

## **INFORMATION TO USERS**

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
800-521-0600





**The Role of *Contextual Factors* in Mediating the Washback of  
High-Stakes Language Assessments on Learners**

by

Ruth Hungerland

A thesis submitted to the  
Faculty of Graduate Studies and Research in  
partial fulfillment of the requirements  
for the degree of Masters of Arts

School of Linguistics and Applied Language Studies  
Carleton University  
Ottawa, Canada

November 2004

© 2004, Ruth Hungerland



Library and  
Archives Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

0-494-06763-2

Your file Votre référence

ISBN:

Our file Notre référence

ISBN:

#### 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.

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

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

#### 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 protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

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

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

\*\*  
**Canada**

## **Abstract**

Many research studies have emphasized the importance of tests in mediating washback on high-stakes language learning environments. Most of these investigations to date have focused primarily on the influence of the test on teachers, whereas few have examined washback on learners. Furthermore, although these studies have identified a variety of factors that may contribute to washback, few have examined the role of extra-test factors, that is, factors associated not with properties of the test, but with the context surrounding the test. The goal of the current study was to explore washback on learners that may be the result of these potentially important *contextual* factors. Using a grounded theory approach, two sets of one-on-one interviews were conducted with six participants preparing to take the CAEL Assessment at the end of their English for Academic Purposes (EAP) program. The findings revealed two distinct but interrelated categories of washback influence, namely, *test-specific* factors, which appear to reflect the specific properties of a test, and, more importantly, a variety of interconnected *contextual* factors. The latter reflected extra-test factors such as learner attitudes toward language learning, affective reactions, past test experience, and other variables. Based on the findings of this study, a *Contextual Washback Model* was delineated to broadly characterize the combined influence of *test-specific* and *contextual* factors on the quality (positive or negative) of test washback on learners.

## **Acknowledgments**

I thank Professor Janna Fox for her mentorship, critical feedback, and enthusiastic encouragement throughout my thesis work. I am especially grateful for our long-distance phone conversations that never failed to renew my inspiration and provided me with a surrogate academic community. I also thank Janet Benger for her generosity and support that allowed me to complete this study in connection to my work. I must also express my gratitude to the six participants for sharing their experiences, opinions, and emotions with me during this study.

I would like to thank Barbara Burnaby for her feedback and direction during the initial phase of my study. I am also grateful to TESL Canada and TESL NL for allowing me to develop and apply my interest in the field of language testing.

I thank Elizabeth Yeoman for accompanying me on a long, long, walk that served to refresh my theoretical ponderings and Gabrielle Alioth for her generosity of ocean views. I am also grateful to my friends (you know who you are) for essential interludes of delicious food, wine, and engrossing conversation.

I would like to express my deepest gratitude to my husband Peter Gaskovski, for his presence, intellectual challenges, editorial support, and unrelenting encouragement. Finally, I thank my mother, Babette Hungerland, for helping me realize that flowers can be just as important and discussion-worthy as a thesis.

## **Table of Contents**

<b>Chapter 1: Introduction.....</b>	<b>1</b>
<b>Chapter 2: Literature Review.....</b>	<b>4</b>
Definitions of Washback.....	5
Review of Washback Studies.....	12
Washback on Teachers.....	16
Washback on Learners.....	19
New Models for Washback.....	23
Methodological Issues in Washback Research.....	26
Summary.....	27
<b>Chapter 3: Research Methodology.....</b>	<b>29</b>
The Methodology of the Current Study: Overview.....	29
Context.....	30
Role of the researcher.....	31
The EAP program.....	32
Students in the EAP program.....	33
Teachers in the EAP program.....	34
The test.....	35
The use of test results.....	36
Participants.....	38
Instruments and Procedures.....	39
Analysis.....	41
<b>Chapter 4: Results and Discussion.....</b>	<b>42</b>
<b>Test-specific Washback Factors.....</b>	<b>43</b>
Pre-test responses:	
Test knowledge.....	43
Test format.....	48

Test difficulty.....	49
Test-specific attitudes.....	51
Post-test responses:	
Test-taking experience.....	54
<b>Contextual Washback Factors.....</b>	<b>57</b>
Pre-test responses:	
Attitudes toward English.....	58
Goals.....	60
Attitudes toward the language learning environment.....	65
Test practice in the classroom.....	67
Teacher role in test practice.....	69
Post-test responses:	
Test preparation strategies.....	71
Parental expectations.....	74
Affective reactions.....	75
Plan of action.....	78
Results.....	82
Fairness.....	85
Fairness for others.....	86
Fairness of the testing situation.....	87
Advice.....	91
Summary.....	94
Introducing the Contextual Washback Model.....	99
<b>Chapter 5: Conclusions and Implications.....</b>	<b>107</b>
References.....	111

## **List of Appendices**

Appendix A.....	Student Interview 1: Guiding Questions
Appendix B.....	Student Interview 2: Guiding Questions
Appendix C.....	CAEL Assessment: Sample Version

## **List of Figures**

Figure 1	Washback Stimulus-Response model.....	7
Figure 2	Synthesis of Alderson and Wall's (1993) 15 Washback Hypotheses.....	9
Figure 3	Bailey's (1996) basic model of washback.....	11
Figure 4	Henrichsen's Hybrid Model of the Diffusion/Implementation Process....	25
Figure 5	Contextual Washback Model.....	101

## **Chapter 1: Introduction**

“The CAEL is a very, very good test. It is better than other tests such as the TOEFL. I can learn some true and useful things; when I enter university I can learn easily. I think the CAEL is important, because if we enter university directly we don’t know how to listen to lectures and will fail credit courses.”

(Participant A)

“The CAEL exam is not as easy as you think. If it is possible go back to China and do the TOEFL.”

(Participant F)

The influence of high-stakes language assessments on the learning environment has received much attention in educational and testing communities during the past two decades. McNamara (2000) draws a distinction between test “washback”, which operates at the micro level and refers to the effect of tests on teaching and learning in classrooms, versus test “impact”, which operates at the macro level and refers to the wider influence of tests on educational systems and society. Numerous studies have identified a variety of factors that shape positive or negative washback on teachers (Burrows, 1999; Alderson and Hamp-Lyons, 1996; Watanabe, 1996; Cheng 1999, 2004), and learners (Shohamy, Donitsa-Schmidt, and Ferman, 1996; Watanabe, 2002; Bailey, 1996). Other studies have focused on test impact on educational systems (Alderson and Wall, 1993; Cheng, 1997; Chapham and Snyder, 2000; Qi, 2004). The findings of these investigations have revealed the complexity of washback processes, and also contributed to increasing understanding of the phenomenon (Wall 2000; Watanabe, 2004). However, most studies to date have

focused primarily on the influence of the test, particularly the washback of the test on teachers, whereas few have examined the role of learners and the test context. The quotes introduced above illustrate that the same test may influence learners differently, suggesting that extra-test factors may play a significant role in mediating washback.

The goal of the current study was to explore washback on learners that may be the result of important contextual factors, such as learner attitudes toward language learning, affective reactions, past test experience, and other variables that shape learner perceptions within a high-stakes testing situation. In order to explore such extra-test factors, two one-on-one interviews were conducted with six participants preparing to take the Canadian Academic English Language (CAEL) Assessment at the end of their English for Academic Purposes (EAP) program. The first set of interviews was held at week seven of the twelve-week program and the second set was held after the completion of the assessment at the end of the term. The present study used a grounded theory approach (Strauss and Corbin, 1998), through which themes and categories of factors influencing participants' perceptions and attitudes toward the test emerged. The findings revealed two distinct but interrelated influences that played a significant and complex role in mediating washback, namely, *test-specific* and *contextual* factors. Based on the findings of this study a *Contextual Washback Model* was proposed to broadly characterize the combined influence of *test-specific* factors and *contextual* factors on the quality (positive or negative) of test washback.

This work is presented in five chapters. The literature review (Chapter 2) presents a short overview of washback models and definitions, investigations of washback on teachers and learners, as well as recent models of washback research. In Chapter 3,

methodological issues are examined by situating the present study in the wider framework of washback literature. In Chapter 4, the qualitative findings of participants' perceptions and attitudes toward their test experience are discussed. Finally, in Chapter 5 these findings are integrated into a conceptual model and the implications and limitations of the study are discussed.

## **Chapter 2: Literature Review**

“And when any of the fugitives of Ephraimites said, “Let me go over,” the men of Gilead said to him, “Are you an Ephraimite?” When he said, “No,” they said to him, “Then say Shibboleth,” and he said “Sibboleth,” for he could not pronounce it right; then they seized him and slew him at the fords of the Jordan.”

(Judges, 12: 5-7)

Spolsky (1997) refers to this “shibboleth test” as the original example of the use of examinations for power and control. This quote also illustrates an extreme instance of a high-stakes test, given the dire consequences associated with test failure. Throughout history tests have been used to prove capabilities, establish credentials, and, after the 17th century introduction of the examination system to Europe by the Jesuits (Madaus, 1988), as a means of controlling school curricula and classrooms. Concern about the influence of tests appears in the literature as early as 1877, when Latham (cited in Cheng, 1999) characterized examinations as an “encroaching power” taking over the syllabus and directing what happens in the classroom.

Traditionally, testing has been a scientific field and the creation of quality tests as well as the accurate measurement of knowledge have thus been of central concern (Shohamy, 2001). During the past two decades, however, language testers have begun to examine the forms and practices of language testing and its connection to teaching and learning practices (Messick, 1989, 1996; Shohamy, 1993, 1997, 1998; Bailey, 1996; Alderson and Wall, 1993). As a result, new research perspectives in language testing have emerged, which go beyond the focus on test design, isolated performances, and the individual, by taking into consideration the societal, systemic, and social impact of tests. At the core of

this paradigm shift in language testing, Alderson (1986, cited in Cheng and Curtis, 2004) identified *washback* as a distinct phenomenon needing further investigation.

The following discussion examines the evolution of the washback phenomenon during the past two decades, reviews pertinent empirical research on the impact of tests on educational systems, teachers, and learners, and addresses the current methodological and conceptual difficulties encountered in the investigation of test washback.

## **Definitions of Washback**

A widespread assumption within the language testing literature holds that tests exert a powerful influence on both teaching and learning behaviours. This influence is commonly referred to as “washback” and numerous definitions have been offered in the educational and testing literatures. Shohamy, Donitsa-Schmidt, and Ferman (1996) simply define washback as “the connections between testing and learning” (p. 298). Buck (1988) explains the concept as “a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to students” (p. 17). Shohamy (1992) goes further in defining washback as “the utilization of external language tests to affect and drive foreign language learning” and emphasizes that “this phenomenon is the result of the strong authority of external testing and the major impact it has on the lives of test takers” (p. 513). Adding to the complexity of washback, Bachman and Palmer (1996) assert that the influence of a test operates both on a micro level (e.g., individual students and teachers) and a macro level (e.g., society and educational systems); consequently, assessments should be evaluated in relation to the

external variables of society's goals and values, test use within the educational system, as well as the potential outcomes of the test in the classroom.

Several different terms are used to refer to washback across disciplinary literatures, including *measurement-driven instruction*, *curriculum alignment*, and *systemic validity*. Outlining the traditional notion of *measurement-driven instruction*, Popham (1987) addresses the relationship between instruction and assessment in which tests act as powerful "curricular magnets" by directing teachers' attention to the content of test items. He argues that if tests are "properly conceived and implemented", then focusing teaching on what they assess is a positive activity. In research related to *curriculum alignment*, performance-based assessments are linked closely to the goals of instruction (Madaus, 1988; Andrews, 1994). The notion of *systemic validity* refers to the integration of assessments into the educational system at all levels and the need to demonstrate that the introduction of a new test can actually improve learning (Frederiksen and Collins, 1989; Fullilove, 1992).

Several authors have linked washback to considerations of test validity (Hughes, 1988; Frederiksen and Collins, 1989; Weir, 1990). In their search for the occurrence of washback, Alderson and Wall (1993) found that "some writers have even gone so far as to suggest that a test's validity should be measured by the degree to which it has a beneficial influence on teaching" (p. 116). For example, Messick (1996) discusses "washback as an instance of the consequential aspect of construct validity, linking positive washback to so-called authentic and direct assessment" (p. 241). Although the role of washback in test validity has become a consideration in test design and evaluation (Green, 1985; Wesche, 1987; Morrow, 1991; Andrews, 1994; Swain, 1995), Alderson and Wall (1993) caution

that the complex nature of washback cannot be directly related to a given test's validity, but that forces such as society, education, and schools likely shape the nature of washback. Importantly, they observe that the quality of the washback effect may be independent of the quality of the test. Thus, any test, good or bad, can be said to have beneficial or detrimental washback.

Burrows (1999) notes that there has been a consistently held view that the introduction of a test necessarily causes changes in teachers and their practices as well as in learners and their learning behaviours. The view that a test causes changes in teaching and learning can be illustrated in a stimulus-response model (see Figure 1, Burrows, 1999). While some researchers view washback as a negative force (Smith, 1991; Andrews and Fullilove, 1994; Noble and Smith, 1994) others believe that it carries potential benefits (Hughes, 1988; Morrow, 1991). This belief in turn has led to the notion that it is possible and desirable to engineer positive washback (Pearson, 1988; Swain, 1985; Bailey, 1996; Hamp-Lyons, 1997).

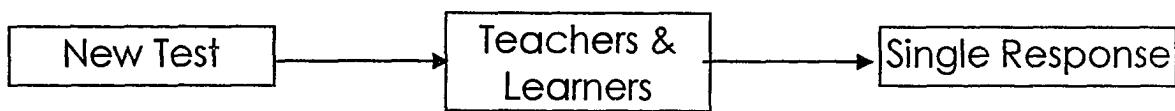


Figure. 1. Washback Stimulus-Response model (Burrows, 1999)

Despite the widely held beliefs regarding the occurrence of washback, most publications before the 1990s were based on anecdote, perception, assertion, or pronounced intentions to create positive washback, but provided no evidence that it had

occurred (Wall, 2000). For example, Alderson and Wall (1993) observe that the concept of washback, and the mechanism by which it operates, has been documented in only a handful of empirical case studies conducted in the Netherlands (Wesdrop, 1982), Turkey (Hughes, 1988), and Nepal (Khaniya, 1990). Similarly, in an exhaustive literature search Watanabe (1996) finds only 10 empirical studies, in contrast to over 500 assertions, about the washback of university entrance exams in Japan.

