10-30) level test takers) and those who took a CAEL Assessment preparation course²² ($\chi^2 = .343$, $p < 0.558$, 1 df). Nor was there a significant relationship between total bandscores and those who studied *The Test-Taker's Guide to the CAEL Assessment* ($\chi^2 = .009$, $p < .927$, 1 df). However, when the frequency of responses for both high and low level test takers was separated by preparation and non-preparation course the contingency tables revealed some interesting

²² Although I observed only the CAEL Assessment preparation course at the university where this study took place and the instructor of the course was familiar with the way the test was marked, I have included all the CAEL preparation courses reported in this statistic and subsequent ones due to the small number of responses. The CAEL Assessment preparation courses reported were from a private tutor, Canadian college, ESL classes, or China.

trends and behavior²³. To begin with, I will address the questions that had a four point Likert scale (questions 9-12). The high and low test takers' responses of strategy use to the Likert types of questions were separated into percentages and observed counts by whether or not they took a CAEL preparation course as seen in Table 12. The cells that present the most interesting observations are shaded light grey and will be discussed in greater detail below. It should be noted that for analytic purposes the cutoff was set at 10% to reveal any meaningful trends.

²³ For the remainder of this section I will refer to test takers who took the CAEL Assessment preparation course and those who did not as preparation and non-preparation test takers. Furthermore, I will label preparation courses other than the ones for the CAEL Assessment as general preparation courses.

	not at all %(#)		a little %(#)		sometimes %(#)		frequently %(#)	
	L	H	L	H	L	H	L	H
1. Read the instructions first								
Preparation course (n = 46)	0% (0)	5% (1)	29% (7)	23% (5)	38% (9)	18% (4)	33% (8)	54% (12)
Non-preparation course (n = 111)	8% (5)	7% (3)	15% (10)	9% (4)	34% (23)	23% (10)	43% (29)	61% (27)
2. Read the question first [R]								
Preparation course (n = 44)	5% (1)	0% (0)	0% (0)	9% (2)	27% (6)	36% (8)	68% (15)	55% (12)
Non-preparation course (n = 111)	5% (3)	5% (2)	9% (6)	2% (1)	27% (18)	21% (9)	60% (40)	73% (32)
3. Underlined the words								
Preparation course (n = 42)	29% (6)	0% (0)	5% (1)	24% (5)	38% (8)	29% (6)	29% (6)	48% (10)
Non-preparation course (n = 109)	11% (7)	28% (12)	23% (15)	14% (6)	39% (26)	21% (9)	27% (18)	37% (16)
4. Made notes in the margins								
Preparation course (n = 45)	13% (3)	32% (7)	35% (8)	27% (6)	52% (12)	32% (7)	0% (0)	9% (2)
Non-preparation course (n = 102)	27% (16)	48% (20)	38% (23)	19% (8)	21% (9)	21% (9)	8% (5)	10% (4)

Questions 1 (Read the Instructions First) & 2 (Read the Questions First for the Reading)

These two questions refer to selective and directed attention strategies; therefore, it is reasonable to assume that the responses would be similar too. However, the opposite is true. There were no significant differences for a chi-square analysis in the response of preparation or non-preparation test takers ($\chi^2 = 5.353$, $p < 0.148$, 3 d.f.²⁴) or high and lower scorers ($\chi^2 = 5.637$, $p < 0.131$, 3 d.f) for “I read the instructions first before completing the task”. And yet, question 1 had a significant correlation²⁵ using Spearman’s correlation coefficient with total bandscores ($r_s = 0.195$, $p < 0.007$), listening ($r_s = 0.158$, $p < 0.023$), reading ($r_s = 0.243$, $p < 0.001$), writing ($r_s = 0.132$, $p < 0.048$) and speaking scores ($r_s = 0.232$, $p < 0.002$). In addition, both high and low level non-preparation test takers responded that they *frequently* “read the instructions first before completing a task” over both levels of preparation test takers, especially for the low test takers which demonstrated a 10% difference. This unforeseen finding may be the result that preparation test takers felt they were well informed of the instructions from the preparation course, so they possibly decided that they did not have to read them and could go on and complete the task at hand. If this was the case, then the test takers were demonstrating non-test wise behavior. Test takers should always read all of the instructions before completing a task to ensure they complete it correctly, thus, avoiding any preventable errors.

Question 2 elicited a different response from these test takers. The shaded cell, *frequently* respondents, for the “I read the questions first before reading the reading

²⁴ The chi-square analysis was conducted regardless of the fact that one or two of the cells did not have the expected count required to meet the assumptions. Therefore, it may not be possible to rely with confidence on these findings.

²⁵ The correlation statistical tool utilized throughout this section is the Spearman correlation. Though the most popular and often the most powerful statistic is ANOVA, the data set did not meet the assumptions of normal distribution and homogeneity of variance for ANOVA (see Field, 2006).

section” row demonstrated an 18% difference between high level test takers who had not taken a preparation course and those who had. In addition, the *sometimes* cell for high level test takers had a 10% percent increase for preparation test takers over non-preparation test takers. Despite these increased percentages there were no significant differences in a chi-square analysis of the responses of preparation or non-preparation test takers ($\chi^2 = 1.303$, $p < 0.728$, 3 d.f.²⁶) or high and low scorers ($\chi^2 = .792$, $p < 0.852$, 3 d.f) for “I read the instructions first before completing the task”. Neither was there a correlation using Spearman’s correlation coefficient for the total bandscore ($r_s = .61$, $p < 0.307$) or for any of the other scores. Though there were no significant findings, a tentative conclusion could be drawn that preparation courses lead to a modification of responses for the *frequency* category of the questionnaire from high level scorers.

Surprisingly, this increased strategy use from high level preparation test takers’ behavior was not present for low level preparation scorers. These results contradict Tian’s (2000) finding that low level test takers tended to read the questions first more than high level test takers. This may be partly due to the different tests used in the studies: Tian conducted her study with the TOEFL while this study focused on the CAEL Assessment. Nonetheless, more research needs to be conducted to explore in more detail the difference between high and low test takers when it comes to reading the questions before reading the text in light of preparation courses.

Interestingly, the relationship between the strategies “I read the questions first before reading the text in the reading section” and “I underlined words and phrases while

²⁶ The chi-square analysis was conducted regardless of the fact that one or two of the cells did not have the expected count required to meet the assumptions. Therefore, it may not be possible to rely with confidence on these findings.

I was reading the text” demonstrated a significant relationship for all test takers ($r_s = 0.211$, $p < 0.042$). Therefore, those test takers who read the questions before reading the text were most likely to underline the words in the text as well.

Questions 3 (Underlined the Words) & 4 (Made Notes in the Margins)

There was a 10% increase of responses for low level test takers compared to high level test takers for the *sometimes* cells for the question “I underlined the words and phrases while reading the text”. On the other hand, the reverse was true for the *frequency* cells, except the difference was 19% for high level test takers over low level test takers. Though the relationship was not significant ($\chi^2 = 4.508$, $p < 0.212$, 3 d.f.²⁷) there seems to be a difference in the frequency of reporting strategy use for low and high level test takers, which needs to be explored in greater detail in future studies. In addition to the difference between low and high level test takers, there was also an 11% increase between high level preparation test takers and high level non-preparation test takers which is visible in the last shaded cell for question 3. Therefore, the preparation course may have enabled high level test takers to underline more frequently. It is unknown why the test takers underlined words or phrases, or if the purposes for high and low level test takers were different. Perhaps in light of Tian’s (2000) findings that low level test takers focused on more word-level reasoning and high level test takers focused on the meaning, this may have been the case with the test takers in this study.

For the question “I made notes in the margins while reading the text”, the number of responses for the *frequently* category was quite low, which was expected since taking

²⁷ The chi-square analysis was conducted regardless of the fact that one or two of the cells did not have the expected count required to meet the assumptions. Therefore, it may not be possible to rely with confidence on these findings.

notes would be time-consuming and should only occur when needed. Secondly, the percentage of low level preparation test takers is more than 30% higher than the low level non-preparation test takers as seen in the shaded cell for question 4. Hence, it seems that the preparation course modified the low level test takers' responses drastically, while there was only a 10% difference for the high level test takers. This was surprising in view of the previously discussed strategies' frequencies that were higher for high level test takers, especially for those who took a preparation course. One possible explanation is that low level test takers were attempting to translate while reading and making notes of their translations, which is an observation that Tian (2000) made of the low level test takers in her study. In general, both low and high level test takers modified their behavior to some degree as a result of taking a preparation course.