Alderson and Wall (1993) note that existing empirical studies show little effect of washback, and are incomplete or inadequate because they do not provide observational data and consequently, lack investigation into “what actually changed in class” (p. 126). Consequently, they conclude that further research on washback is needed and suggest that such investigations must include classroom observations in order to incorporate an ethnographic perspective into the identification of particular features, characteristics, and contexts that account for washback. Furthermore, such research must aim to develop “increasing specifications of the Washback Hypothesis” (p. 127), as well as integrate findings in the areas of motivation, performance, and educational innovation. Figure 2 is a visual representation of Alderson and Wall’s (1993) Washback Hypotheses, which are generally presented as a list of fifteen statements on who is affected by a test as well as how a test influences teaching, learning, and attitudes. It illustrates the hypotheses that an important test may influence some or all teachers and learners, qualities of teaching and learning, and attitudes concerning method and content of teaching and learning.

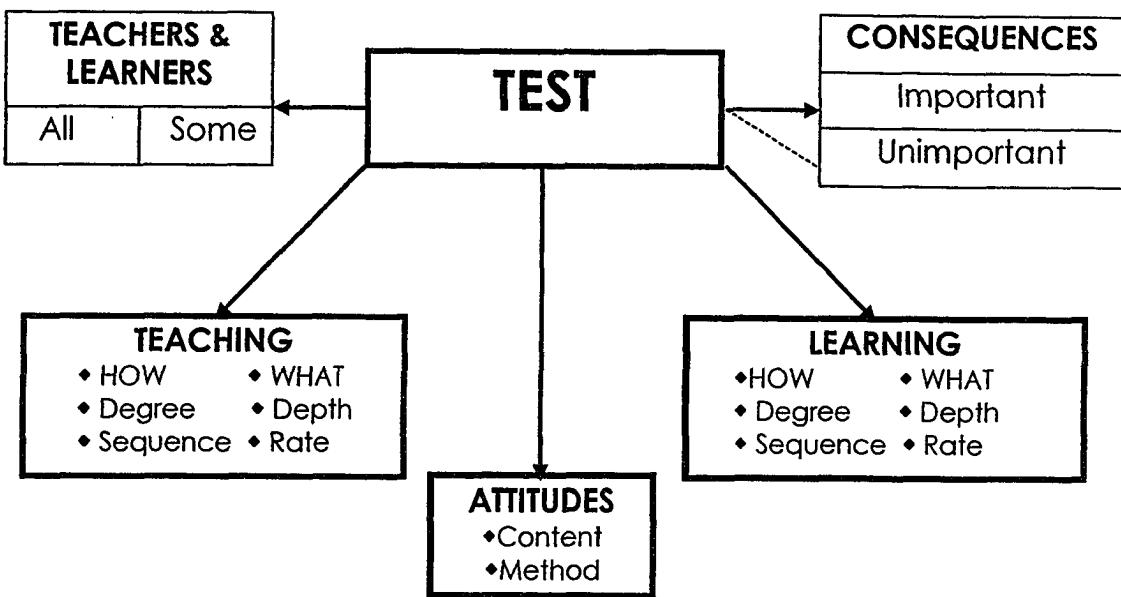


Figure 2. Synthesis of Alderson and Wall's (1993) 15 washback hypotheses.

Alderson and Wall's hypotheses expand the washback Stimulus-Response model (see Figure 1, Burrows, 1999) by identifying various possibilities for responses to a test. However, as Figure 2 illustrates, these hypotheses are unidirectional and are focused on the test as a central force in affecting teachers and learners and their respective behaviours. Furthermore, the hypotheses do not express the potential for interaction among the various participants and processes involved in the testing situation.

Adding to the theoretical discussion, Hughes (1993, cited in Bailey, 1996) proposed a trichotomous model, distinguishing between washback to the *participants*, the *process*, and the *products* of an educational system. In this framework, *participants* include language learners and teachers, administrators, materials developers, and publishers, "all of whose perceptions and attitudes toward their work may be affected by a test" (p. 2).

The term *process* covers any actions taken by the participants, which may contribute to the process of learning, including materials development, syllabus design, changes in teaching methods or content, learning and/or test taking strategies. Finally, *product* refers to what is learned (facts, skills etc.) as well as the quality of learning.

The nature of a test may first affect the perceptions and attitudes of the participants towards their teaching and learning tasks. These perceptions and attitudes in turn may affect what the participants do in carrying out their work (process), including practicing the kind of items that are to be found in the test, which will affect the learning outcomes, the product of that work. (Hughes 1993, cited in Bailey, 1996, p. 262)

Bailey (1996, 1999) uses Hughes' framework to develop a basic model of washback (see Figure 3), in which the test directly influences the participants engaged in various processes, which result in products connected to each category of participants. The dotted lines represent possible influences on the test, an effect van Lier (1988) termed "washforward". Although participants and products are recognized as possible factors affecting the test, the model assumes that the test to be the central force in the washback process. In addition, the model does not allow for the possibility that the processes that interact between the participants and the products may also play a role in mediating washback.

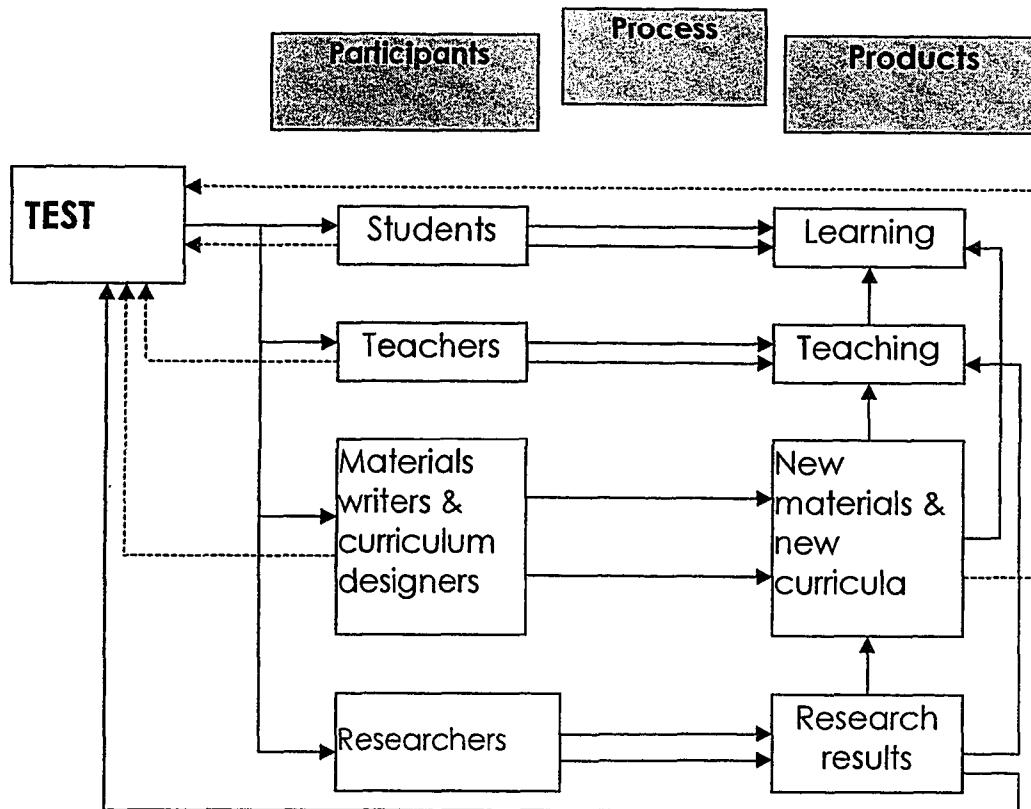


Figure 3. A basic model of washback (Bailey, 1996).

Bailey (1996) integrates Hughes' (1993) framework with Alderson and Wall's (1993) list of washback hypotheses. In particular, she makes a further distinction between "washback to the learner" (p. 264), which results from supplying test-related information to the test-takers, versus "washback to the programme" (p. 266), which results from providing test-related information to all of the other participants in the education system. Bailey also observes that the paucity of empirical research on washback can be attributed in part to the complexity of the concept itself and, in particular, to the difficulty of isolating washback from most other properties of teaching and learning, on the basis of a non-random sample of subjects, and deducing what can be evidentially linked to the test (see also Messick, 1996).

## **Review of Washback Studies**

Theoretical discussions of washback are complemented by empirical studies documenting the washback of different types of tests employed in different cultural settings. The Sri Lankan impact study (Wall and Alderson, 1993) is considered a landmark investigation, in part because one of its key characteristics was method triangulation, involving the careful observation of teacher behaviour. The authors were commissioned to investigate the impact of a new examination and textbook series, intended as “levers for change” (p. 45), on English language teachers and classrooms. The findings indicated that, although there was evidence of washback on the content of teaching, there was no evidence of washback on methodology. The authors therefore suggested that “the examination does not determine *how* teachers teach, however much it may influence *what* they teach” (p. 127). Furthermore, the researchers identified several factors apart from the examination that shaped the form of teaching, including: teachers’ understanding of the nature of the exam; lack of adequate training; and (perceived) ability to implement changes in methodology. Commenting on the value of observations in facilitating the understanding of interview responses, raising new questions, and revealing actual classroom practices, Alderson and Wall (1993) conceded that although the findings were useful, they did not reveal “*why* teachers do what they do” (p. 62). Furthermore, the authors uncovered important limitations in using exams to bring about systemic change: “We now believe an exam on its own cannot reinforce an approach to teaching the educational system has not adequately prepared its teachers for” (p. 67).

In an examination of the washback effect over time, Shohamy, Donitsa-Schmidt, and Ferman (1996) observed the effects of the introduction of two new Israeli national tests, namely, Arabic as a Second Language (ASL), and English as a Foreign Language (EFL). The washback of these tests had previously been investigated by Shohamy (1993), who reported that both exams influenced teaching and learning activities; in particular, the tests directed attention to areas that had previously not been taught. She observed that during the preparation period for the ASL test, class-activities became “test-like”, the classroom atmosphere became tense, and that “students and teachers appeared highly motivated to master the materials” (p. 301). Interestingly however, such teaching and learning behaviours ceased abruptly after the administration of the test. The author also found significant differences between experienced teachers, who used the test as a main source of guidance for teaching oral language, and novice teachers, who employed a variety of additional oral activities. Revisiting the washback of the two tests three years later, Shohamy, Donitsa-Schmidt, and Ferman (1996) observed that the effect of the ASL exam had decreased substantially, whereas the washback of the EFL exam had increased over time. The authors identified several factors that contributed to changes in the washback effect, including the stakes of the test, status of the language being assessed, purpose of the test, format, as well as the skills tested.

In their investigation of TOEFL preparation courses in the United States, Alderson and Hamp-Lyons (1996) aimed to explore the “folklore” of the negative effect of TOEFL on classroom content and teacher methodology. Through student and teacher interviews as well as classroom observation, they hoped to gain an understanding of how teachers feel about preparing students for the test, while also obtaining an inside look into test

preparation classrooms. The authors learned that most teachers have a negative attitude towards the TOEFL and to teaching an “inauthentic and non-communicative test” (p. 285). Further, they found patterns of difference between TOEFL and non-TOEFL classes, including fewer digressions, less student questioning, more teacher talk, less pair-work, and less laughter in TOEFL classes. However, Alderson and Hamp-Lyons also concluded that, although TOEFL affects both what and how teachers teach, “the effect is not the same in degree or kind from teacher to teacher” (p. 295). Most interestingly, they stated that it could be tempting to infer that the test alone does not cause washback but rather that stakeholders themselves cause the washback. The authors’ findings led them to suggest an important qualification to the Washback Hypotheses (Alderson and Wall, 1993), namely that “the existence of a test by itself does not guarantee washback, either positive or negative” (p. 296).

In a large-scale quantitative and qualitative empirical study Cheng (1997) aimed to explore which stakeholders and areas of teaching and learning would be most affected by the change of a high-stakes test within the Hong Kong secondary school system. The author introduced the term “washback intensity” to refer to “the degree of washback effect in an area or a number of areas of teaching and learning affected by an examination” (p. 43). Cheng reports findings similar to Alderson and Wall (1993), that is, that teaching content, by means of revised examination test books, produced the most pronounced washback effect. Furthermore, although results from teacher surveys indicated a positive attitude toward the new exam, the question whether the attitude would cause actual changes in teaching methodology remained. Cheng (1997) concluded that several issues would need to be considered in future investigations to determine the form and direction

of washback. Her recommendations include, first, the analysis of test properties and stakes, then the examination of the educational context, and finally an investigation of the degree of “superficiality” (p. 52) of the washback effect on teaching and learning.

Most recently, Qi (2004) investigated the intended washback effects of a high-stakes test, the National Matriculation English Test (NMET), in China. The test constructors anticipated that changes in the examination would bring about a positive washback effect through the shift from passive language knowledge to the active use and practice of English in classrooms. Although the test designers had included communicative elements into the exam, the multiple-choice format remained unchanged because of reliability concerns. The author found that classroom activities continued to be dominated by the de-contextualized learning of linguistic knowledge, the drilling of grammatical rules, and the teaching of isolated words. In particular, activities both in and outside of class were marked by the excessive use of multiple-choice items, despite test constructors’ explicit objection to the use of this format in teaching. The author offered the following explanation, based on teacher informants, to account for the apparently powerful impact of exam format: “students and parents will complain if the exercises at school do not look like the NMET” (p. 186). Qi concluded that, 15 years after the introduction of the exam, the intended positive washback was limited, manifested mainly at the format level, and that the NMET is an inefficient tool for inducing positive pedagogical changes (see Andrews, 1995). The author speculated that the intended washback may have been hindered by several factors addressed in other empirical studies, including teacher beliefs and background, teaching style, teaching experience (Wall & Alderson, 1993; Watanabe,

1996; Alderson & Hamp-Lyons, 1996; Cheng, 1998) as well as lack of communication between test constructors and test users (Wall & Alderson, 1993).

### **Washback on Teachers**

In order to explore the role of teaching methods in the washback process, Watanabe (1996) conducted classroom observations and teacher interviews in a Japanese exam preparatory school. The procedure employed in the study followed Alderson and Wall's (1993) recommendation that questions raised during observation be followed by interviews in order to gain insight into *why* teachers did what they did. The author reported similar findings to Alderson and Hamp-Lyons (1996), observing that the content of entrance examinations appeared to induce washback on some teachers but not others. Watanabe (1996) proposed that teachers' educational background, beliefs about teaching methods, as well as the timing of the observations may all be factors contributing to washback differences between teachers. In a more recent study, Watanabe (2004) revisited the role of the teacher in mediating washback and added that the degree of teachers' familiarity with a range of teaching methods might also be a factor influencing the process of producing washback: "It seems to be crucial then to identify empirically a range of effective teaching methods to improve authentic or real life language skills as well as to help students pass the examination" (p. 140). Furthermore, the author noted that psychological factors also play a role: "teachers may be placing undue blame on the presence of the examination for what they are doing" (p. 140). Importantly, he maintains that holding various unproven assumptions about the examination may hinder the generation of beneficial washback. The author proposed that engineering beneficial

washback would need to begin at the level of individual teachers, involving pre-service or in-service teacher training. An important component of this course would include “*re-attribution training*” to aid teachers in viewing the examinations not as an external, uncontrollable factors, but instead as internal, controllable factors. Watanabe (2004) concluded that helping teachers place the examination on a controllable dimension would add to their perceived validity of the test and may in turn affect their classroom practices.

Focusing on the role of the teachers in the washback process, Burrows (1999) conducted interviews and observations in the Australian Adult Migrant English Program (AMEP). The author was interested in learning whether the introduction of a new classroom-based assessment had changed teaching methods and content. Her findings were similar to those of Alderson and Hamp-Lyons (1996), Watanabe (1996), and Shohamy (1996), in that the teachers responded to the assessment implementation in different ways, “these responses differing in both kind and degree” (p. 7). Employing Woods’ (1996) framework of expressed beliefs, assumptions and knowledge (“BAK”), Burrows (1999) categorized teachers as falling into three main groups. The first group are “resisters”, who have educational and philosophical beliefs that conflict with the theoretical and educational foundations of the new assessment tool. The second group are “adopters”, whose educational and philosophical beliefs are in general agreement with the nature of the new assessment tool. Finally, the third group “adapters and partial adapters”, have ambivalent feelings toward the new exam because their beliefs, assumptions, and knowledge, although not in conflict with the bases of the assessment, are nevertheless challenged. Burrows (1999) concluded by emphasizing “the importance of teachers as a

medium through which any change brought about by the implementation of an assessment system or test must pass” (p. 20).

On the basis of international examples of high-stakes testing in educational systems, Chapham and Snyder (2000) examined the success of national testing in improving classroom instruction. The authors reported that one way that governments attempt to control the educational system is by changing national tests so that teachers are coerced to alter their teaching practices, which in turn assures their students’ success on those tests. They reported two examples of government-directed changes in the examination systems of Trinidad and Tobago and Uganda. Although these efforts produced some positive instructional changes, the authors emphasized that “it is not the examination itself that influences teachers’ behaviour, but teachers’ beliefs about those changes” (p. 462). Furthermore, in order for teachers to act upon changes in assessment practices, it is essential that they understand the nature of the change and how they can effectively prepare students for the new assessment. The authors noted, however, that changes in teachers’ behaviours do not necessarily translate into changes in student learning. Importantly, Chapham and Snyder (2000) point out that classroom instruction is not a static activity, but instead a combination of dynamic behaviours that do not have a linear, direct relationship with changes in test format or content.

Finally, in an investigation of washback in a public examination change in Hong Kong on teachers’ perceptions and actions, Cheng (1999,2004) used a combined research framework to learn how the new examination would influence classroom teaching, and how it would affect teachers’ attitudes and behaviours. The author found that, overall, teachers reacted positively to the change in examination and that the perceptions of the

policy makers were matched by those of the teachers. However, Cheng (2004) noted that although teachers displayed positive attitudes and a willingness to alter their classroom activities to accommodate the new examination, significant changes in their actual teaching methodology were not observed over the initial two years, suggesting no pronounced washback effect. That is, policy makers had successfully changed the *what* in teaching and learning, yet the examination had resulted in only minimal changes in the *how* of teaching and learning (p. 164). Interestingly, the importance of the public examination, although forceful enough to change classroom content, did not achieve significant changes in teachers' beliefs and attitudes about how to conduct their teaching and how students learn.

### **Washback on Learners**

As Wall (2000) observed, most findings in washback research to date relate to what Hughes (1994, cited in Wall, 2004) called the "processes" of teaching: "the selection of content (skills, teaching materials, exam preparation materials), the methodology teachers used and the ways in which they assessed their own students" (p. 502). In contrast, relatively little research has focused on the attitudes and behaviours of language learners as well as the "products" of learning. In particular, it has not been demonstrated that students learn more or better because they have prepared for a given test (Wall, 2000). Bachman and Palmer (1996) proposed that test-takers can be affected by: "the experience of taking, and in some cases preparing for the test; the feedback they receive about their performance on the test; and the decisions that may be made about them on the basis of

the test” (p. 31). The following studies are examples of research that focuses on the role of learners in mediating washback.

Bailey (1996) identified the direct influence of test-derived information on test-takers as “washback to the learner” (p. 264), and hypothesized that test-takers could engage in a variety of activities when preparing for a high-stakes test. These behaviours could include practicing items similar to the test, studying grammar and vocabulary rules, using test-taking strategies, attending test-preparation courses, and skipping language classes to prepare for the test (p. 265). The author suggested that using a selection of such behaviours would promote either beneficial or negative washback, depending on the actual language development promoted by these processes. As an example of washback on learners, Bailey (1996) referred to Buck’s (1988) study of listening tasks and associated washback on student preparation practices. She noted that in terms of washback, Buck considers inadequate the task-types that are commonly encountered in listening assessments including “noise tests”, listening cloze, tape-recorded sentence repetition, and dictation.

As part of their investigation of test impact over time, Shohamy, Donitsa-Schmidt, and Ferman (1996) administered questionnaires to students who were preparing to take one of two exams, which were considered either low-stakes (Arabic as a Second Language, ASL) or high-stakes (English as a Foreign Language, EFL). The authors found significant differences between the two student groups concerning aspects of test awareness, preparation time, and attitudes toward the effect of the test. They reported that of the students who would attend the ASL examination, 63% were unaware of its existence, 90% did not know what material would be included, and 77% said that their

parents did not know of the test (p. 304). With regard to the perceived consequences of the test result, 52 % of students believed that it would greatly affect their final grade and also held the view that their performance would influence future success in their studies. However, overall students did not believe that the examination raised the status of ASL and, indeed, 65% of students felt that success in the test was not important. In contrast, the researchers observed very different impact results of the EFL examination on learners. While students received no information from their teachers about the ASL exam, they were informed in detail of the changes to the EFL exam at the very beginning of the term, which is confirmed by the questionnaire responses showing that 97 % of students and 66% of parents were aware of the changes to the exam. This awareness was mirrored in the high anxiety that it produced in 96% of students, 70% of whom believed that EFL test results could influence their success in future studies. Consequently, 82% of students regarded the exam as very important (p. 311). These findings appear to indicate that factors including exam awareness, perceived importance of results, as well as the status of the subject mediate the test impact on learners.

Cheng (1998) conducted a longitudinal study on the effect of a public examination change in Hong Kong secondary schools on students' views and attitudes toward their language learning activities. The design of the study, which included questionnaires administered to students under both the new and old examination, allowed the author to explore the relation between the introduction of a new test and the process of language learning. Cheng reported findings similar to her study of washback on teachers perceptions and attitudes (1999, 2004), in that the new examination "informed students' *what*, but not *how*" (p.297). In particular, students' reported motivation to learn English,

learning approaches, and views on the examination, did not exhibit significant changes over the initial two-year period following the introduction of the new exam. The author thus concluded that more profound changes to teacher methodologies and classroom materials are necessary to produce an associated change in learning processes.

Most recently, Watanabe (2002) addressed washback on the learner in relation to different models of motivation. Similar to Burrows' (1999) investigation of test washback on teachers, he proposed that factors internal to the test taker, such as belief, assumption, and knowledge ("BAK", Woods, 1996), determine the intensity of washback. Watanabe then introduced an attribution theory model of motivation as an additional mediating factor in shaping washback on the learner. This attribution theory assumes that people refer to four sets of attributions in explaining their successes and failures: ability, effort, luck, and the perceived difficulty of the task (Williams and Burden, as cited in Watanabe, 2002). Of these factors, ability and effort are regarded as internal, while luck and difficulty are seen as external factors. This model of attribution can be applied to students' perceptions of their test performance. For example, students who believe that their failure was caused by lack of ability (uncontrollable) can be helped to attribute the cause of their poor performance to lack of effort (controllable). Watanabe suggested that *retribution* training of this kind would facilitate beneficial washback. The author presented an additional "flow" model of motivation, asserting that if students are to be motivated by means of testing, the difficulty level of the assessment must be appropriately challenging. If a test is too difficult, students may lose confidence and not prepare at all; conversely, if a test is too easy, students may not bother to engage in test practice either. The final model of motivation the author discussed, involved the "functional autonomy of motives", in

which learners are motivated by either an instrumental motive or an integrative motive for performing well on a test. Watanabe (2002) suggested that an initial instrumental motive - to pass the test- may not necessarily lead to an integrative motive - actual interest in the target language. Finally, the author concluded that washback is a subjective phenomenon, perceived differently depending on the individual and the context involved.

Bailey (1999) observes that, although language learners are key participants in the testing process, there is relatively little research that documents their point of view and their washback-related behaviour before and after tests. Given that language learners are those most directly affected by potential washback, Watanabe's (2002) approach to washback on learners signals an important change in the direction of research by affirming that test-takers must form an integral part in the investigation of the nature of washback.

### **New Models for Washback**

Researchers have identified a wide variety of factors that influence washback. Brown (1997) has organized factors that have been suggested to influence the process of washback into four categories including: *prestige factors*, such as the status of the test, status of the subject of the test, perception of quality and importance of test, or levels of stakes (Shohamy et al., 1996); *test factors*, such as test methods, test content, purpose of the test, format of the test, skills being tested, practicality, or use of test results (Gates, 1995; Shohamy et al. 1996); *people factors*, such as teacher ability and background, beliefs about teaching and learning methods, or student anxiety (Alderson and Wall, 1993); and *curriculum factors*, such as the match of current teaching practices, types of

classroom activities, or generation of new teaching materials (Alderson and Hamp-Lyons, 1996). Watanabe (2004) adds to this list *micro-context factors* (e.g., school environment in which test preparation occurs) and *macro-context factors*, which refer to the larger society in which the test is used.

Wall (2000) observed that the challenge now facing researchers is how to categorize findings and to develop a model that considers all variables while remaining “coherent and user-friendly” (p. 503). For that purpose Wall (2000) proposed Heinrichsen’s (1989) “Hybrid Model of the Diffusion/ Implementation Process”, which divides the process of innovation into Antecedents (i.e., conditions in the educational system before the introduction of a test), Process (i.e., operating factors during test implementation), and Consequences (i.e., types of outcomes created by the interaction of Antecedents and Process). Although this model cannot predict the form that test impact will take in a given context, it allows test developers and implementers to make judgments about the benefits and risks involved in a new testing project.

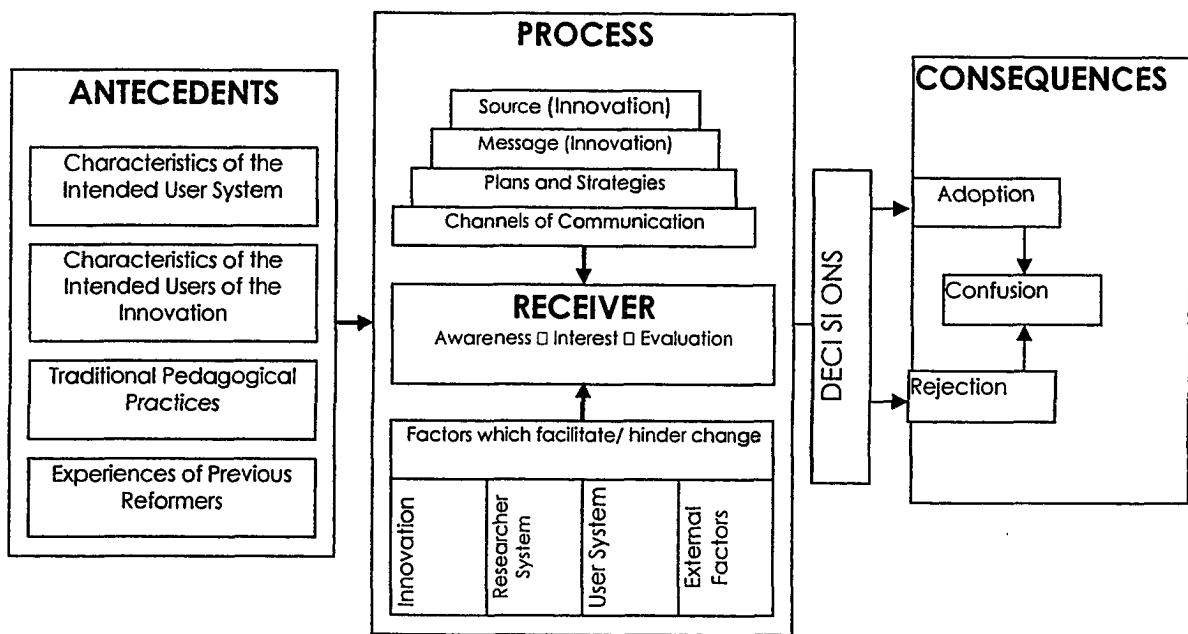


Figure 4. Henrichsen's Hybrid Model of the Diffusion/Implementation Process as presented in Wall (2000).