An additional explanation for the overall increase between preparation and non-preparation test takers was perhaps the increased awareness of being permitted to write in the reading booklet in the test that a preparation course might bring. At the beginning of the written test all test-takers were told that they could write in their reading booklets. Despite this, low level test takers may have been too shy to ask for clarification while high level test takers may not or they were more likely to understand what was said. Indeed, Alexandra mentioned that it was important information that she learned as a result of taking the preparation course, since she would probably not be allowed to do that on a test back at home.

Therefore, the differences in results could also be in relation to the understanding of what was permitted and not permitted in the CAEL Assessment. Further to the point, the strategy also had a significant relationship between the question "I used suggestions

from a preparation course” ($r_s = 0.227, p < 0.004$). This related to the fact that preparation courses often inform students of the layout of the test and what is allowed. In this case, test takers were permitted to write in the reading booklet. Moreover, the relationship between “I used the suggestions from a preparation book” and “I made notes in the margins while reading the text” was approaching significance ($r_s = 0.133, p < 0.054$). Thus, the difference between the course and the book may be saliency: the information presented in the course may be more noticeable than what is explained in the preparation book.

Now I will discuss the yes and no questions on the questionnaire and the relationship between high and low level test takers’ responses and of those who took a preparation course and those who did not, which are presented in Table 13²⁸.

Table 13 High and Low Level Preparation and Non-Preparation Test Takers who did and did not Take a Preparation Course

	No %(#)		Yes %(#)	
	L	H	L	H
5. Read all of the text [R]				
Preparation course (n = 47)	64% (16)	46% (10)	36% (9)	55% (12)
Non-preparation course (n = 110)	49% (32)	68% (30)	52% (34)	32% (14)
6. Review answers at the end [R]				
Preparation course (n = 46)	29% (7)	32% (7)	71% (17)	68% (15)
Non-preparation course (n = 111)	50% (33)	30% (13)	50% (33)	70% (31)
7. Wrote answers in point form [R]				
Preparation course (n = 47)	28% (7)	36% (8)	72% (18)	64% (14)
Non-preparation course (n = 106)	38% (24)	38% (16)	63% (40)	62% (26)

²⁸ There were no significant differences found when analyzing the questions in Table 13 with chi-square for those who had taken a preparation course and those who did not, and by low and high test takers, therefore I will only discuss the differences in percentage between the four groups.

	No %(#)		Yes %(#)	
	L	H	L	H
8. Wrote answers in point form [L]				
Preparation course (n = 47)	24% (6)	36% (8)	76% (19)	65% (14)
Non-preparation course (n = 107)	30% (19)	26% (11)	70% (45)	75% (32)
9. Review answers at the end [L]				
Preparation course (n = 47)	28% (7)	9% (2)	72% (18)	91% (20)
Non-preparation course (n = 108)	34% (22)	26% (11)	66% (43)	74% (32)
10. Aware of the essay				
Preparation course (n = 47)	4% (1)	9% (2)	96% (24)	91% (20)
Non-preparation course (n = 112)	16% (11)	18% (8)	84% (57)	82% (36)
11. Wrote essay using listening				
Preparation course (n = 47)	36% (9)	36% (8)	64% (16)	64% (14)
Non-preparation course (n = 110)	46% (30)	43% (19)	55% (36)	57% (25)
12. Planned essay				
Preparation course (n = 42)	29% (6)	14% (3)	71% (15)	86% (18)
Non-preparation course (n = 108)	27% (18)	21% (9)	73% (48)	79% (33)

Questions 5 (Read all of the text), 6 (Review Answers at the End of the Reading), & 7 (Wrote Answers in Point Form for the Reading)

From a test-wisness perspective all test takers should have read the questions first and therefore not read all of the text but skim and scan for answers instead. However, the results did not reflect this, especially for the high level preparation and low level non-preparation groups. There was an 18% difference between high level preparation and non-preparation test takers who responded *no*, with the non-preparation group being the larger of the two. On the other hand, there was a 15% difference between low preparation and low non-preparation test takers who responded *no*, with the preparation group being the larger of the two. It would seem that preparation courses lead

to opposite strategy use by high and low level test takers. Possibly, preparation courses may have taught the high level test takers reading strategies that enabled them to read all of the readings with relative ease. Conversely, preparation courses may have indicated to low level students that it might not be feasible to read all of the text, but rely on other strategies to find the answers, such as searching for keywords or drawing on the subtitles to answer the questions as Charles and Alexandra recalled doing in the post-test interviews. Thus, it seems that preparation courses could have played a major role in these high and low level test taker's responses to read all of the text or not.

For question 6, "review the answers at the end of the reading section", the only group that had a noticeable difference in percentage from the others was the low level non-preparation group, as indicated with the shaded box in Table 8. Therefore, it seems that the preparation course may have benefited the low level test takers for this question, but not the high level test takers. It is plausible that high level test takers have a high enough proficiency level to complete the reading questions and still have enough time to review them at the end, regardless if they take a preparation course or not, which may relate to Tian's (2000) finding that high-level test takers employed more repair strategies than low level test takers. Furthermore, it is interesting to note that the percentage for the low level non-preparation group is roughly the same percentage as seen in the previous question, "I read all of the text", for the *no* category which could indicate an opposing relationship. This seems logical and demonstrates test-wiseness if they are used in conjunction: low test takers may save time by not reading all of the text, thereby, possibly having time at the end to review their questions and ensure that they avoid preventable errors.

Interestingly, for question 7, “I wrote my answers in point form for the reading answers”, there was no difference between the low level and high level preparation and non-preparation groups, but there was a 10% increase in percentage for the low level preparation group compared to the high level preparation group, and a 9% increase over low level non-preparation groups. This seems to indicate that the low level group benefited from the preparation course the most, which could reflect a number of reasons. First of all, from Millman et al.’s (1965) test-wiseness perspective all test takers should use their time effectively. Since it is permitted on the CAEL Assessment to write answers in point form, all test takers should use this strategy, but especially low level test takers. This would have minimized low level test takers’ time spent on each question, and they possibly have had more time at the end to review their answers and avoid preventable errors.

Questions 8 (Wrote Answers in Point Form for the Listening) & 9 (Review Answers at the End of the Listening Section)

Surprisingly, the high level test-takers who took a preparation class had the lowest percentage of the four groups that responded *yes* to the question “wrote answers in point form while listening to the lecture”. From a test-wiseness perspective the high level test takers should have written their answers in point form to save time and to ensure that they did not miss a question²⁹. For question 9, more than 60% of all the groups reviewed their answers at the end. This percentage is rather low considering the test takers were provided with the time once the lecture was turned off and were explicitly told to review their answers. Conversely, 20-35% of 3 out of the 4 groups responded that they did not

²⁹ Responses to the listening section are in real-time. If test takers stop to write out full-sentences, they may miss information necessary to answer subsequent questions.

use the time at the end to review, which displayed extremely poor judgment on their part from Millman's et al.'s (1965) test-wisness perspective. The test takers may have been like Ben who chose not to use the time provided in the pre-test interview, which could have been due to the topic of the test or test anxiety. Moreover, the test takers could have used the time to go back to the previous sections, as Alexandra stated in the post-test interview that she went back to the reading section during the time given for the writing. Additionally, there was a 20% difference between low and high preparation test takers, which was the same difference as between high preparation and non-preparation test takers. Therefore, high level test takers seemed to benefit the most from the preparation course and were able to review their answers better than the low level test takers, perhaps because they were able to remember what was said better.

Questions 10 (Aware of the Essay) & 11 (Wrote the Essay Using the Listening)

For question 10, there was a 10% difference between non-preparation and preparation test takers, whether they were in the high or low level group. Thus, the preparation course seems to have had a positive effect on the test takers' awareness during the listening and the reading section that they would have to write the essay. Unfortunately, it was unknown whether the increased awareness lent itself to noticing elements or questions to write the essay with. It was evident, however, that the preparation course was most likely a source of information regarding the CAEL Assessment and that test takers are permitted to use the listening questions to write the essay.

Question 12 (Planned the Essay)

In the final question discussed, the highest percentage for the *yes* category was for the high level preparation test takers. Therefore, the high level groups appeared to have benefited in planning their essays from taking the preparation course, however, the difference between the preparation and non-preparation high level test takers was only 7%. Therefore, if there was a difference, it was most likely minimal. On the other hand, there was a 15% difference between the low and high level preparation test takers, while there was only a 6% difference between the low and high level non-preparation test takers. Overall, preparation courses seemed to encourage test takers to plan their essays.