Watanabe (2004) contributed a five-dimensional model for the categorization of washback findings including *specificity*, *intensity*, *length*, *intentionality*, and *value*. Washback *specificity* may range from general, which can be produced by any test, to specific, which is contingent on a particular feature of the test or test type. The *intensity* of washback may be strong, determining all actions and behaviours in the classroom, or weak, affecting only some activities, teachers, and students. Watanabe noted that research findings tend toward the weak end of the continuum and that the stakes of the test are a mediating factor in washback intensity. The *length* of examination influence may be of short or long duration. *Intentionality* separates washback into intentional and unintentional examination effects. This dimension has been linked to validity judgments in terms of social consequences of test interpretation and use (Messick, 1989). Watanabe noted that

McNamara (1996) also supports this view, stressing the importance of collecting evidence on “the intended and unintended effects of assessments on the ways teachers and students spend their time and think about the goals of education” (p. 22). Finally, the *value* of washback can be either positive, usually linked to intended effects, or negative, which may be linked with both intended and unintended washback. The author further noted that it is important to identify the audience for evaluation studies of washback, because judgments of positive or negative washback may be dependent on the perspectives of the school system, teachers, and learners.

### **Methodological Issues in Washback Research**

As the review of literature revealed, there are numerous factors that mediate the process of washback (Alderson and Wall 1993; Alderson and Hamp-Lyons, 1996; Shohamy et al. 1996; Brown, 1997), which in turn operate on several dimensions (Watanabe, 2004). Consequently, in attempting to disentangle the complexity of the phenomenon, a variety of methodologies may be employed for investigations aiming to develop further specifications of washback. Although testing has traditionally been considered a scientific field, influenced and dominated by quantitative, psychometric/measurement approaches to research (Shohamy, 2001), qualitative and ethnographic methodologies are increasingly being used. In his discussion of methodology in washback studies, Watanabe (2004) maintains that a traditional experimental approach would require the creation of artificial test conditions which, unlike natural situations, would lead participants to view the test as having few educational consequences and thus,

not producing useful and meaningful findings. The shift toward qualitative approaches signals the need to gain insight into the *why* of behaviours and beliefs, in addition to the observable “facts” of washback processes. In addition, qualitative methods such as interviews, surveys, and questionnaires, complement observations, aid in the interpretation of data, and importantly direct researchers towards pertinent new questions to be addressed (Alderson and Wall, 1993).

## **Summary**

The influence of testing on teaching and learning is widely assumed and regarded as significant. Before the 1990s few empirical studies had been conducted to document the exact nature of washback or the mechanisms by which it operates. More recently, researchers have begun to investigate the form of washback in a variety of contexts. These studies have led to a more complex understanding of washback and brought into consideration factors outside the immediate testing situation. However, there is no consensus on what factors influence the process and form of washback, nor do researchers consistently produce similar findings. In order to categorize findings and facilitate research several frameworks have been introduced, providing researchers with conceptual tools to examine the washback/impact phenomenon from different perspectives and facilitate further specification of the concept. Importantly, the classification of the multiple factors involved in washback plays an important role in establishing frameworks for studies investigating the effect of washback on teaching and learning behaviours.

The foregoing significant studies have focused almost exclusively on the impact of tests on educational systems (at the macro level) and washback on teachers and their classroom behaviours (at the micro level). Although several researchers have begun to address how test washback affects learners (Cheng, 1998; Shohamy, Donitsa-Schmidt, & Ferman, 1996; Watanabe 2002), we still know very little about learners' perceptions toward tests and test-taking (see Wall, 2000). Importantly, the studies discussed have generally focused on how a *test* affects the educational system and its participants, whereas little attention has been granted to the role of *learners* and their respective *contexts* in mediating washback. Given that learners represent the central reason for testing and teaching, it is essential that their role in test washback be further investigated.

## **Chapter 3: Research Methodology**

The purpose of the current study was to investigate washback on learners that may be the result of important *contextual* factors, such as learner attitudes toward language learning, affective reactions, past test experience, and other variables that may be present within a high-stakes testing situation. In order to investigate these various *contextual* factors, the following research questions framed this study: What washback factors may be identified through an exploration of learners' perceptions toward a high-stakes test? Which of these factors appear to reflect variables in the context surrounding the test (i.e., *contextual* factors), and which appear to reflect those specific to the test (i.e., *test-specific* factors)? On the basis of these findings, is it possible to delineate a model that broadens current understandings of washback processes and informs future research?

### **The Methodology of the Current Study: Overview**

In order to reach a better understanding of washback in language testing, there is a need to further explore the phenomenon within a qualitative framework. For that purpose, the current study employed a *grounded theory* approach (Strauss and Corbin, 1998) in order to explore contextual washback factors through learners' perceptions toward a high-stakes test. One of the main tenets of the *grounded theory* approach is that theory should be allowed to emerge from data, rather than tested, in order to more accurately reflect reality. For this purpose the researcher aims to build theory through a variety of "coding procedures" (p. 13) that allow for creative and responsive analysis of data. In developing

theory, the researcher therefore considers a range of meanings and avoids taking one stand toward the data through “microanalysis” (p. 57). During this process, data is examined specifically by focusing on words, phrases, and sentences while continuously asking general and theoretical questions to “flush out concepts and their relationships”(p. 66), as well as develop concepts and classifications of events. In addition to asking questions to build theory, “theoretical comparisons” (p. 73) are made to systematically move from the level of description to abstraction. Through the application of these procedures, the researcher is thus able to group specific data items into categories and concepts, define their properties, and discover patterns that represent the foundations for building theory.

As this was a preliminary investigation, the main focus was to develop constructs of *contextual* washback factors as expressed through learners’ perceptions toward a high-stakes test. The current findings may then serve in the development of a larger research framework which would include a variety of participants (learners, teachers, administrators) and methods (observation, interviews, and questionnaires).

The detailed description of the research context is essential in understanding the role of both the participants and the test in this study. Such “thick description” (Geertz, 1973) is further necessary to establish transferability, or generalizability, of findings of the present study to other contexts (Brown, 2002).

### **Context**

In order to situate the present study in the educational context in which it was conducted, the following section will provide an overview of: the researcher role, the English for Academic Purposes (EAP) program, the student population, teachers, the nature of the test, and the use of test results.

### **Role of the researcher**

Watanabe (2004) notes that because almost all researchers have themselves been in the position of a test-taker, it is very probable that their own experience may influence or even sway their research. He thus suggests that “to increase the degree of reliability or ‘trustworthiness’ (Eisenhart & Howe, 1992) of the research, it is important to make one’s base line explicit” (p. 23). Watanabe refers to Allwright and Bailey’s (1991) definition of “base line” to describe the framework of experiences, attitudes, and education that may affect researchers’ observations. In order to make my base line explicit, I feel it is necessary to briefly outline my role and involvement in the current research context:

I began teaching in the English for Academic Purposes (EAP) program at a large Canadian university in 2001 and have acted as program coordinator since the fall of 2002. In my role as coordinator, I am responsible for:

- a) English language placement and proficiency testing for EAP program students, as well as international undergraduate and graduate students.
- b) Developing and maintaining language support programs, including the Conversation Partner Program and the Homestay Program, as well as organizing social activities.
- c) Providing teacher support for curriculum development and resources.
- d) Providing student support for cultural, social, and academic adjustment.

My responsibilities as coordinator therefore result in a tight connection with the EAP program administration, teachers, and to a lesser extent, the students.

At the time I joined the EAP program as coordinator, the EAP curriculum was not reflected in the content or the construct of the *exit test* (i.e., the language proficiency test administered at the end of the program). However, given that performance on this test

either granted or denied students access to the university, they often demanded that teachers devote class time to test preparation. Consequently, teachers felt dissatisfied because they believed that practicing for the test, rather than EAP, would not adequately prepare students for university studies. During that time I had been working on a critical review of language testing practices, and proposed that the program adopt a test that is closely linked to actual target language use (TLU) (Bachman and Palmer, 1996) in order to: (1) create a better “fit” between the test and curriculum/program goals; (2) draw more valid inferences about language performance from the test; and (3) allow students and teachers to engage in test practice that reflects current language learning theories. In response to this suggestion, the Canadian Academic English Language (CAEL) Assessment (Fox, 2000; 2004) was introduced in the winter term of 2003.

### **The EAP program**

The study considered in this thesis was conducted in the English for Academic Purposes (EAP) program at a large Canadian university during the winter term of 2004. The EAP program is offered three times a year and is designed to prepare students for the demands of an English-medium university setting. The course lasts twelve weeks and is divided into morning and afternoon classes. The morning classes focus on integrated English skills, including essay writing, academic reading, and listening to lectures; the afternoon classes focus on academic speaking and casual listening, including discussion, debate, and formal presentations. The Pre-CAEL ADAPT (see “The test”) is used as a student placement tool at the beginning of the term. Throughout the course, student performance is evaluated according to the grading regulations of the university on the

basis of quality of completed assignments, in-class assessments, and participation. At the end of the course, students complete the CAEL Assessment, which is a standardized English language proficiency test accepted by the university (see “The test”).

### **Students in the EAP program**

The information provided in this section refers to the student population at the time of the current study in the winter term of 2004.

a) *Student Enrollment:*

A total of 103 students were enrolled in the EAP Program and were divided into seven classes, which were labeled from A (lowest) to G (highest). The number of students per class ranged from 12 to 17.

a) *Country of Origin:*

Of the 103 students in the EAP Program, eighty-eight were from the People’s Republic of China (85.4%), ten were from Korea (9.7%), four were from the United Arab Emirates (3.8%), and one was from Turkey (0.9%).

b) *Age:*

The age range of students was between 18 and 24 years.

c) *Gender:*

There were 43 female students (41.7%) and 60 male students (58.2%).

d) *Years of English Study:*

Students from the People’s Republic of China had completed on average 9 to 11 years of English instruction, beginning studies in junior middle school and continuing through to college or university studies. The level of instruction in the EAP program requires that

students have “intermediate” (CAEL band level 20-40) proficiency in English. This information may be provided through available language proficiency test scores, or if unavailable, self-assessment reports may be used.

e) *Purpose of English Study:*

Most students enrolled in the EAP program have provisional acceptance to the university and are attending the course to improve their English language skills to a level that satisfies the language proficiency requirement. At the time that the current research was conducted, a majority of the EAP students were enrolled in a joint Chinese/Canadian university Bachelor of Technology program. They had completed three years of their degree at a Chinese university and needed to complete their last year at a Canadian university. Before their arrival at the Canadian university, students had completed English language tests in China, but had failed to satisfy the language proficiency requirement. Students were thus informed that they would need to improve their English language skills before continuing with their studies at the Canadian university.

**Teachers in the EAP program**

Teachers in the EAP program have a degree (Masters or one-year diploma) in English as a Second Language Teaching, Applied Linguistics, or other closely related field. The level of teaching experience ranges from novice to highly experienced credit level instructors and teacher educators. Teachers use both EAP textbooks and self-developed materials in their classrooms and are encouraged to include authentic texts and listening materials.

With the introduction of the CAEL Assessment in the winter term 2003, teachers attended CAEL “orientation” meetings during which they familiarized themselves with the test philosophy, format, and rating system. All teachers were also interested in receiving CAEL rater training in order to serve as CAEL essay raters for the final assessment at the end of the term. Having gained an understanding of the test, teachers included practice tests in classroom activities, demonstrated to students how their curriculum is reflected in test tasks as well as the university setting, and some developed their own CAEL practice materials. This integration of the test into classroom activities was based on the assumption that “in the CAEL Assessment context, teaching to the test does not undermine student learning, rather it supports and complements it” (Fox 2000, p.12). All teachers administered the same CAEL practice test at the midpoint of the course, but did not otherwise receive specific instructions on how to use CAEL practice materials in the classroom.

### **The Test**

The EAP program uses two externally developed tests: the Pre-CAEL ADAPT for placement purposes at the beginning of the program, and the CAEL Assessment for final language proficiency testing.

The stated purpose of the CAEL Assessment is to test and describe students’ ability to use English for academic study at English-medium colleges and universities. The CAEL was developed in 1987 at Carleton University as an alternative to other standardized tests, which failed to effectively identify students able to engage in academic study (Fox, 2000; 2004).

The CAEL is a topic-based performance test, comprised of an integrated set of language activities including reading academic articles, listening to a lecture while taking notes, and responding to the learned information in the form of an essay. The test includes language tasks, activities, and content areas that are representative of those commonly encountered in introductory university courses (see Appendix C for sample test version).

The CAEL is a criterion-referenced test and results are reported in descriptive band scores from 10-90 for each of the subtests of reading, listening, and writing. The official version of the CAEL includes a speaking sub-test, but for practical reasons this component is not administered within the EAP program.

In addition to the CAEL, the Pre-CAEL ADAPT is used as a placement tool at the beginning of the term with the aim of creating a cohesive testing process. This assessment is an academic diagnostic and placement test which is based on the CAEL model, but is specifically designed for students at lower-intermediate levels of English proficiency. The results of the Pre-CAEL ADAPT are used to place students in appropriate class levels and also provide diagnostic information on academic English reading, writing, and listening skills.

### **The use of test results**

As discussed above, most students who enroll in the EAP program have provisional acceptance to the university and, in order to gain full acceptance, must provide proof of English proficiency in the form of a standardized test score. The completion and level of performance in the EAP program is not considered in the university admissions process. However, if a student's CAEL Assessment result is discordant with teacher evaluations

and in-class performance, available information may be used to provide the student with an opportunity to complete another test (i.e., CAN-Test).

Students who achieve an overall CAEL band score of 40 are allowed to register for one university credit course in math, chemistry, economics, or physics, under the condition that they attend a non-credit EAP class during the morning session. This arrangement is based on the recognition that once students have attained a threshold level of language proficiency, it is advantageous to both students and the university that they commence their academic programs, although their language skills are below the level required for full-time study (Fox, 2004).

Students who achieve an overall CAEL band score of 50 are allowed to register for credit courses and must attend a non-credit EAP writing course (English 102f). Students at this level of proficiency are strongly encouraged to register for no more than three credit courses; however, due to the nature of the university's registration regulations, students are not prevented from registering for a full course load.

Students who achieve an overall CAEL band score of 60 are allowed to register for credit courses and must attend a credit EAP writing course (English 1020).

Students who achieve an overall CAEL band score of 70 or more are allowed to register for credit courses and must complete two English credits as part of general degree graduation requirements.

Students who achieve an overall CAEL band score between 10 and 30 are not allowed to register for credit courses and are advised that they will require additional EAP studies before achieving a language proficiency level necessary for credit courses.

Students who did not “pass” (i.e., attain band level 50 or above) the CAEL assessment, or who are dissatisfied with their result can provide an alternative standardized language test score, such as the CAN-Test, TOEFL, IELTS, or MELAB, and may be admitted if the result meets the university language proficiency requirements.

## **Participants**

The participants<sup>1</sup> in this study were student volunteers recruited from all levels of EAP classes who were interested in sharing their experiences and attitudes about language tests. After hearing and reading a brief description of the study, forty-seven students provided their contact information to receive more information. After receiving a more detailed description of the study over e-mail, ten student volunteers from different class levels agreed to be interviewed, six of whom were available at the time the interviews were conducted.

The participants for this study, then, were six adult learners between the ages of 20 and 25, of whom four were female and two male. All were from the People’s Republic of China. At the time of the study, all of the participants were attending the EAP program for the first time and were at four different class levels: intermediate-CAEL band 30 (learner A), high intermediate- CAEL band 30+ (learner B), low advanced-CAEL band 40 (learners C and D), and advanced CAEL band 40+(learners E and F). The participants had

---

<sup>1</sup> The current research was carried out in accordance with the Tri-Council Policy Statement, “Ethical Conduct for Research Involving Humans”. An application dossier, including all the ethical considerations of dealing with human subjects in research, was submitted to the Ethics Committee of the university and Ethics Clearance was approved.

relatively homogenous backgrounds as they shared the same mother tongue, culture, and current learning environment.

Furthermore:

- All six participants had completed nine to eleven years of English language studies in China.
- All six participants intended to enter a Canadian university after fulfilling the English language proficiency requirement.
- At the time of the study, five of the participants had been living in Canada for four months, while the sixth participant had resided here for one and a half years.
- Five participants were enrolled in the joint Chinese/Canadian university Bachelor of Technology Program, which required them to fulfill the final year of their degree at the Canadian university.

### **Instruments and Procedures**

The data collection mode in the study was comprised of two one-on-one interviews that were conducted with each of the six participants during the course of one twelve-week term; one occurred at week seven and one occurred after the completion of the CAEL at the end of the term at week twelve. Each interview lasted approximately 45-60 minutes, was carried out in English, tape-recorded, and subsequently transcribed.

Structured questions provided a framework for each interview, but additional questions were allowed to evolve during the interaction. This has been the preferred type of interview for many qualitative researchers, as it is flexible and produces rich data

(Nunan, 1992; Woods, 1996). These structured questions were informed by Henrichsen's Hybrid Model of the Diffusion/Implementation Process (see Figure 4, Chapter 2), a framework proposed by Wall (2000) for the categorization of washback factors identified in studies to date. This model represents the process of innovation in three components including *antecedents*, *process*, and *consequences*. These three components were used to initially structure questions for the interviews, including background questions, test preparation and practice, test-taking experience, as well as test results and consequences.

The interviews included three types of questions. First, "background questions" were asked during the initial phase of the interview in order to capture some grounding ideas of participants' language learning experience, motivation for learning English, and past experiences with testing. Second, "attitude and perception questions" were then explored in order to gain insight into participants' experiences in the EAP program in connection to the test, as well as factors external to the EAP program influencing their beliefs and attitudes about tests. Third, "follow-up" questions were added as the data collection and analysis evolved. These questions were aimed at identifying the relationship between the findings from early interview stages and changes in perceptions and attitudes which occurred after the completion of the test. Accordingly, the data collection phase and analysis of findings coincided, such that the results of the initial interviews framed questions for the second set of interviews.

## **Analysis**

Of central concern during the early stages of data analysis was to avoid premature reductionism of data and to expose assumptions that may be taken for granted by the researcher. Therefore, the initial analysis phase involved a process of “line-by-line analysis” (Strauss and Corbin, 1998) in order to conceptualize and categorize statements, events, and outcomes from the data collected. This procedure allowed the researcher to examine participants’ responses line by line, while labeling each statement or event with relevant concepts such as: attitude toward English, test preparation, anxiety, peer talk, fairness, etc. Throughout the course of analysis, recurring themes and concepts emerged through further analytic and interpretive examination of data. These themes and concepts were used to continuously shape additional questions and provided theoretical foundations.

## **Chapter 4: Results and Discussion**

This chapter presents the results of the analysis of washback factors that were identified through an exploration of learners' perceptions of a high-stakes test. This exploration served to examine the role of *contextual* washback factors, such as learner attitudes toward language learning, affective reactions, past test experiences, and other variables that may be present in a high-stakes testing situation. The analysis of data generated by the six participants revealed two distinct categories of washback factors that were identified through learners' perceptions toward the CAEL Assessment: factors that are *test-specific* and extra-test factors that I shall term *contextual* factors. Although these two categories are interrelated, for the purpose of this discussion they will first be addressed separately and their connections explored subsequently. Each of the two categories of washback factors include subcategories of themes that shall be explored in the following discussion.

Although a great variety of factors were identified by single participants, the purpose of the current study was to identify common themes among participants, under the assumption that these commonalities would likely represent the most important factors for researchers to explore. Consequently, all of the themes that are discussed were addressed by at least three of the six participants. In terms of data, these themes are illustrated by a selection of participants' quotes. It is important to indicate, that these quotes by themselves are not meant to be proof of positive or negative washback. Instead, what these quotes reflect are what participants' themselves found significant in a high-stakes testing situation, and what we may therefore want to take into consideration in formulating

hypotheses about what factors are important in mediating washback on learners.

Finally, the two categories of washback factors are divided into pre-test and post-test responses in order to reflect the interview structure. This division is also relevant because of the possibility that participants' test-taking experience and test performance may have further shaped or elaborated their perceptions.

## **Test-Specific Washback Factors**

The washback factors discussed in this section all relate to the CAEL in particular. These factors were categorized as *test-specific* because they appeared to be closely connected to the construct and format of the assessment tool itself. The five themes identified were test knowledge, format, difficulty, test-specific attitudes, and test-taking experience.

### **Pre-test responses**

The test-specific washback factors that are discussed in the following section emerged from the analysis of the first set of interviews with participants, which were conducted at week seven of the twelve-week EAP program.

### **Test knowledge**

The CAEL Assessment was developed in 1987 at Carleton University and has since been introduced in many parts of the world to be used as part of the admissions process mainly to universities and colleges in Canada, but also in Europe and the United States.

However, administration of the test in the People's Republic of China has begun only in recent years and all of the participants in this study noted that they had never heard of the CAEL before joining their EAP program in Canada. However, all of them knew about the TOEFL and five of the participants had at one time either been preparing to take the test or had already taken it. Indeed, the following comments illustrate participants' general belief that preparing for the TOEFL is a natural and obvious path to follow for English learners, especially because it serves as a tool for access to universities abroad.

E-(adv.)<sup>2</sup> : "In China, all students know about the TOEFL and prepare for it."

B-(high-int.): "The TOEFL is just a proof for international students. It just gives them a chance to study abroad."

Participants, who had been preparing for the TOEFL, indicated that having learned about the assessment and developed familiarity with it, led them to perceive it as less difficult as well as more "comfortable" than the CAEL.

E-(adv.): "The CAEL is new so the TOEFL is easier to them. They are more familiar with the TOEFL than with the CAEL. Students say that they would rather do the TOEFL; the CAEL is more difficult and they have heard and learned more about the TOEFL."

---

<sup>2</sup> Quotes are identified by participant code (A-F) and class level intermediate (int.); high-intermediate (high-int.); low-advanced (low-adv.); and advanced (adv.).

The importance of test familiarity was also addressed by a participant who had not prepared for the TOEFL and felt that, even though it was possibly “easier”, she would choose the CAEL because she was more familiar with it.

A-(int.): “I am not familiar with the TOEFL because I didn’t do that practice. Maybe it is easier than the CAEL, but now that I know how to do the CAEL I would choose it.”

The comments offered by participants E and A illustrate that test knowledge and familiarity were closely connected to their preferences for a given test. Also, the length of time that was required by participants to develop familiarity with a new test differed depending on previous test experience. For example, although the CAEL was novel to both participant E and A when they first joined the EAP program, participant E did not yet feel comfortable with the CAEL, because his years of experience with the TOEFL far outweighed his newly-acquired knowledge about the CAEL; on the other hand, participant A had never practiced for the TOEFL and thus indicated that she felt comfortable with the CAEL after having learned about it during her time in the EAP program.

At the time the interviews were being conducted, all of the participants had been attending the EAP program for seven weeks and thus had had some time to gain familiarity with the CAEL through their teachers, class activities, and peers. As described in Chapter 3, although all of the teachers attended the same “CAEL orientation” sessions, they did not receive specific instructions on using test practice materials in the classroom. In addition, while some of the participants’ classes included students who had previously

attended the CAEL, others were in classes where none of the students had taken the assessment. Consequently, the type and amount of information that participants received about the CAEL from their teachers and peers may have varied. Nevertheless, all of them offered similar descriptions of the assessment.

E-(adv.): “The CAEL is very near to studying university lectures and courses.”

B-(high-int.): “The CAEL requires not just one part of the English language, it requires the whole parts: listening, reading, and writing. In addition, I think that the CAEL is just like a course. The whole test is just one topic and you listen and you can know lots of information and then you can write about it. When you are finished this exercise you will know the full meaning about this topic.”

F-(adv.): “I have not entered the university classes to know what I need for university studies, but some of my friends told me that the CAEL test is like the class in university.”

The aspect that figured most prominently in participants’ descriptions of the CAEL was that it is a topic-based test. While some participants viewed this feature as a reflection of what university courses are like, others viewed it as a source of confusion.

D-(low-adv.): “I don’t know what will be the area of focus in the CAEL test. The first practice test is sleep, the second time it’s academic listening, and then brain chemistry; so I just don’t know the focus. Although we don’t need the background of this topic, sometimes if we are familiar I think I can easily understand and get a high score and the student will be more comfortable to take the test.”

Although participants realized that they did not *need* to know anything about the topic of the test in order to perform well, they nevertheless felt much more comfortable about familiar topics and perceived them to be less difficult.

E-(adv.): “During the test the topic was not so difficult. Maybe in my own experience I watched a lot, read a lot about these kinds of things.”

The role of topic in test performance was examined by Jennings, Fox, Graves, and Shohamy (1999) because of a concern that a given topic may positively or negatively affect some test-takers’ results. They suggested that factors such as “the test taker’s interest in the topic, prior knowledge of the topic, and the perceived relevance of the topic may have an effect on the test-taker’s performance” (p. 427) and thus represent an important consideration with regard to construct-irrelevant variance of topic-based tests. The researchers also acknowledged that test-takers could possibly understand the readings and lecture better with prior knowledge of the topic and thus reflected participants’ perceptions that a familiar topic was more comfortable and less difficult.

### **Test format**

In describing their knowledge of the CAEL, many participants made reference to the TOEFL in order to make comparisons and explain differences they perceived in the test formats. The three main components of the TOEFL include listening comprehension, structure and written expression, and reading comprehension, all of which are in a multiple-choice format. The fourth component of the TOEFL is the Test of Written English (TWE), which requires test-takers to compose an essay responding to a topic statement. Some participants noted that the essay format of the TOEFL allowed them to memorize sentences and structures that they could use for any given topic. This format was similar to the writing that they had practiced in their English classes in China, where the memorization of model essays was practiced to allow them to “write more smoothly” during a test.

B-(high-int.): “In the TOEFL test we have some model tests, we have some model essays that we can read to remember some good sentences. You can just remember this model and you will get a higher score. Maybe this is good for us but for the CAEL test we cannot do it.”

Another comparison of the two language tests made by participants C and D referred to the different formats of the listening section.

C-(low adv.): “The TOEFL will be like two people, one person says: “is the weather rainy?” and the other person says: “oh no, it’s nice”, and then it’s finished

and the next question will be changed. This is not like in the CAEL where you have one topic from listening to writing.”

D-(low-adv.): “In the CAEL there is a long talk. When we take notes we will need to answer questions and if we cannot focus we will miss lots. In the TOEFL test there are short talks and you just answer one question and go to the next. In that situation the TOEFL will be more comfortable.”

In their descriptions of their knowledge and understanding of the CAEL, all participants made comparisons to the TOEFL, but also to the English Band level tests that they were required to attend in China and which they regarded to be similar to the TOEFL. The differences that participants observed between the two tests, in particular with regard to test format, served to illustrate which parts of the CAEL they perceived as new and different from their previous testing experiences. Within their comparisons, it is notable that their familiarity with a particular testing format played a central role in their preference for and comfort with a test. In this regard, Shohamy (1982) suggested that practice and familiarity with a given testing procedure may affect attitude and anxiety. This relationship between test familiarity and affective factors will be further addressed in relation to participants' practice and test-taking experiences in the discussion of *contextual washback factors*.

### **Test difficulty**

In their comparisons, participants made reference not only to test format differences

that they identified between the CAEL and the TOEFL, but also to test difficulty levels. As the following array of quotes illustrates, there was no consensus among participants on which of the two tests they perceived to have a higher degree of difficulty.

B-(high-int.): “The CAEL is more difficult than the TOEFL.”

C-(low-adv.): “I think the CAEL is similar to the TOEFL.”

F-(adv.): “I think the TOEFL is harder than the CAEL.”