In summary, both high and low level test takers benefited from the preparation courses, especially for the strategies of:

- reading the questions first
- underlining words in the reading
- making notes in the margins

Though high level test takers seemed to benefit the most from:

- reading the questions first for the reading
- underlining words in the reading
- reading all of the text
- reviewing answers at the end of the listening
- planning the essay

Alternatively, the low level test takers seemed to benefit the most from:

- reviewing the answers at the end of the reading
- writing answers in point form for the listening

Thus, it appears that high level test takers benefited the most from taking a preparation course. Since I do not consider test takers' previous test experience, more investigation is needed before definite conclusions can be drawn. Nonetheless, the fact that high level test takers benefited more is logical considering Fox et al's. (1997)

findings of high level test takers in their study responding to pre-teaching before taking the CAEL Assessment. Though high level test takers appear to modify their strategies the most, low level test takers did appear to modify their strategy use as a result of taking a preparation course as well, especially for reviewing their answers at the end of the reading and writing their answers in point form for the listening.

In the next and final chapter of this study, I will discuss the limitations and implications of this study and summarize my findings in relation to my research questions.

CHAPTER 5: CONCLUSION

This study investigated the role that preparation, specifically a preparation course, plays in test takers' strategy use and performance. The research was informed by Alderson and Wall's (1993) Washback Hypothesis and their notion that that tests can influence teaching and learning, O'Malley and Chamot's (1990) framework of learning strategies, and Millman's (1965) principles of test-wiseness. To explore the role of preparation in test takers' strategy use, this study examined the strategies discussed in a preparation course, what test takers reported doing on a practice test, what some of those same test takers recalled doing during the live administration of the test, and finally a post-test questionnaire. In this final chapter, I discuss the limitations of the study, briefly summarize the findings and how they relate to the research questions put forth in the introduction. I will then discuss the methodological and pedagogical implications of this study. Finally, I will suggest further research that needs to be addressed, so we can better understand the relationship between preparation courses and test takers' strategy use.

Limitations of the Study

As with any study there are limitations, and this study is no exception. There are several limitations of the study that may have affected its external and internal validity.

Self-Report

The use of self-report and verbal protocols to investigate the conscious processing of an individual is debatable and seen to some as not possible. However, as suggested by Cohen (1998), verbal protocols should be used in combination with other forms of research methods as was done in this study. Furthermore, there is debate about whether conscious mental processes, such as strategies, are obtainable through self-report

methods. However, this study does not claim that the data elicited through the three self-report methods (verbal reports, stimulated recall, and post-test questionnaire) are a direct reflection of test takers' mental process and in doing so I recognize the threat to the internal validity that such methods can bring.

Strategies Considered

The strategies discussed throughout the paper are not a comprehensive list of strategies though they came from multiple data sources. This is because the focus of this study was on the strategies that test takers use while taking the CAEL Assessment. In addition, the verbal report method used in the pre-test interviews was limited by the number of strategies that the test takers could recall and then discuss. I am aware of the fact that the test takers most likely used many more strategies than were reported. Furthermore, the questionnaire used in this study was developed from the course observations and does not claim to be a comprehensive list of strategies used during the CAEL Assessment.

Population Generalizability

This study occurred within a specific educational context and as a result the findings are not generalizable to other student populations or other settings. Moreover, this study was an exploratory study and as a result consisted of four methods, two of which were implied with a small sample size. In addition, the focus of study was on the CAEL Assessment, consequently the strategies used are directed to that particular test. Therefore, I strongly warn against generalizing the findings of this study to any other population or any other high-stakes language test.

In Depth Analysis

The scale of this study was extensive and exploratory in nature. Furthermore, its scope was both its greatest strength and its greatest weakness. On the one hand, I was unable to go into explicit detail of the four skills covered by the CAEL Assessment, which could be treated as four separate studies in the future. On the other hand, the extensive coverage of the study unveiled the key finding that some of the test takers may have resorted to non-test-wisness strategies, which would not have been possible with only the pre-test or post-test interviews. Furthermore, the tracking nature of the methodology utilized in this studied was not previously found in the literature and should be a critical component of future strategy research. In the next section, I will summarize the findings of this study as they relate to its research questions. I begin the section with the main research question, and then follow with the four sub-research questions that correlate to the four sections in Chapter 4.

The Research Questions

The guiding research question for this study was:

What was the role of preparation in test takers' strategy use and test performance?

First of all, preparation appears to play a critical role in test takers' strategy use on a practice test versus a live situation. Secondly, the relationship between preparation and performance is not as straightforward as preparation companies such as Kaplan Inc. or Princeton Review claim them to be. Thirdly, the preparation that high level test takers require may be different from what low level test takers require. This was suggested by the questionnaire data: test takers with higher scores seemed to benefit more than those with lower scores, so it would seem that preparation has the potential to play an important

role in test takers' performance. In addition, the post-test interviews seemed to indicate that low level test takers such as Ben and Alexandra drew on strategies which were not test-wise strategies, but on the practice test they did not report using such strategies. The strategies that they reported using were suggested to them in the preparation course. Therefore, the findings from this study seem to suggest that preparation and test takers' strategy use is extremely complex. Below are the research questions in italics with the summary of findings underneath each question.

What strategies were suggested to the students in a preparation course?

The test takers were exposed to numerous strategies in the preparation course. The highest number of strategies suggested to them was writing, followed by listening and then reading. Surprisingly, the lowest number of strategies suggested were for speaking, which most likely relates to the test's specifications and weighting on the writing versus other skills. Furthermore, the test takers were presented with more language use strategies than test-wiseness strategies overall. In addition, the majority of the strategies suggested were a cross-over between the two strategy descriptions suggesting that the course could possibly provide the opportunity for students to learn. The presence of so many learning strategies suggests that the course provided an environment that supported learning and possibly indicates the presence of positive washback.

Did the observed strategies from the preparation course appear in the pre-test interview with the test takers from the course; and if so, which ones?

In general, the test takers made use of most of the strategies that were suggested to them, with a few exceptions such as Edward's time management strategy.

Interestingly, the majority of the strategies used by the test takers were test-wiseness strategies, which was slightly different than what was presented to them. Two of the test takers would sometimes draw on more than one strategy to answer a question. In addition, Ben and Donald exhibited non-test-wise behavior by not having the appropriate documents. As a consequence, they were unable to write the test. Ben missed the bus and therefore arrived too late to sit for his oral language test, but this was before he took the CAEL Assessment preparation course. However, Donald did not have the appropriate documentation to sit for the test even after he had taken the CAEL Assessment preparation course. The teacher had mentioned to the students how important bringing the appropriate documentation was, but for Donald this point was maybe not stressed enough.

What strategies did the test takers report using from the preparation course and the pre-test interview in the post-test interview? What was different?

The post-test interviews elicited different strategies than what was seen in the practice environment. An obvious difference was the use of deductive reasoning or resourcing strategies, some of which had not been previously seen in the preparation course or in the pre-test interview. Furthermore, Ben, who appeared to be one of the least strategic test takers in the pre-test interview, exhibited the greatest degree of modification in his strategy use. Charles recalled fewer strategies, but did not seem to vary as much as the others. Alexandra recalled using the most strategies; however, she also reported using a non-test-wiseness strategy that most likely contributed to the fact that she was unable to finish her essay. What is most important to draw from the post-test interviews was the difference in strategies reported in the two types of interviews. The difference may have

been due to anxiety about a live test versus a practice test and the test takers' resorting to non-test-wise strategies as a result.

Was there a difference between the strategy use of high and low level test-takers' who took and did not take (or did not take) a preparation course?

On the whole, both levels of test takers seemed to benefit by taking the preparation course, but the high level test takers may have benefited the most, which was logical considering they would have had more language knowledge and therefore were better able to apply the strategies. However, there were some interesting responses from low level preparation test takers compared to non-preparation test takers. This may indicate that different level test takers have different needs when it comes to preparation courses.