Participant B, who perceived the CAEL to be more difficult than the TOEFL, explained that certain strategies that she used for the TOEFL, such as guessing and memorizing essay models, which allowed her to complete even questions whose answers she did not know, could not be used for the CAEL. However, her estimation of the CAEL as a more difficult test did not automatically translate into a preference for the TOEFL because she felt it was “easier”.

B-(high-int.): “In the TOEFL you just guess something, maybe it’s easy to get a high mark but I don’t think that it’s useful.”

Participant C stated that, although the CAEL and the TOEFL had different formats, she felt that both asked test-takers to complete similar tasks (such as understanding an article, listening to speakers, answering questions, etc.) and had similar difficulty levels. Finally, participant F considered the TOEFL to be more difficult than the CAEL in part because she felt that a test-taker was required to use “common sense” in the essay

component of the TOEFL, whereas a test-taker could use information from the reading and listening sections of the CAEL to answer the essay question. Furthermore, she felt that the reading comprehension section of the TOEFL required a test-taker to understand every sentence, whereas using skimming and scanning strategies was sufficient to answer reading questions for the CAEL. However, although participant F felt that the TOEFL was more difficult, she did not necessarily mean that the test measured “higher English”.

F-(adv.): “If you pass the TOEFL it just means that you are familiar with the TOEFL, but if you pass the CAEL it means you have the ability and you can go to university classes.”

Importantly, although participants B and F stated opposite views on the test difficulty levels of the TOEFL and the CAEL, both indicated that passing the TOEFL was less useful or not meaningful. Thus, for these participants, the meaningfulness of the test result was more central in shaping their attitude toward the test than their estimation of the difficulty level.

#### Test-specific attitudes

The language assessment experience of all of the participants had been defined by multiple-choice tests and “topic-response” essays. Therefore, the topic-based, integrated skills format of the CAEL represented a great contrast to their familiar test framework. The participants generally regarded this difference favourably and expressed a positive attitude toward the more “authentic” format of the CAEL. This attitude was particularly

expressed in participants' descriptions of the CAEL tasks as "true learning" and "learning useful things", as well as the connection they observed between the tasks and the university setting that they were preparing to enter.

A-(int.): "The CAEL is a very, very good test. It is better than other tests such as the TOEFL. I can learn some true and useful things, when I enter university I can learn easily. I think the CAEL is important, because if we enter university directly we don't know how to listen to lectures and will fail credit courses."

C-(low-adv.): "If you get a high mark in the TOEFL then it's just one test. It means that you can do this test well; it does not mean that you can learn English with your professor very well. We don't learn some true things from choosing multiple-choice. In the CAEL I can learn a lot of things. You listen to some information from the lectures and take notes; you can learn something."

These statements reflect a theme that was also observed by Shohamy (1982), who found that students preferred testing procedures which are similar to real life communicative situations in both form and content, and which they perceived as a learning tool. Importantly, these participants felt that learning "true and useful things" was a necessary part of their preparation for successful university studies. Both participants A and C commented that mastering the skills required for the CAEL, including taking notes, reading quickly, and integrating information for an essay response, would facilitate their learning in university. This belief was also represented in a study of test fairness by

Sambell, McDowell, and Brown (1997), who found that students perceived test tasks as relevant or authentic if importance was placed on what they judged as valuable and transferable skills. Participant B extended the relevance of the CAEL in preparing for university studies by suggesting that succeeding in the test would also increase learners' confidence in their English abilities.

B-(high-int.): "I think it is a 'fire test' because the CAEL is just like a course.

When I first heard about it I thought that it was difficult and different, but when I studied for these months I think it is really useful. If we have this test it can prove our English is the same as local students. If we pass the CAEL it can give us confidence."

These comments demonstrate that participants' positive attitude toward the CAEL was shaped by their perception that it promoted "true and useful" learning. Furthermore, as participant B noted, even though she perceived the CAEL as different and difficult, she also judged it as useful because it tested her readiness for university.

### **Post-test responses**

The test-specific washback factors that are discussed in the following section emerged from the analysis of the second set of interviews with participants, which were conducted after the completion of the CAEL at week twelve of the EAP program.

### **Test-taking experience**

When relating their test-taking experience with the CAEL, all six participants addressed two issues that were most salient to them: the differences between the practice and final CAEL versions and the mismatch of their reading strategies with the reading comprehension questions presented in the test.

The recently published CAEL Assessment Preparation Guide includes several practice CAEL tests and two of these practice versions are also available on the official CAEL website. All of the participants in this study had completed at least two practice tests taken from the preparation guide during their EAP classes; one participant had repeatedly practiced with the test versions available on the CAEL website. None of the participants had purchased the preparation guide. A common assumption reported by all participants was that the practice tests not only reflected the components and question format of the “real” CAEL, but were also representative of the difficulty level. However, when they attended the final CAEL, they perceived great differences between the tests they practiced with and the test they encountered.

F-(adv.): “The practice exams were not very similar. The real CAEL is harder and had a lot more summary parts; in the practice there were not so many summary parts.”

B-(high-int.): “I feel that this test was more difficult than the practice, especially the reading was very difficult. I could understand the question’s meaning, but I could not find the answer and usually when we did the practice, the answer was

easy to find. If the practice is more difficult then the CAEL will be better.”

Importantly, participants were not as much concerned about the difficulty level of the final CAEL as they were about the *differences* in difficulty compared to the practice tests. They felt that they would have been able to prepare better if the practice tests matched the difficulty level of the final CAEL. The following comments by participants F and A reveal that the perceived increase in difficulty level could in part be attributed to an unfamiliar question format and a consequent mismatch of the strategies they used in the test.

F-(adv.): “In the practice test we just need to skim and scan, but with this one it didn’t work because we should summarize. A lot of things we must do ourselves, not just take from the reading, it is very different.”

A-(int.): “Using our strategies for this CAEL is not very useful because some answers I couldn’t find. For some questions I could not find the keywords in the reading.”

All of the participants reported that using the skimming and scanning strategies that they had practiced during their classes did not help them find the answers in the reading section. Participant C commented that she had always felt that the reading section was quite easy, because all that was required was to find the keywords and copy the appropriate sentences. She noted that, in contrast to the practice tests, there were several questions in the final test that called for a summary of the ideas presented in the articles.

Consequently, she felt that “you should spend time thinking, not finding the answers in the article.” Participant F also made reference to the differences in question format and expressed the following comment in frustration and anger:

F-(adv.): “There was a big form in the last part of the reading and it is not concentrated in one reading but in three parts, so I think it made me surprised. I almost did not know what I should do. But there were a lot of blanks that I did not fill in and a lot of my friends told me that all of them felt surprised and it’s awful. The reading was most difficult only because it’s different from the way we have done it before.”

There was agreement among all of the participants that, even though they usually perceived the listening or the writing section as the most challenging part of the CAEL, the reading component was most difficult part of the final test because of the type of comprehension questions posed.

B-(high-int.): “Usually, when we made the practice I think that listening is difficult for me, but this time reading is more difficult for me.”

D-(low-adv.): “Usually, the most difficult part is writing, but this time because of the type of question it is reading.”

The findings reported in this section indicate that familiarity with test format played a

central role in participants' perceptions of the CAEL. Although the integrated skills format of the CAEL differed from the test formats that participants had previously experienced, they expressed overall positive attitudes about learning "true and useful things". However, these positive attitudes were somewhat clouded by the test-taking experience. Participants believed that by completing the practice tests they had familiarized themselves with the format and difficulty level of the final CAEL; they thus expressed surprise and frustration when they encountered unexpected differences between the practice versus the final test versions. The difficulties that participants reported for the reading section may in part be explained by a narrow focus of skimming and scanning strategies in their test practice. Finally, being familiar with the difficulty level of the "real" CAEL was more important to participants than the actual difficulty level overall.

### **Contextual Washback Factors**

The washback factors discussed in the following section emerged from the analysis of the themes relating to the wider *contextual* factors of the testing situation. In contrast to the *test-specific* washback factors, which relate to a specific test, *contextual* factors refer to extra-test factors that are likely to influence washback on learners in any given high-stakes assessment situation. In aiming to identify the contextual factors through learners' perceptions toward the CAEL from a macro-level perspective, it was considered important to first explore participants' attitudes toward English, the role of English in achieving their goals, as well as their attitudes toward studying in the EAP program. This foundation seemed to frame participants' observations on test practice as well as their test-taking

experiences, which in turn appeared to inform their judgments of test consequences and fairness.

### **Pre-test responses**

The contextual washback factors that are discussed in the following section emerged from the analysis of the first set of interviews with participants, which were conducted at week seven of the twelve-week EAP program.

### **Attitudes toward English**

Five of the participants expressed their attitude toward English in terms of the overall importance and usefulness of the language. At the most fundamental level, the comments of participants A, C, and E indicated that English played an essential role not only for their individual goals, but also for their society at a global level.

A-(int.): “English is very important to me, to all Chinese people.”

C-(low-adv.): “In China all students study English. You have to, of course, because English is important for our goals.”

E-(adv.): “I want to master English; it is very useful. Even though Chinese is spoken by most people in the world, English is a worldwide language.”

It is notable that the importance and usefulness of learning English in all of the above

comments was linked to the level of Chinese society as a whole. Participants B and D further noted that English would be essential for them beyond their studies, for areas such as business and international communication. The view that English plays an important role at a societal level may also carry implications for participants' motivation and success in language learning. For instance, while investigating the relationship between learners' goals and their success in second language learning, Gilette (1994) found that "a learner's goal depends on a learner's social history and the use value ascribed to foreign languages in his or her environment" (p. 210).

While all participants were hoping to complete a university education, some further noted that their overall goal was not only to obtain a degree, but also to master the English language.

E-(adv.): "My father told me that my job is to master the language, not just getting a degree."

F-(adv.): "English is more important than a degree. In China my advantage is not the knowledge of science but my advantage is the language I learned."

Although these participants' comments indicate that their purpose for learning English transcended the attainment of a degree, their motivation was nevertheless directed from the outside, from a parent or society. Studies in learner motivation traditionally draw a distinction between factors that influence motivation from within the individual (intrinsic) and those present in the environment around the individual (extrinsic) (Gardner and

Lambert, 1959). Pierce's (1995) suggestion that motivation should be more appropriately viewed in terms of "investment", indicating that learners invest in learning a second language in order to "increase their cultural capital" (p. 17). This observation offers another useful perspective on participant F's comment that language, not scientific knowledge, would serve her in a competitive job market. Overall, participants' responses indicated that they generally had a positive attitude toward learning English, but were motivated by mainly pragmatic reasons (see LoCastro, 2001).

### **Goals**

All of the six participants expressed similar views on the overall role of tests in their learning. They emphasized that tests provided them with a clear aim to which they could direct their efforts, acted as a motivator, and also served as evidence for the hard work they had invested in their studies.

F-(adv.): "The CAEL is affecting me positively of course, it is the clear aim I want to reach. I think that tests are necessary because all people are naturally lazy. So we have to somehow set specific goals, to set a clear aim, so that we can work very hard for it."

D-(low-adv.): "My father always says that if you do something and you don't have a test, you will not do well because you have no stress to force you to do something."

C-(low-adv.): "I think the CAEL is like an equipment to show how hard you study."

Participant F's comment that the CAEL represented a clear goal that she wanted to attain, exemplifies how an important test can serve as a powerful motivator. In particular the test acts as an external force that these participants perceived in some sense as a "tool" to push them to work harder. The relevance of setting clear objectives in learning achievement is emphasized in the goal-setting theory proposed by Locke and Latham (1984). The premise of the theory is that specific and challenging goals will lead learners to approach tasks with increased attention and perseverance and will thus result in better performance. This perspective also emphasizes that specific goals, rather than vague or general goals, are more reliable in directing learners' actions and result in clear expectations.

In addition to acting as a motivating force, three of the participants also expressed the relevance of the CAEL in providing feedback on the work that they had invested during the term as well as in demonstrating improvement.

B-(high-int.): "We need a test to prove which level we are, and if we have some improvement a test can be a demonstration."

D-(low-adv.): "We know that we do hard work all semester and need some feedback."

E-(adv.): “I was excited to take tests because tests would show my ability in English; tests are a way to examine your ability.”

Participant E appeared particularly enthusiastic about taking tests and reported that he had generally performed very highly in tests in China. His favorable attitude is reflected in the finding that successful students hold positive views of standardized tests because they consider them as an opportunity to evaluate their English proficiency (Gan, Humphreys, and Hamp-Lyons, 2004).

The preceding comments illustrate that participants generally held the CAEL in positive regard, because it represented a clear goal, served as a motivator, and also provided evidence and feedback for hard work. At the same time, however, passing the test, which was participants’ immediate goal, also represented a barrier preventing the participants from reaching their ultimate goal of entering university.

B-(high-int.): “Their goal is not to pass the CAEL test, but before they enter university they must pass this test. No matter which level I am, I think this has given me stress because it will decide whether I can enter university. It is a very big problem for us; it can directly influence our future.”

Participant B expressed frustration that a test was simply imposed on students, especially because it directly affected students’ opportunity to gain access to university. The following comment by participant D also illustrates that obtaining a test score was his first priority in order to reach his ultimate goal of attending university in Canada.

Interestingly, he was concerned only with a “passing number” of any given test that would grant him access.

D-(low-adv.): “I came to Canada and just wanted to attend university. So the first thing I just wanted to get my TOEFL score. I need a chance to get into university, so any way that I will get into the university I will try. Here I am the oldest one, so I need to get into the university soon, as fast as I can.”

As is mentioned by participant D, time is of central importance for students who are still waiting to begin their university studies. For example, participant B noted that she felt time pressure to pass the CAEL quickly because it would allow her to enter university sooner, graduate earlier, and pursue a graduate degree earlier. Five of the participants raised the concern of time pressure experienced because of their current situation compared to their desired or expected situation. For example, participant E stated that he felt time pressure because all his friends in China entered university last year, whereas he was still trying to begin his studies. Participant D noted that many of his classmates were already taking courses at the university and he felt pressure to catch up to them.