Implications of the Study

From a methodological perspective, using language use and test-wiseness strategy frameworks in combination offers meaningful insight into the workings of strategies that are taught in courses or when investigating the test taking process. Another methodological consideration for future studies is the investigation into high-stakes live testing situation as opposed to a non-threatening practice test environment. As was seen in this study, the two can sometimes differ. The final methodological implication of this study is the consideration that different interview formats may vary from participant to participant as seemed to be the case with Ben and Charles. Therefore, it is recommended that whenever possible the two methods be employed to offer the most meaningful insight into strategy use. This methodological approach is fundamental to understanding what test takers actually do on the test. Furthermore, when used together they offer a

better source of triangulation and validity for the results in considering the overall test taking process.

Pedagogically, this study provided insight for teachers who may believe that teaching to the test could be seen as a hindrance to learning. With a test such as the CAEL Assessment, teaching to the test may not be a deviation from teaching learning strategies; the only difference may be that the strategies are applied towards a test rather than a learning task. In addition, I think it might be beneficial to have the students go through the practice tests to better understand the purpose of the strategy. This way, they can decide for themselves whether it will be beneficial for them, which in turn will allow them to realize what strategies best suit them and will be most effective. Generally, preparation courses for the CAEL Assessment could potentially support students' awareness of academic genres that they will encounter in their mainstream classes. This is due to the number of language use strategies suggested in the classroom observations, pre-test interviews and post-test interviews. That being said, every course taught differs from teacher to teacher. Furthermore, preparation teachers should be conscious that the needs of a low level test taker may be very different from those of a high level test taker and the class should be taught accordingly. However, much more research is required before we can know exactly how proficiency affects potential learning and or test taking behavior on the CAEL Assessment.

Finally, additional future recommendations are as follows, but not limited to: 1) the use of a needs assessment of students' weaknesses and strengths in a preparation course (or language support course), 2) an emphasis should not only be placed on the writing in a preparation course but the importance of the essay should be stressed in

relation to the total bandscore, 3) teachers that advise students on effective test behavior (whether or not it is for the CAEL Assessment) should discuss the pressure of time that the students might feel during the live test and then suggest the best strategies to use when they feel overwhelmed, 4) the test specifications for the speaking portion of the CAEL assessment should be examined, whether or not tasks could be included that would elicit a wider range of strategies than was seen in the study, and 5) domain specifications for the CAEL Assessment should be re-assessed to evaluate the weight of the sub-tests and consider whether more weight should be placed on the speaking.

Future Directions

More research needs to be conducted on test takers' strategy use of integrated topic-based tests such as the CAEL Assessment. To better understand test takers' strategy use on the CAEL Assessment a questionnaire should be developed from the pre- and post-test interviews that encompass questions pertaining to resourcing or deductive reasoning strategies that were not previously predicted. Furthermore, it should include questions that relate to Millman et al.'s (1965) framework for test-wiseness, which is typically applied to multiple choice tests. As was seen in this study, however, it can be used for discussing other styles of tests. In addition to test-wiseness, research also needs to be conducted on a longitudinal basis that investigates test takers' strategy use at different intervals during the preparation and testing process (before and after they take the test).

Finally, the future of language testing is not with multiple choice tests; thus, more research should be focused on test takers' strategies for integrated-topic based tests such as the CAEL Assessment. There is so much more to do.

REFERENCES

- Afflerbach, P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, 25, 31-46.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: a study of washback. *Language Testing*, 13, 280-297.
- Alderson, J. C., & D. Wall. (1993). Does washback really exist? *Applied Linguistics*, 14, 115-129.
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal*, 75, 460-72.
- Al-Naddabi, Z. (2006). *ESL students' language learning repertoires in EAP and ESP contexts: Perceived success and pedagogical mediation*. Unpublished master's thesis, Carleton University, Ottawa, Ontario, Canada.
- Aminkova, E. (2004). *Test-takers' perceptions: Behind the construct of a test*. Unpublished master's thesis, Carleton University, Ottawa, Ontario, Canada.
- Bachman, L. (2006a, June). *A research use argument: An alternative paradigm for empirical research in Applied Linguistics*. Paper presented at the 2006 joint AAAL and ACLA/CAAL conference, Montreal, Canada.
- Bachman, L. (2006b). Generalizability: A journey into the nature of empirical research in applied linguistics. In M. Chalhoub-Deville, C.A. Chapelle and P. Duff (Eds.), *Inference and generalizability in applied linguistics: Multiple perspectives* (pp. 165-207). Amsterdam: John Benjamins Publishing Company.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

- Bailey, K. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13, 257-279.
- Becker, H. S. (1958). Problems of inference and proof in participant observation. *American Sociological Review*, 23, 652-660.
- Becker, H. S. (1998). *Tricks of the trade: How to think about your research while you're doing it*. Chicago, IL: University of Chicago Press.
- Brown, A. (2006, July). *Validating the revised IELTS speaking test*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia
- CAEL Assessment. (n.d). *Which educational institutions accept CAEL Assessment scores?* Retrieved May 22, 2006, <http://cael.ca/edu/who.shtml>
- Chamot, A. U. (2005). Language learning strategies instruction: Current issues and research. *Annual Review of Applied Linguistics*, 25, 112-130.
- Chamot, A. U., El-Dinary, P. B., & Robbins, J. (1999). Children's learning strategies in immersion classrooms. *The Modern Language Journal*, 83, 319-341.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11, 38-54.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24, 279-301.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press: Cambridge, UK.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, A. Curtis, Y. Watanabe (Eds.),

- Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Cheng, A. Curtis, Y. Watanabe (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. London: Longman.
- Cohen, A. D. (in press). The coming of age of research on test-taking strategies. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, and C. Doe (Eds.), *Language testing reconsidered*. Ottawa, Canada.
- Cohen, A. D., Weaver, S., & Li, T-Y. (1998). The impact of strategies-based instruction on speaking a foreign language. In A. D. Cohen, *Strategies in learning and using a second language* (pp. 107-156). London: Longman.
- Cumming, A., Grant, L., Mulcahy-Ernt, P., & Powers, D. (2005). A teacher-verification study of speaking and writing prototype tasks for a new TOEFL. *Language Testing*, 21, 107-145.
- Davies, A. (1997). Demands of being professional in language testing. *Language Testing*, 14, 328-339.
- Diamond, J. J., & Evans, W. J. (1972). An investigation of the cognitive correlates of test-wiseness. *Journal of Educational Measurement*, 9, 145-150.
- Douglas, D., & Hegelheimer, V. (2005, July). *Cognitive processes and use of knowledge in performing new TOEFL listening tasks*. Paper presented at the 2005 Language Testing and Research Colloquium (LTRC), Ottawa, Ontario.

- Douglas, D., & Hegelheimer, V. (2006, July). *Learner strategies and knowledge in TOEFL listening tasks: A validation study*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- Eckes, T. (2006, July). *Raters' perceptions of scoring criteria in writing and speaking performance assessments*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *Modern Language Journal*, 87, 222-241.
- Field, A. (2005). *Discovering statistics using SPSS (2nd ed)*. London: Sage Publications
- Fox, J. (2001). *It's all about meaning: L2 test validation in and through the landscape of an evolving construct*. Unpublished PhD thesis. McGill University. Montreal, Quebec, Canada.
- Fox, J. (Ed.). (2002). *Test score and user's manual: The CAEL Assessment*. Ottawa: Carleton University Press.
- Fox, J. (2003). From products to process: An ecological approach to bias detection. *International Journal of Language Testing*, 3, 21-47.
- Fox, J. (2004). Test decisions over time: Tracking validity. *Language Testing*, 21, 437-465.
- Fox, J. (2005). Rethinking second language acquisition requirements: Problems with language-residency criteria and the need for language assessment and support. *Language Assessment Quarterly*, 2, 85-115.

- Fox, J. A., Pychyl, T. A., & Zumbo, B. D. (1993). Psychometric properties of the CAEL assessment, I: Overview of development, format, and scoring procedures. *Carleton Papers in Applied Language Studies*, 10, 1-12.
- Fox, J., Pychyl, T., & Zumbo, B. (1997). An investigation of background knowledge in the assessment of language proficiency. In A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Luoma, (Eds.), *Current developments and alternatives in language assessment: Proceedings of LTRC 96* (pp. 368-383). Jyväskylä, Finland: University of Jyväskylä.
- Fraser, W., & Brisson, M. (Eds.). (2003). *The CAEL assessment test takers' preparation guide*. CAEL and Penumbra Press: Carleton University.
- Frederiksen, J. R., & Collins, A. (1989). A system approach to educational testing. *Educational Researcher*, 18, 27-32.
- Green, R., & Wall, D. (2005). Language testing in the military problems, politics and progress. *Language Testing*, 22, 379-398.
- Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardised achievement test scores and the origins of test score pollution. *Educational Research*, 20, 2-7.
- Hamp-Lyons, L. (1997). Washback, impact and validity: ethical concerns. *Language Testing*, 14, 295-303.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge, UK: Cambridge University Press.
- Jamieson, J. (2006, June). *The value of a mixed-model design in CALL research*. Paper presented at the 2006 joint AAAL and ACLA/CAAL conference, Montreal, Canada.

- Kaplan. (2006). *Why choose Kaplan*. Retrieved August 25, 2006, from <http://www.kaptest.com/repository/templates/Lev4InitDroplet.jhtml?ProductId=&lev4Parent=/www/KapTest/docs/repository/content/College/SAT-Subject-Tests/why-kaplan>
- Kojic-Sabo, I., & Lightbrown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *Modern Language Journal*, 83, 176-192.
- Lazaraton, A., & Taylor, L. (in press). Qualitative research methods in language test development. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, and C. Doe (Eds.), *Language testing reconsidered*. Ottawa, Canada.
- Lemke, J. (1995). *Textual politics: Discourse and social dynamics*. London: Taylor and Francis.
- Liao, Y. (2006, July). *Validating the GEPT: Lexico-grammatical ability as a predictor of L2 listening ability*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- Luxia, Q. (2005). Stakeholders' conflicting aims undermine the washback function of a High-stakes test. *Language Testing*, 22, 142-173.
- Luxia, Q., & Qingsi, L. (2006, July). *Applying the 'diffusion-of-innovation' framework to an examination reform in China*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- McNamara, T. (2000). *Language Testing*. Oxford, England: New York.
- Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? *Educational Measurement: Issues and Practice*, 8, 14-22.

- Messick, S. (1982). Issues of effectiveness and equity in the coaching controversy: Implications for educational and testing practice. *Educational Psychologist, 17*(2), 67-91.
- Messick, S. (1989). *Validity*. In R.L. Linn (ed.) *Educational Measurement* (3rd ed.) (pp. 13-103). Macmillan: New York.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*, 241-56.
- Millman, J., Bishop, C. H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement, 25*, 707-726.
- Myers, D. G., & Spencer, S. J. (2001). *Social Psychology*. Toronto, Ontario: McGraw-Hill.
- Nankatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal, 89*, 76-91.
- Nevo, N. (1989). Test-taking strategies on a multiple-choice test of reading comprehension. *Language Testing, 6*, 199-215.
- O'Loughlin, K. (2006, July). *The use of IELTS for university selection: An Australian study*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge, England: Cambridge University Press.

- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston, MA: Heinle & Heinle Publishers.
- Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL). *System*, 23, 1-23.
- Oxford dictionary of current English* (3rd ed.). (2001). Oxford: Oxford University Press.
- Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*, 20, 26-56.
- Popham, W. J. (1991). Appropriateness of teachers' test preparation practices. *Educational Measurement: Issues and Practice*, 10, 12-15.
- Powers, D. E., & Rock, D. A. (1999). Effects of coaching on SAT I: Reasoning test scores. *Journal of Educational Measurement*, 36, 93-118.
- Purpura, J. E. (1998). Investigating the effects of strategy use and second language test performance with high- and low-ability test takers: a structural equation modeling approach. *Language Testing*, 15, 333-379.
- Purpura, J. E. (1997). An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning*, 47, 289-325.
- Purpura, J. E. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach*. Cambridge: Cambridge University Press.

- Roessingh, H., & Field, D. (2000). Time, timing, timetabling: Critical elements of successful graduation of high school ESL learners. *TESL Canada Journal*, 18, 17-31.
- Rogers, T. W., & Bateson, D. J. (1991). The influence of test-wiseness on performance of high-school seniors on school leaving examinations. *Applied Measurement in Education*, 4, 159-183.
- Rubin, J. (1975). What the “good language learner“ can teach us. *TESOL Quarterly*, 9, 41-51.
- Rubin, J. (1981). The study of cognitive processes in second language learning. *Applied Linguistics*, 2, 117-131.
- Sarnacki, R. E. (1972). An examination of test-wiseness in the cognitive test domain. *Review of Educational Research*, 49, 252-279.
- Saville, N. (2006, July). *A model for investigation the impact of language assessment within a national educational context*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.
- Saville, N., & Hawkey, R. (2004). The IELTS impact study: Investigating washback on teaching materials. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing* (pp. 73-96). Mahwah, NJ: Laurence Erlbaum.
- Shepard, L. A. (1990) Inflated test score gains: Is the problem old norms or teaching the test? *Educational Measurement: Issues and Practice*, 9, 15-22.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Singapore: Pearson Education Limited.

- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing, 13*, 298-317.
- SLALS. (2005). *ESL Credit Courses: CAEL Test Results and What they Mean to Graduate Students*. Retrieved August 10, from:
http://www.carleton.ca/slals/esl/esla/cael_results_graduates.htm
- Smith, M. L. (1990). Meanings of test preparation. *American Educational Research Journal, 28*, 521-542.
- Spolsky, B. (1997). The ethics of gatekeeping tests: what have we learned in a hundred years? *Language Testing, 14*, 242-247.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Tashakkori, A., & Teddlie, C. (2003). Issues dilemmas in teaching research methods courses in social and behavioural sciences: US perspective. *Social Research Methodology, 6*, 61-77.
- Tian, S. (2000). *TOEFL reading comprehension: Strategies used by Taiwanese students with coaching-school training*. Unpublished doctoral dissertation. Teachers College, Columbia University.
- Tsagari, C. (1994). *Method effect on testing reading comprehension: How far can we go?* ERIC ED424768.
- Vongpumivitch, V. (2006, July). *An impact study of Taiwan's general proficiency English test (GEPT)*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia

- Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching: A case study using insights from testing and innovation theory*. Cambridge University Press: Cambridge, UK.
- Wall, D., & Alderson, J. C. (1993). Examining washback: the Sri Lankan impact study. *Language Testing, 10*, 41-69.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research, *Language Testing, 13*, 318-333.
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and Methods* (pp. 129-146). Mahwah, NJ: Laurence Erlbaum.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning, 50*, 203-243.
- Yang, P. (2000). *Effects of test-wiseness upon performance on the Test of English as a Foreign Language*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta, Canada.
- Yu, E. (2006, July). *The impact of test delivery modes and test-taker characteristics on speaking test performance*. Paper presented at the 2006 Language Testing Research Colloquium (LTRC), Melbourne, Australia.

APPENDICES

Appendix I: Description of CAEL Assessment Band Scores (Fox, 2002)

- 90 Expert User:** Demonstrates exceptional competency required for academic English use. Is fluent, accurate, flexible and adaptable in the academic setting.
- 80 Adept User:** Demonstrates high level of competency required for academic English use. Is fluent, accurate, flexible and adaptable in the academic setting.
- 70 Very Competent User:** Demonstrates effective competency in using academic English. Is generally fluent, accurate, and flexible in the academic setting.
- 60 Competent User:** Demonstrates satisfactory competency in using academic English. Minor limitations in fluency, accuracy, and flexibility in the academic setting.
- 50 Competent but Limited User:** Demonstrates a degree of control in using academic English but fluency, accuracy and flexibility are somewhat limited in the academic setting.
- 40 Marginally Competent User:** Demonstrates uneven control in using academic English. Fluency, accuracy, and flexibility are impediments to overall competence in the academic setting.
- 30 Limited User:** Demonstrates constrained competency in academic English use. Noticeable problems in fluency accuracy, and not sufficiently flexible in the academic setting.
- 20 Very Limited User:** Demonstrates severely constrained competency in academic English use. Insufficient fluency, accuracy, and flexibility in the academic setting.
- 10 Non User:** Demonstrates very little ability to use or understand English in the academic setting.

Appendix II: Ethics Certificate



Office of Research Services
5th Floor, Tory Building
1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: (613) 520-2516
Fax: (613) 520-2521
www.carleton.ca/ors

Certificate of Ethics Approval

This is to certify that the Carleton University Research Ethics Committee has examined the application for ethical approval for the research project entitled **Test-takers' Strategies for CAEL Assessment** submitted by **Christine Doe, M.A. candidate** in the **School of Linguistics and Applied Language Studies** and under the supervision of **Professor Janna Fox, School of Linguistics and Applied Language Studies**.

The committee found this project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

This certification is valid for one year from the date indicated below.