Although all of the participants were in their first term of EAP studies, time pressure and desire to pass the CAEL quickly was a great concern to all of them. Several of their classmates however, were attending their second or even third term in the EAP program and still had not been able to achieve a passing score (i.e., band level 50 or above) on the CAEL that would allow them to begin university studies. As the following comments illustrate, for these students the CAEL not only represented a barrier blocking their goal, but also a source of anger.

A-(int.): "Most people feel angry with the CAEL because they are prevented from entering university and they want to finish their courses quickly. Maybe their English level is good enough so they feel upset that they didn't pass it."

B-(high-int.): "Some students have studied for one year and they haven't passed the CAEL. They are very angry, they have strong stress."

The strongly negative emotions expressed by students who had failed the CAEL on two or even three previous occasions are an important factor to consider with regard to the learning context. It may reasonably be assumed that anger expressed among students about the CAEL may reduce motivation for language learning as a consequence of cynicism and other counter-productive feelings.

Although all of the participants expressed a desire to pass the CAEL quickly because of their experienced time pressure, participant A's comment exemplifies a potential conflict.

A-(int.): "We want to finish our courses quickly. If we can't pass the CAEL we will spend more time here and we don't want to do that. It's very conflicting; we want to pass quickly but we also need the skills for university."

Thus, on the one hand, participant A wanted to pass the test quickly, but on the other hand, she recognized that she also needed the skills to do well in university. This view was

also reflected in the attitudes that participants held toward studying in the EAP program.

#### **Attitudes toward the language learning environment**

All of the participants stated that studying in the EAP program was preparing them to attend courses in university. Importantly, as the following comments by participant C and D illustrate, they regarded the abilities and skills developed in the EAP classes as essential ingredients for performing well in their university courses.

C-(low-adv.): "If you take a course but you don't have the skills and you can't understand the professor that is a problem. Studying ESL is very necessary. If you stay here to study ESL it lets you warm up."

D-(low-adv.): "If you go into university before you understand what the professor is saying, you cannot study and pass the exams. Now this situation is best for me because my oral English and my listening need more practice; so I think ESL is useful for me."

Although all of the participants generally felt that the EAP program was overall useful and necessary for them, they regarded the role of their classes in the development of their listening and writing abilities to be of greatest importance.

A-(int.): "I think that ESL class is very useful for us. Before I entered the class I don't know how to write an essay, I don't know how to listen and take notes."

Participant A further observed that, even though she had practiced writing and listening skills in her English classes in China, the form of that practice was very different. In particular, the writing tasks were based on models that students were required to imitate, while listening practice consisted mostly of short conversations with multiple-choice responses. Participant E reported that, when he first attended his EAP classes, he felt confused about the differences he encountered between Chinese and North-American writing styles, but was also eager and interested in acquiring this new writing knowledge.

E-(adv.): “The ESL program has many areas that I never touched on before. The most important help from ESL class I think is writing; it’s very helpful and formal. If we don’t take ESL class we may not know how to write. We know our ways, Chinese ways, but I think it’s too different, we won’t get a good grade from that.”

When discussing the EAP program, all participants talked most emphatically about the importance of having gained knowledge in writing. For example, participant B reported that before attending the EAP program, she would use only her own thoughts to write an essay and had never been required to integrate or summarize information from articles in her writing. Both participants D and C stated that they felt confused about writing essays at the beginning of the term, but after seven weeks of studying in the program, had gained confidence and knowledge in writing. Interestingly, participant A noted that she felt that students could study all other skills at home by themselves, but definitely needed a teacher to help them with writing. The importance that participants

ascribed to writing skills may also influence the attitudes of peers toward writing.

B: "At the beginning I think: just pass the CAEL. But now some friend told me that if you have good writing skills and a high level of English, you can study credit courses easier. He told me that writing is very important, so I began to realize that it is important for me."

### **Test practice in the classroom**

The attitudes that participants expressed toward the EAP program generally indicated that they felt it was useful, necessary, and helped them develop the skills that they would require for university studies. Another aspect of the EAP program that participants discussed was its role in preparing them for the CAEL.

A-(int.): "If I don't enter the ESL class, I don't think that I can pass the test."

E-(adv.): "The only way to study for the CAEL is to study ESL. What the teacher does with us will be useful for the CAEL."

The perception that studying in the EAP program was closely connected to studying for the CAEL appears to suggest that participants viewed the academic tasks practiced in class as generally reflective of the assessment tasks. Consequently, there appears to be an overlap of what participants regarded as "test practice" and the activities that were part of their standard class curriculum.

While all of the participants addressed the overall importance of the EAP classes in preparing them for the test, they emphasized different aspects of class activities that they regarded as test preparation. For instance, participant B stated that CAEL practice in the classroom was mostly focused on the essay component, which she felt was most important in the test. On the other hand, participant C felt that listening to tapes of lectures and taking notes was the most useful CAEL preparation during class. In addition to holding different views of the form of classroom test practice, participants also offered varied opinions about the amount of test practice in their classes.

B-(high-int.): "Every week we have one essay, but that is not enough practice because essays are very important for the CAEL test."

D-(low-adv.): "We did not have enough test practice, just two up until now."

F-(adv.): "I think the practice is adequate, we don't need a lot. If we have a lot of CAEL practice every week it's not like studying English, it's like preparing for a specific test."

These different viewpoints on what represented an adequate amount of test practice may reflect participants' comfort level with their abilities as well as the test itself. For example, participant D who did not feel that there had been enough test practice during class provided the following comment:

D-(low-adv.): “I have no idea how to prepare for the CAEL. On the web-site there are only two simple practice tests. It’s not enough information for me to prepare for this test.”

Because he felt that there was not enough information available on the CAEL, he reported feeling uncomfortable with the test and confused about how to approach test preparation. Interestingly, this participant had extensively prepared for the TOEFL test before joining the EAP program. As part of the TOEFL preparation classes he had attended, he was required to complete several practice tests on a weekly basis. Thus, his previous experience likely influenced the amount of practice tests he expected to access as well as his overall mode of test preparation.

#### **Teacher role in test practice**

As noted previously, because teachers were not required to follow any strict guidelines for CAEL preparation, it is possible that teachers took varied approaches to test practice. Furthermore, although teachers all attended the same “CAEL Orientation” meetings, it is likely that they also held differing beliefs and attitudes toward the test. The following statements further indicate that participants also had different perceptions of the role their teacher played in preparing them for the CAEL.

B-(high-int.): “Our teacher does not talk about the CAEL; he just tells us that writing is very important.”

C-(low-adv.): “He talks about it very often. From the beginning we know that the test has three parts.”

E-(adv.): “The teacher says that whatever we are studying will be useful for the CAEL test.”

These statements illustrate notable variations in how participants viewed “teacher talk” about the CAEL. Although, in the current study, such perceptions could not be confirmed with observational data, differences between teachers in test washback have been well documented (see Burrows, 1999; Watanabe, 1996; Cheng, 1999, 2004).

Three of the participants reported that their teacher played a very important role in their test preparation, because they felt that teachers know their learners’ areas of weakness and could thus best help in improving their English skills. For example, participant D regularly met with his teacher during office hours to receive feedback on his writing and was thus able to focus his improvements. Participant C also met with her teacher and benefited not only from her feedback, but also gained confidence in her language learning through these interactions. Notably, Gan, Humphreys, and Hamp-Lyons (2004) found that a majority of successful students maintain a good rapport with their English teachers.

### **Post-test responses**

The contextual washback factors that are discussed in the following section emerged from the analysis of the second set of interviews with participants, which were conducted

after the completion of the CAEL at week twelve of the EAP program.

### **Test preparation strategies**

The amount and type of test practice activities that occurred in the EAP classes were mainly directed by the teacher. Therefore, it is possible that participants directed their own test practice differently when outside of class. However, five of the participants reported that they did not feel the need to engage in additional “test practice” at home throughout the term, but focused rather on more “general” language learning activities. These activities included listening to the radio, reading magazines or newspapers articles, and watching English movies, in addition to completing their homework assignments. For example, participant F stated that it was more important to improve her overall English abilities in order to do well on the test, rather than practicing for the test specifically. Only participant D reported specific test practice, for which he had printed the two CAEL practice tests from the internet and had completed them several times.

However, during the days before the final CAEL, all participants did begin to focus on test preparation outside the classroom. Importantly, the following statements suggest that preparation activities varied according to participants’ general learning preferences and past experiences with taking tests, rather than being determined by the format of the test they were preparing for. For example, participant E placed great emphasis on learning vocabulary, which was a practice he acquired while preparing for the TOEFL test. He felt that memorizing vocabulary was also important for the CAEL, because it might help him in understanding the readings.

E-(adv.): “For both the TOEFL and the CAEL you have to practice grammar points, and the meaning of vocabulary. I just opened my books, memorized some vocabulary, thought about what the teacher told us and what we learned. I summarized the classes, thought about what kinds of abilities I got, but not think very hard.”

Although participant E memorized some vocabulary and reviewed his classes, he also felt that it was important not to think too hard before the test. Indeed, all of the participants held the view that it was important to be relaxed before the test because it would help them to perform better. Participant B stated that there was not much purpose in engaging in more intensive preparation before the test. She committed two or three evenings before the test to preparation, mainly reminding herself of writing structures that she wanted to use for her essay.

B-(high-int.): “I can’t prepare anything for the CAEL test. I think it is a collection of my English studies. I just tried to relax. Before the test I did some practice listening and writing, but just to see some essay structure and transition words.”

Participant F focused on reviewing test strategies that she had learned during a CAEL workshop that she attended the week before the test, but was most concerned about feeling relaxed when attending the test. Thus, the evening before the test she stopped all review activities and watched an English movie. Similarly, participant C did not engage in any test preparation but aimed to place herself in an English “frame of mind”.

C-(low-adv.): "Two or three days before the test I tried not to speak Chinese. I tried to watch TV and listen to the radio, read newspapers, do more things in English."

During this last phase of preparation for the CAEL, one common source of anxiety that participants reported was a sense of uncertainty about their current level of English proficiency and about their performance on the test. For example, participant B stated that, even though she had been doing well in class, she was unsure about her overall English ability.

B-(high-int.): "I didn't know whether I have improved and I can't say which level I am. I haven't the confidence to attend the last test. If I have more confidence I can relax."

Participant C expressed her uncertainty about what exactly she needed to do in order to be successful in the CAEL. She was also unsure about the level that she might achieve in the test. Both of the participants indicated that they would have felt much more comfortable about attending the test, if they had had some sense of where their current ability level was in relation to the test. Because none of the participants in the study had attended a CAEL before, another source of uncertainty was whether the final assessment would match their expectations.

E-(adv.): "A friend gave me a CAEL test that was very hard compared to what I took in class. I am confused about what the CAEL is really like."

### **Parental expectations**

Another common source of anxiety during the last days of test preparation were parental expectations. Four of the participants stated that they experienced considerable stress because of the pressure that their parents placed on them. For instance, participant E said that his father held the belief that if a student lived abroad and studied really hard, he would be "the same as a native speaker within one or two months". Naturally this participant knew that, no matter how hard he worked, he would probably disappoint his father's expectations. Participant B noted that her parents did not understand the importance of the test that she would need to pass, and simply said the same thing as many other parents: "study hard and try your best". Participant F also commented on her parents' lack of understanding of the type of assessment she would need to successfully complete to gain access to the university.

F-(adv.): "They don't understand English, so they don't understand what kind of English test I will need to pass. Every time they call me they ask "Have you got permission to enter university?" This question forces me to work hard."

Participant B stated that she did not tell her parents when exactly she would attend the test, nor did she speak with them during the week of the test. She was afraid that they would worry about her, and that this would place additional pressure on her to perform

well. Parental pressure was especially fierce for participant C, who said that her parents were angry with her because she was not yet studying her major and told her: “if you can’t pass the test go back to China!” Only participant A reported that she found a source of comfort and confidence in her parents.

A-(int.): “Before the test my father called me and said “don’t be nervous, trying your best is ok. I believe you can pass it.”

### **Affective reactions**

Not surprisingly, most of the participants expressed that they felt nervous on the day of the test. One exception was participant F, who reported that the day felt like a regular day and that she was *too* relaxed. She indeed wished that she had been more nervous in order to be “in the right mood to write the CAEL”. While a certain degree of nervousness is to be expected and serves to focus attention on the task at hand, most of the participants were upset at the nervousness they experienced because the difficulty of the test and the reading comprehension questions did not match their expectations.

B-(high-int.): “We thought that the test is the same as the practice and when we came to the test our confidence was lower than before.”

A-(int.): “During the test I did not do well on the reading and that made me feel nervous. I think that I can’t pass the CAEL, but most people did not feel very well.”

C-(low-adv.): “I felt really, really bad. I felt that I could not pass the test because the reading was hard, harder than before.”

F-(adv.): “I don’t remember the listening part but the reading part is terrible, it’s painful! I think that our teacher gave us easier practice exams and we thought that this should be the average level and that we could pass it. But at last it’s harder than what we expected.”

As previously discussed, participants felt that the type of reading questions in the final CAEL were very different from the questions that they had encountered during their practice tests. Participants stated that this difference in questions caused them great confusion and frustration. In this regard, an interesting consideration is offered by Cohen (cited in Bradshaw, 1990), who suggested that it may be the difficulty of the test technique or “the amount of cognitive load involved and the frustration experienced”(p.26), which results in negative reactions to a test, rather than the difficulty caused by low English proficiency.

In addition to confusion and frustration, participant D also believed that once emotional reactions began to interfere with the completion of a given test task, they would interfere with performance on subsequent tasks.

D-(low-adv.): “If you can imagine that long lecture and lots of stress, maybe you will get nervous if you missed the first question and get into a mess at the next.”

It is important to emphasize that participants reported not only that they were anxious, but they experienced *increased* anxiety because their expectations were not met. The relevance of anxiety levels is highlighted by Wine (cited in Bradshaw, 1990), who found that test-takers who were most anxious about their performance were “likely to be diverted from task-relevant activities and this diversion will disrupt information processing” (p. 15). Significantly, the author notes that this interference with the processing of information may lead to lower performance and, consequently, inaccurate measurement of test-takers’ actual ability levels.