- No conditions apply**
 Conditions apply (see below)

Date: 28 October 2005

Leslie J. MacDonald-Hicks
Research Ethics Committee Coordinator
For the Chair of the Carleton University Research Ethics Committee
Prof. Antonio Gualtieri

Appendix III: Post-Test Questionnaire

For Office Use Only

2

Total Bandscore: _____ Writing: _____ Reading: _____ Listening: _____ Speaking: _____

CAEL Assessment Questionnaire

Do not write your name anywhere. Test ID: _____ (will be deleted after the scores are released)

When did you come to Canada: _____ (mm/year) When did you start studying English in Canada: _____ (mm/year)

How long have you studied English _____ (years) Home Country: _____ 1st Language: _____

Sex: _____ (M/F) Age: _____ Where have you studied English? _____ (countries/years) Future University & Major: _____

What did you do to prepare for the CAEL Assessment?

Please answer the following items by circling YES or NO

- | | | |
|--|-----|----|
| 1. I took a CAEL preparation course..... | YES | NO |
| a. If yes, where? _____ | | |
| 2. I took a preparation course for a different test (TOEFL, IELTS, etc) | YES | NO |
| 3. I looked at the strategies/suggestions in <i>The CAEL Assessment Test Taker's Preparation Guide</i> | YES | NO |
| 4. I took a CAEL practice test..... | YES | NO |
| 5. I have taken the CAEL Test before..... | YES | NO |
| a. If yes, how many times? _____ b. What was your last bandscore? _____ | | |
| 6. Please list anything else you felt helped prepare you for the CAEL Test _____ | | |

What strategies/skills did you use during the CAEL Assessment?

Please answer the following items by circling the best response.

- | | Not at all | A little | Sometimes | Frequently |
|---|------------|----------|-----------|------------|
| 7. I used the suggestions from a preparation book | 1 | 2 | 3 | 4 |
| 8. I used the suggestions from a preparation course | 1 | 2 | 3 | 4 |
| 9. I read the instructions before completing each task. | 1 | 2 | 3 | 4 |
| 10. I read the questions first before reading the text in the reading section. | 1 | 2 | 3 | 4 |
| 11. I underlined words and phrases while I was reading the text. | 1 | 2 | 3 | 4 |
| 12. I made notes in the margins while reading the text. | 1 | 2 | 3 | 4 |
| 13. I read all of the text of the two readings..... | | | YES | NO |
| 14. I wrote my answers in point form for the reading questions..... | | | YES | NO |
| 15. I used the time at the end of the reading section to review my reading answers..... | | | YES | NO |
| 16. I wrote my answers in point form while listening to the lecture..... | | | YES | NO |
| 17. I used the time at the end of the listening section to review my listening answers..... | | | YES | NO |
| 18. While doing the reading and listening, I was aware that I would use this information for the essay..... | | | YES | NO |
| 19. I wrote the essay using information from the listening answers..... | | | YES | NO |
| 20. I planned my essay in writing before I wrote it..... | | | YES | NO |
| 21. In the essay, I used information from: (Check one of the following) | | | | |

The Listening & Both Readings Only The Listening Only The Two Readings Only One Reading Other: _____

22. Now that you have finished the test, how do you feel about your performance? Not Good Satisfactory Good Very Good

Appendix IV: Interview Questions for Pre-test Interviews

1. How are you today?
2. Where are you from?
3. How long have you studied English for?
4. What kind of classes/courses?
5. How long have you been in Canada?
6. Have you studied English before somewhere else?
7. What level are you at in the IC Program?
8. Have you taken the CAEL before?
 - a. If yes, what was your overall score?
 - b. Do you remember what was your specific bandscores?
9. When did you take it?
10. Why did you decide to take this course?
11. What score do you hope to get on the CAEL?

Appendix V: Practice Test



Sample Version – Topic: Brain Chemistry

In the CAEL Assessment, you will be asked to do the things you would be expected to do in a university course. You will listen to a university lecture and take notes on a topic. You will read articles which provide you with additional information about the topic and answer questions about them. Finally, you will be asked to write a response to a question about the topic using information from the readings and the lecture.

SUMMARY

TASK	TIME	POINTS
Listening	20 minutes	32 points
Reading 1	15 minutes	13 points
Reading 2	30 minutes	25 points
Writing	45 minutes	Level placement

TOTAL TIME: 1 HOUR AND 50 MINUTES

All of the information in this test relates to the topic of the effect of drugs on the neurological system. You will need information from the lecture and from the readings to answer the following question in the writing section:

How does an understanding of brain chemistry help doctors work with people who have personality disorders?

Listening Section (20 minutes / 32 points) *

During this section of the test, you will listen to a lecture and *answer questions at the same time as you are listening*. The questions follow the sequence of the lecture.

The tape will only be played one time. It will not stop. After the lecture is finished, you will have 5 minutes to review your answers and complete any questions you may have missed.

Before the tape begins, you have three minutes to read the questions in the listening section of this test.

Listening Questions:

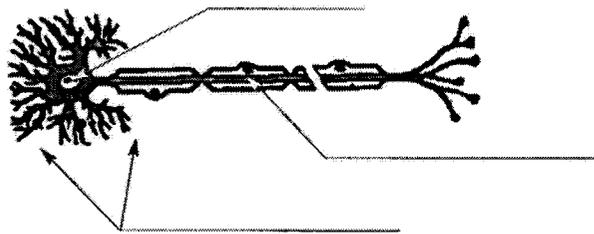
1. What is the topic of today's lecture? Check (3) one. (1 point)

- The brain and behaviour
 The nervous sisters
 The nervous system
 Cellular phones

2. What is another word for nerve cell? (1 point)

3. What is the major purpose of the individual nerve cell? (1 point)

4. Place the following words on the picture below as the lecturer describes the three main parts of the nerve cell: *soma*, *dendrites*, *axon*. (3 points)



GO ON TO THE NEXT PAGE

* For the Answer Key, see pp. 160

5. What are the three main functions of living systems? (3 points)

5.1 _____

5.2 _____

5.3 _____

6. Where do the following neural activities take place in the body? (3 points)

6.1 Information about the world travels to your cortex:

6.2 Command messages are sent to your muscles and glands:

6.3 Decisions are made about what to do with incoming messages:

7. Take notes as the lecturer describes how messages are passed from one neuron to the next. (6 points)

8. Where is a synapse located? (2 points)

GO ON TO THE NEXT PAGE

Reading 2: The Biology of Personality (30 minutes / 25 points)

1. What is the main idea of this article? (1 point)

2. Where do you think this article comes from? Check (3) one. (1 point)

- A newspaper
- An academic journal
- A brochure in a doctor's office
- A magazine

3. What factors other than early experiences might contribute to the development of a personality disorder? (2 points)

3.1

3.2

4. What two technologies can now allow us to understand how mechanisms in the brain determine differences in personality? (2 points)

4.1

4.2

5. Why is it useful to understand some of the biological vulnerabilities someone might have to a particular personality trait? (2 points)

GO ON TO THE NEXT PAGE

6. Complete the chart by providing the relevant information in the blanks: (8 points)

TRAIT	PROBLEM IN BRAIN CHEMISTRY	BEHAVIOUR / PERSONALITY
<ul style="list-style-type: none"> • Impulsivity 	<ul style="list-style-type: none"> • _____ • _____ <p>(1 point)</p>	<ul style="list-style-type: none"> • Acts without thinking • Irritable and gets into fights • Poor relationships • Alcoholism / drug abuse
<ul style="list-style-type: none"> • Emotional reactivity / sensitivity 	<ul style="list-style-type: none"> • Overactive norepinephrine systems 	<ul style="list-style-type: none"> • _____ • _____ • _____ • _____ <p>(3 points)</p>
<ul style="list-style-type: none"> • _____ <p>(1 point)</p>	<ul style="list-style-type: none"> • _____ • _____ <p>(1 point)</p>	<ul style="list-style-type: none"> • _____ • _____ • Suspicious • _____ • _____ <p>(2 points)</p>

GO ON TO THE NEXT PAGE

The Biology of Personality

New research at the Mt. Sinai School of Medicine is showing that the development of personality is not only based on formative life experiences, but on biological factors as well. As reported by Dr. Larry J. Siever in The Frontiers of Pharmacology, this research is leading to a number of new drug treatments for personality disorders.

Everybody has characteristics that make them unique. Some of us are tall, others are short. Some are naturally optimistic and happy, while others have a tendency to be pessimistic and angry.

Although many of us take for granted that our personal appearances are shaped by genetic and biological factors, what you may not realize is that scientific research is demonstrating that genetics and biology have an inherent effect on personality too.

Parents, teachers and influential people in our lives help to shape our early experiences and these moments, in turn, mold our personality. However, advances in modern science have now brought us to the point where we can begin to fully explore what Larry J. Siever, MD calls the "biology of temperament"—that is, the ways in which brain functions determine the manner an individual reacts to his or her environmental surroundings.

"New tools to probe brain chemicals, as well as new imaging techniques to actually visualize how the brain works, allow scientists to get a clearer idea of how individual differences in brain functioning might contribute to differences in personality."

Depending on the temperament that you're born with, you might be predisposed to cope with life's difficulties easily, or you might struggle immensely, reacting to situations angrily or defensively. In the laboratory at the Mt. Sinai School of Medicine, researchers "try to identify and isolate certain traits to discover how they relate to personality and its related disorders and discomforts."

Understanding the biological basis for personality disorders also helps researchers to identify and create medication that is useful in the treatment of those disorders. "While a knowledge of the biology of the brain and its relation to personality will not allow us to predict an individual's behaviour, it may offer us a vocabulary to understand why some people are more prone to outbursts of anger, for instance, while others are painfully shy and fearful of social contact."

As elaborated on by Dr. Siever, in *The Frontiers of Pharmacology*, the program at the Mt. Sinai School of Medicine focuses "on three personality traits—impulsivity, emotional reactivity (which can also be thought of as sensitivity or irritability), and eccentricity—in people whose disturbed thoughts, feelings and / or behaviour are severe enough that they result in problems at work or in their relationships."

Each of these traits can be traced back to specific problems in brain chemistry and these problems, in turn, result in various behaviour and personality disorders. For instance, Dr. Siever explains that impulsivity can be frequently associated with "abnormalities in one of the brain's messenger systems—the

release and uptake of serotonin," which can cause people to "act without thinking." Individuals suffering from impulsivity "may be irritable and get into fights; experience stormy, unstable relationships; or have trouble with alcohol and drugs."

Similarly, individuals with overactive norepinephrine systems are often prone to emotional reactivity and extreme sensitivity. Early studies conducted by Dr. Siever and his colleagues have found that "people with overactive norepinephrine systems may be more likely to be the risk-takers of the world, to be irritable or in other ways overreactive to their environment."

The third personality trait scrutinized by researchers at the Mt. Sinai School of Medicine is eccentricity. According to Dr. Siever, paranoid and schizophrenic individuals usually show an increase in a brain chemical called dopamine, which helps to both set in motion and regulate our thoughts. High levels of dopamine in the brain "may lead people to appear somewhat odd and experience reality in a peculiar fashion." Some may also appear strange or demonstrate suspiciousness towards others' actions and motives.

Research into personality disorders stemming from biological factors such as brain chemistry is occurring at renowned labs and institutes across the world. Not only do scientists want to understand how neurotransmitters like serotonin, norepinephrine and dopamine lead to changes in personality, they also want, and need, to know how to measure levels of neurotransmitters that cause personality disorders.

For example, Dr. Larry Siever and a fellow researcher, Emi Coccaro, have found that the effects of fenfluramine (a serotonin-produc-

ing drug used to treat patients suffering from obesity) can further our understanding of impulsivity disorders. In Dr. Siever's estimation, fenfluramine makes the brain release serotonin, which "stimulates the production of a hormone called prolactin. By measuring levels of prolactin in blood samples taken before and after fenfluramine is given, we may get a reflection of serotonin activity in the brain."

In addition to studies examining serotonin activity in the brain, Dr. Siever and his colleagues have also focused their attention on the neurotransmitter norepinephrine. According to Dr. Siever, norepinephrine "acts on other brain systems to enhance the signal of incoming information." In other words, this neurotransmitter lets the brain know that important information is coming in and "is particularly reactive to events that act as stressors." People with both hyperactive and inactive norepinephrine systems experience either symptoms of impulsive behaviour or depression, respectively.

In his research, Dr. Siever has found the high blood pressure drug clonidine useful when testing the effectiveness of norepinephrine transmitters. Dr. Siever has found that clonidine causes a norepinephrine receptor to increase its production and release a growth hormone. This enables his research team to measure levels of growth hormones in blood samples before the administration of clonidine, as well as after. The comparison allows Dr. Siever and his colleagues to see how well norepinephrine transmitters are functioning.

According to Siever, it has also been discovered that eccentric individuals with low levels of the chemical dopamine experience problems with the way in which they retain information in the brain or in 'working memory'. In his research on dopamine, Siever has ex-

perimented with amphetamines that cause an increase of dopamine production in the system. In order to test whether or not an increase of dopamine in the brain would assist individuals with their ability to process immediate information, Dr. Siever employed the Wisconsin Card Sorting Test (WCST) that he describes as "a kind of solitaire game in which subjects must sort a deck of cards marked with four different colours, numbers, and shapes." During the test, instructions for sorting the cards are continually altered.

The results of the research with WCST illustrate that eccentric personalities improve their performance after their levels of dopamine have increased with the help of amphetamines. It is the hope of Dr. Siever and his fellow researchers that their studies with levels of dopamine may eventually help people who suffer from eccentric personality disorders.

Appendix VI: Semi-Structured Interview Questions – Instructor

1. What do you think is the purpose of the course?
2. Do you think this kind of course is useful?
3. How do you approach the various lesson plans?
4. What do you hope the students will be able to take away from the course?
5. What do you think is a strategy?
6. In class you gave an analogy of driving a car for the difference between what skills and strategies?

Appendix VII: Post-test Interview Questions

Questions about the Oral Language Test (OLT):

1. How are you today?
2. You took the CAEL Assessment this weekend?
3. When was your Oral Exam?
4. Do you remember what room you were in?
5. What was the first question?
6. What was the second task? How did you do that?
7. What was the third task? How did you do that?
8. What about the fourth task? How did you do that?
9. What about for the final task? How did you do that?

Questions about the Written Assessment:

10. What was the day like?
11. Do you remember the room?
12. Where were you sitting?
13. What was the topic of the test?
14. When the monitor read the writing prompt? What did you do?
15. Do you remember what the first reading?
16. Do you remember the second reading?
17. What did you do for the five minutes before the listening?
18. What did you do during the listening?
19. What were you doing at the end of the listening?
20. What did you do for the writing?

Questions about the Preparation:

21. What did you do between the first interview and from when you took the test?
22. Did you do anything differently because of the first interview?
23. What do you think was most helpful from the preparation course?

Appendix VIII: Sample Narrative of Classroom Observations

[Date has been deleted]

Class A

The beginning of the class, there was a one on one discussion about the students writing. The instructor started with the positive aspects of both student's writing. The instructor first went over the CAEL score. However, both students according to the instructor need to find relevant information from readings, but said that they would go over that later in the course.

The instructor then went to talk about the idea of an academic paper format. He stated that had they written in this format they would have gotten a better mark. The idea of an introduction was briefly introduced, according to the instructor the students should agree or disagree with the prompt at this point. The instructor also talked to the students the idea about reader-centered writing versus writer-centered writing. With the reader-centered the reader is the customer and it is the writer's job to make sure that the reader understands is being said in the writing. Whereas with the writer-centered it is the reader's job to try understand what it is that the writer is writing about. The instructor talked about the idea of genre and that there is a certain expectation that readers have when they read something. With academic writing, it is more formal. The instructor asked the students whether they had reader or writer-centered style of writing in the home countries. Two of the students stated that they had writer-centered while the other indicated that his was reader-centered.

At this point, the instructor introduced the idea of reading strategies. He told them to read aloud the strategies listed in the main preparation book on page 16.

The first strategy has to do with titles and subtitles. The instructor asked why is it important to guess what the reading is about?

The second strategy is to read the questions first. According to the instructor, the director will tell them to scrap it. The instructor also mentioned that the preparation book's reading level is really not that of a second language learner at this level. The instructor suggested that it would be better for the students to read the subtitles and captions to begin to understand the reading.

The third strategy is "determine what type of information is required in the answer". According to instructor, this is why it is important to read the whole question. He used the example of how a student in the other class made a mistake because she did not read the question fully before answering.

The fourth strategy is "note if the questions are organized in some way". The preparation book stated that "questions are divided into entitled, for example, "reading for main ideas" or to answer the questions. The discussion in class lead to talking about charts. The instructor told the students that they should look for words "such as" and "the

following”. He also told the students that charts shouldn’t scare them because charts shouldn’t scare them and just to look for keywords and then the lists that follow.

The fifth strategy “check the point value of each question”. The instructor reiterated what was in the preparation book said – that the students should look at the value because this will indicate how much information you have to include in your answer.

The sixth strategy “make your answers brief”. According to the instructor, the students should not write in full sentences. The instructor also gave the example that if all is required from a question is ‘yes’ or ‘no’ then that is all you need to put as the answer.

For the seventh strategy to “review your answers”, the instructor

The class then moved on the last reading for the diagnostic test. They were able to go through six questions before the director of the CAEL assessment came to talk to them.

For the last part of the class, the director of CAEL came to talk to the class. He told them to think about their performance and focus on what went wrong the last time. Also he told them that ultimately they should focus on improving their language and the best way to do that is to practice. Then he went on to talk about the writing scores. He said that if they are to get a 60 or 70 that they must incorporate the listening, but to get a 40 you have to show that you understand the reading. Also it is important that your voice comes out, whereas for a 30 there is no voice and it is more like a formula. In addition, he said that you have to answer the prompt and you should sound like you are interested in it. Regarding the 40 bandscore, the director said that there has to be some level of fluency and vocabulary and structures. He also said that you have to ultimately read the readings and be able to understand them.

Strategies that were observed during this Class:

Reading:

- Look for titles and subtitles (the better understand the reading)
- Determine what kind of information is required in the answer, this is why you should read the whole question
- Look at if the questions are organized in some way
 - charts (look for words like “such as” and “the following”
 - just look for the keywords and then the lists that follow
- Check the point value of each question (this will indicate how much information you have to include in your answer)
- Make your answers brief (not to write in full sentence), also if all is required is ‘yes’ or ‘no’ then write only that
- Review your answers

Appendix IX: Script

Before you leave there is an optional questionnaire for you to complete. This questionnaire is part a larger research project that is being carried out by a master's student. There are two pieces of paper that are stapled together, which have been handed out to you. The top one is the questionnaire, and the bottom is an informed consent form. The questionnaire will be more meaningful when in combination with your score, however, the CAEL testing unit wants to ensure that you remain anonymous. Therefore, if you sign the informed consent, you are agreeing to let the CAEL administration release your test score. Once the scores are released, the CAEL testing unit will write the score on the questionnaire that is attached to the appropriate informed consent form. The questionnaires will be given to the researcher, while the CAEL testing unit will keep the informed consent forms.

Appendix X: O'Malley and Chamot's (1990) Description of Strategies

Planning: Previewing the organizing concept or principle of an anticipated learning task (advance organization); proposing strategies for handling an upcoming task; generating a plan for the parts, sequence, main ideas, or language functions to be used in handling a task.

Directed Attention: Deciding in advance to attend in general to a learning task and to ignore irrelevant distracters; maintaining attention during task execution.

Selective Attention: Deciding in advance to attend to specific aspects of language input or situational details that assist in the performance of a task; attending to specific aspects of language input during task execution.

Self-Management: Understanding the conditions that help one successfully accomplish language tasks and arranging for the presence of those conditions; controlling one's language performance to maximize the use of what is already known.

Self-Monitoring: Checking, verifying, or correcting one's comprehension or performance in the course of a language task.

Problem Identification: Explicitly identifying the central point needing resolution in a task or identifying an aspect of the task that hinders its successful completion.

Self-Evaluation: Checking the outcomes of one's own language performance against an internal measure of completeness and accuracy; checking one's language repertoire, strategy use, or ability to perform the task at hand.

Repetition: Repeating a chunk of language (a word or phrase) in the course of performing a language task.

Resourcing: Using available reference sources of information about the target language, including dictionaries, textbooks, and prior work.

Grouping: Ordering, classifying, or labeling material used in a language task based on common attributes; recalling information based on grouping previously done.

Deduction/Induction: Consciously applying learned or self-developed rules to produce or understand the target language.

Substitution: Selecting alternative approaches, revised plans, or different words or phrases to accomplish a language task.

Elaboration: Relating new information to prior knowledge; relating different parts of new information to each other; making meaningful personal associations to information presented.

Summarization: Making a mental or written summary of language and information presented in a task.

Translation: Rendering ideas from one language to another in a relatively verbatim manner.

Transfer: Using previously acquired linguistic knowledge to facilitate a language task.

Inferencing: Using available information to guess the meanings or usage of unfamiliar language items associated with a language task, to predict outcomes, or to fill in missing information.

O'Malley and Chamot (1990, p. 137-138)

Appendix XI: Millman, Bishop, and Ebel's (1965) Outline of Test-wiseness Principles

I. Elements independent of test constructor or test purpose

A. Time-using strategy

1. Begin to work as rapidly as possible with reasonable assurance of accuracy.
2. Set up a schedule for progress through the test
3. Omit or guess at items which resist a quick response
4. Mark omitted items, or items which could use further consideration, to assure easy relocation

B. Error-avoidance strategy

1. Pay careful attention to directions, determining clearly the nature of the task and the intended basis for response.
2. Pay careful attention to the items, determining clearly the nature of the question.
3. Ask examiner for clarification when necessary, if it is permitted.
Check all answers.

C. Guessing strategy

1. Always guess if right answers only are scored.
2. Always guess if the correction for guessing is less severe than a "correction for guessing" formula that gives an expected score of zero for random responding
3. Always guess even if the usual correction or a more severe penalty for guessing is employed, whenever elimination of options provides sufficient chance of profiting

D. Deductive reasoning strategy

1. Eliminate options which are known to be incorrect and choose from among the remaining options.
2. Choose neither or both of two options which imply the correctness of each other
3. Choose neither or one (but not both) of two statements which, if correct, would imply the incorrectness of the other.
4. Restrict choice to those options which encompass all two or more given statements known to be correct.
5. Utilize relevant content information in other test items and options.

II. Elements dependent upon the test constructor or purpose

A. Intent consideration strategy

1. Interpret and answer questions in view of previous idiosyncratic emphases of the test constructor or in view of the test purpose.

2. Answer items as the test constructor intended.
3. Adopt the level of sophistication that is expected.
4. Consider the relevance of specific detail.

B. Cue-using strategy

1. Recognize and make use of resemblances between the options and an aspect of the stem.
2. Consider the relevancy of specific detail when answering a given item
3. Recognize and make use of specific determiners
4. Recognize and make use of resemblances between the options and an aspect of the stem
5. Consider the subject matter and difficulty of neighboring items when interpreting and answering a given item.

(Millman et al., 1965, p. 711-713)

Appendix XII: Transcription Notation Symbols

From Lazaraton and Taylor (in press)

1. **unfilled pauses or gaps** - periods of silence, timed in tenths of a second by counting "beats" of elapsed time. Micropauses, those of less than .2 seconds, are symbolized (.); longer pauses appear as a time within parentheses: (.5) is five tenths of a second.
2. **colon (:)** - a lengthened sound or syllable; more colons prolong the stretch.
3. **dash (-)** - a cut-off, usually a glottal stop.
4. **.hhh** - an inbreath; **.hhh!** - strong inhalation.
5. **hhh** - exhalation; **hhh!** - strong exhalation.
6. **hah, huh, heh, hnh** - all represent laughter, depending on the sounds produced. All can be followed by an (!), signifying stronger laughter.
7. **(hhh)** - breathiness within a word.
8. **punctuation**: markers of intonation rather than clausal structure; a period (.) is falling intonation, a question mark (?) is rising intonation, a comma (,) is continuing intonation. A question mark followed by a comma (?,,) represents rising intonation, but is weaker than a (?). An exclamation mark (!) is animated intonation.
9. **equal sign (=)** - a latched utterance, no interval between utterances.
10. **brackets ([])** - overlapping talk, where utterances start and/or end simultaneously.
11. **percent signs (% %)** - quiet talk.
12. **carat (^)** - a marked rising shift in pitch.
13. **arrows (> <)** - the talk speeds up, **arrows (< >)** - the talk slows down.
14. **psk** - a lip smack, **tch** - a tongue click.
15. **underlining or CAPS** - a word or SOUND is emphasized.
16. **arrow (--->)** - a feature of interest to the analyst.
17. **empty parentheses ()** - transcription doubt, uncertainty; words within parentheses are uncertain.
18. **double parentheses (())** - non-vocal action, details of scene.