In addition to the concerns discussed above, participants also felt that their performance was hampered by time pressure experienced during the test. Interestingly, none of the following three participants referred to the need for time simply to complete questions, but rather lacked the time to think and consider their responses carefully.

C-(low-adv.): “There is too much to do and we don’t have time to think carefully and write very exactly.”

A-(int.): “We just listen and take notes directly. We don’t have any time to think.”

F-(adv.): “I maybe lack the skill of writing an essay in a short time. If I want to write a good essay I need more time. I think just the time is the problem.”

This concern for additional time during test-taking was also mirrored in findings by

Jennings, Fox, Graves, and Shohamy (1999) who reported that test-takers considered the amount of time available to complete the test as the most significant factor affecting their performance.

In the current study, five of the participants, on the basis of their overall test experience, concluded that they performed poorly and that they would likely not pass the test. The comments of participant C and E also illustrate that many other students who took the test had similarly negative reactions about their performance.

C-(low-adv.): "We said that we will be returning. It's too hard; we can't pass it. Almost all the students said they did very badly."

E-(adv.): "Some said it's harder than the last test, very hard. For most of them the result won't be very good, so they are prepared to do ESL next semester,"

### **Plan of action**

The perception that they had performed poorly on the CAEL caused the participants considerable worries about failure. The comments of participant D and B reflect that they were concerned about missing their only chance to gain access to university for the next term.

D-(low-adv.): "We worry about what we are going to do if we fail. Sometimes I think, and my classmates also think, if you have just one chance to take the CAEL and you fail, people are going to be very disappointed."

B-(high-int.): "It is very serious. I am always worried about my test result because this is the only test. If I can't pass it, I need to study another semester and it will take time."

The remaining four participants expressed similar concerns about the consequences of possible failure, which was evidenced in their careful consideration of the actions they had planned to take if they did not achieve a sufficient test result to enter university. Before the final results of the test were available, all of the participants had already decided what plan of action they would follow depending on their result. For example, three of the participants planned to write the TOEFL or another test in order to obtain a score that would grant them access to university.

C-(low-adv.): "If I can't pass the CAEL I will study for the TOEFL; yes that is my plan. When I see the CAEL mark one is pass, one is not pass. If it is pass it is good and I will stay here. If I can't pass it I will go back to China to study for the TOEFL. I didn't tell my parents, this is my plan."

D-(low-adv.): "I don't know what will happen when I take this test. If I will fail I will fail and I will find another way to get a higher score and try and get into university. One of my classmates did not feel good about the test, so the next day he went to take the TOEFL test."

F-(adv.): "The whole term I would learn for the TOEFL and not learn ESL. I would go back to China and learn for the TOEFL there."

In the event of failure these participants maintained that, even though they felt it had been valuable to attend the EAP program and held a generally positive attitude toward the CAEL, they would nevertheless choose another test that they could possibly pass more easily. For example, participant F stated that if she were to attend another term of the EAP program, she still would have no guarantee that she would pass the CAEL at the end of the term. In contrast, she felt that studying for the TOEFL was more efficient because she could solely focus on test practice. The reason that participants F and C gave for returning to China to study for the TOEFL, rather than staying in Canada, was that the test preparation classes in China would be cheaper and more intensive than in Canada. In her plan of action, participant A had also decided not to repeat the EAP program, but instead would prepare for the CAEL on her own.

A-(int.): "You can go home and you know how to practice for the CAEL, you can do it by yourself and use a short time to do it. It's faster and more effective to prepare by yourself."

She felt that returning to class was not useful, because the teacher would focus on weaker students instead of pushing the higher students to perform at their best. Practicing at home would allow her to proceed at a pace that she felt would suit her best and she would not be frustrated by slower students. Participant E stated that he would return to the

EAP program, but would ensure that he had at least two chances to achieve a passing test score.

E-(adv.): “I would feel upset and very bad if I failed. I think I would stay and do ESL again, the next test I would pass. If I have time I will do both the CAEL and the TOEFL. I was always planning to prepare for the TOEFL outside of class.”

Only participant B noted that she would repeat the same path of studying if she did not pass the CAEL.

B-(high-int.): “If I can’t pass I will do another ESL, so I should work as quickly as possible and maybe next time I can pass it.”

Notably, in the range of plans that five of the participants had forged in the event they did not gain access to the university, all included attending a different or additional test. While it is conceivable that the courses of action that participants had chosen reflected their beliefs about the meaning of failure, this may not necessarily be the case. For example, participant B’s decision to attend another term of the EAP program if she could not pass the test may have been based on her opinion that such a result would indicate the need for further development of her writing skills as well as additional English practice. However, participant F, who also stated that not passing the exam would mean that she still lacked the necessary abilities for university studies, felt strongly about returning to China to practice for the TOEFL. Thus, while these participants held similar beliefs about

the meaning of test failure, they had decided on different courses of action in the event of failure.

In addition to reflecting a test-taker's lack of English ability for university studies, participant E emphasized that a poor result would also indicate a lack of effort.

E-(adv.): "If I do bad on the CAEL then I don't think that I studied very hard in ESL. I did my best and the result will be the true result of myself. If I fail the CAEL it's because my ability in English is not good enough and I still have to study. The result is what I really am; it is true. It's the result I worked for, how much or how hard I worked, so I don't say that the result is not what I really got. I will accept it and do my best next time. "

The significance that this participant ascribed to effort is also reported in numerous studies which have found that Asian cultures attribute success to effort, and failure to lack of effort (Biggs, 1996).

## **Results**

In the final phase of the test experience, the concordance between result expectations and the actual result appeared to represent an important factor affecting participants' perceptions of test fairness. Participants had been asked about the result they had anticipated and whether the final result had matched their expectations. In estimating their performance, participants drew a distinction between the result they really desired and one that they could reasonably expect. Although several participants wished that they could

achieve a band score 70 or 80, there was a common assumption that the highest score they could possibly achieve would be an overall band of 60 and that scores above 60 were beyond their reach. The following comment by participant C illustrates her classmates' feelings about the difficulty of achieving their desired result in the CAEL.

C-(low-adv.): "There is a joke about the CAEL, we usually say KILL because you need 60 out of 90 to pass."

In relating the compatibility of their expectations and their test results, participants reported a range of reactions. For example, participant B had initially expected that she could achieve a band score of 60 but, feeling that she performed very poorly on the test, she had lowered her expectations to 40 or possibly 50. Therefore, she was very surprised and delighted when she received a result of 60, which met her initial expectations. The following three participants all achieved an overall band score of 50, but their individual expectations led them to receive the result with differing reactions.

C-(low-adv.): "I was so surprised to get a 50. I thought I would be returning. I did it pretty bad."

A-(int.): "70 is too high, so I thought that I could pass a 60. Passing a 50 is good and I can choose some credit courses, so I didn't feel nervous.... I got a 50 and it's not my expected result, I wanted to get a 60."

F-(adv.): “50 is not a good result. I thought I will pass it and the result is no good; it’s not the same as I thought. I hate the exam. I hate the result.”

Test-takers who achieve an overall band score of 50 in the CAEL have the opportunity to register for three credit courses in addition to attending a non-credit EAP writing course. While participant C was very content with having achieved this result, participant A experienced some disappointment about not having gained unrestricted admission to university. Indeed, she felt regret that she had not taken the TOEFL in addition to the CAEL, because attending two tests might have offered her another chance to receive a full passing score. Participant F, however, was truly angry about having achieved a band score of 50. Her expectations had been very different because of the grades that she had received in her class. Furthermore, her parents had expressed disappointment with her performance, and she experienced great guilt and remorse for not having studied harder.

F-(adv.): “I feel pressure from my parents. They are not satisfied with my result and said that maybe I didn’t work hard enough.”

The result that participant F achieved is considered “successful” by teachers and university administrators, because a band score of 50 allows a student to begin university studies. However, because she would still need to attend one non-credit EAP writing course, participant F felt that she had failed the test. Furthermore, her parents’ judgment of her performance affected her deeply. Her sentiment is reflected in findings by Biggs and

Watkins (1996), who observed that in Asian countries “standards of excellence and what constitutes success may be determined both by the individual and by significant others including family, the group, or the society as a whole” (p. 274).

### **Fairness**

Participants’ judgment of test fairness was composed of several factors including: the concordance of participants’ achievement expectations and their test results; compatibility of test results with perceived ability levels; fairness of others’ test results; and overall fairness of a “single-instance” testing situation.

Participant F experienced frustration and anger about the mismatch of her usual class performance compared to her test performance, which fueled her skepticism about the ability of the test to accurately reflect her language ability.

F-(adv.): “I don’t know if the result can measure my skill. When I wrote the essays in class I got very high marks, so I think that my English skill is high enough. But at last, in the CAEL exam I didn’t get high marks.”

Of the six participants, five felt that the test generally reflected their listening and writing abilities, but judged the reading score they received in the test to be below their actual performance level.

D-(low-adv.): “The result is kind of fair. Sometimes the reading score I get is not the real ability. This time my reading was low. If I did another test I think my reading would be fine.”

C-(low-adv.): “I know I am reading very poorly in this test. That does not mean that before the final test I did reading very badly; I did reading very well before the final test.”

The indication that five of the participants did not feel that the reading score accurately reflected their true ability could be explained in part by the increased nervousness that they experienced during this component of the test. For example, Bradshaw (1990) suggested that performance of some test-takers “may be disrupted by aspects of a test which increase the amount of anxiety which they experience, or which distract them in some other way from their task” (p. 15). Furthermore, as discussed previously, the apparently narrow skimming and scanning focus of participants’ reading strategies did not seem suitable for certain question types and may thus also have contributed to a lower reading performance. Notably, the overall perception that their reading ability was not accurately reflected in their test result led most participants to question the fairness of the reading section of the CAEL.

### **Fairness for others**

Another area where participants expressed judgments of fairness concerned the comparison of their results with the results and ability levels of their classmates. For example, participants A and D were concerned that some students, whose ability they perceived to be similar to or higher than their own, did not achieve a good result in the CAEL.

A-(int.): "For me it's fair. For some other people it's not fair. Some students' English level is better than mine, but they didn't pass the CAEL."

D-(low-adv.): "We know that my friend is a good student and he did the ESL class very well, but he got just one credit course."

In contrast, participant F expressed frustration that, although she consistently received one of the highest grades in her class, several of her classmates performed better than her in the test.

F-(adv.): "The classmates that passed the exam THINK that it's fair, and some other people that did not pass the exam don't think it's fair. Some of my classmates passed the exam and I don't think that their skill is higher than mine in the class. In my class my score is the highest, so I think that I should get the best mark of my class but I didn't."

She felt that students who received lower grades than her in the EAP program did not deserve to achieve a higher score than her in the final test. To her, this was an instance of unfairness and led her to further question whether the test accurately reflected learners' ability levels.

#### **Fairness of the testing situation**

Participants were also concerned about the fairness of assessing learners' English

abilities under very stressful conditions. In particular, participants suggested that a mismatch between ability level and test result can be attributed to the nervousness and stress experienced by many students.

A-(int.): “Some students who work very well did not pass the CAEL. It’s not because their English level is not good, it’s because they feel very nervous.”

B-(high-int.): “Maybe that day we have some physical condition; maybe we have some problem. Usually we have a high mark but when we don’t have a good condition it will influence our score.”

D-(low-adv.): “Sometimes if we get too much stress, we cannot show what we learned and just get a mess and we have no chance to show our ability. Sometimes good students have a lot of pressure and they will be nervous for the test; so when they are doing the test they will make a mess.”

The concern that participants expressed about the influence of a learner’s emotional state on test performance has also been discussed by Bradshaw (1990). She emphasized that considerations of affective conditions should be of central importance in the testing process, given the evidence that an “excessive degree of anxiety can have debilitating effects on the performance of some test-takers” (p. 15). In order to determine the role of test-taker variables in test performance, Shohamy (1982) also recommended that factors such as anxiety and attitudes toward the testing experience must be further examined. She

observed that these affective factors are often ignored by test developers and users, although they may strongly influence achievement and performance on a test, and could lead to inaccurate assessment.

Finally, participants questioned the fairness of making judgments about a students' overall English ability on the basis of a single test performance.

C-(low-adv.): "The test can just show THIS time. It is not just your knowledge about English, it also depends on your feeling and the test's topic."

Although participants were concerned about the role of affective factors and the test topic when making fairness judgments, they stated that test fairness was also compromised by the value ascribed to a single test performance. In particular, they considered it unfair because some students, who did not invest effort in their English studies throughout the term, were given an equal opportunity to gain access to university as students who worked very diligently. Furthermore, they believed that valuing only one instance of their performance would not acknowledge all the hard work that they had invested in their English studies. Also, even though they did receive a grade from their EAP class, this grade could not be used for university admission purposes. Therefore, they felt that in the event of failure on the CAEL, all of their work would seem worthless and wasted.

C-(low-adv.): "It should not just be the final exam. It depends on whether you are a good student or not, whether you work hard and participate."

D-(low-adv.): “Sometimes we have full attendance, do our homework on time and we think we are doing well. But we cannot enter university because it depends on the final test, so it’s kind of unfair. It is not fair to the students who come to school every day and work very hard and if they fail in the final their three months of hard work will be gone.”

The sentiment of these comments is also captured in findings by Cheng (1998), who reported that students agreed most strongly with the statement that “examinations should NOT be used as sole determinant of students’ grades” (p. 296). Participants’ overall concerns about the fairness of the test condition as well as the single instance of assessment were also reflected in an investigation of test fairness by Sambell, McDowell, and Brown (1997). Students in this study felt that “end-point summative assessments...were actually considerably down to luck rather than accurately assessing present performance” (p. 362). Furthermore, students did not believe that examinations reflected their true ability, but maintained that success in exams depended on: whether they felt ill on that day; whether they had a tendency to panic; or whether they experienced extreme levels of stress. Another finding in this study that also reflected the attitudes of participants C and D, was that students believed success depended more on “consistent application and hard work, not a last minute burst of effort or sheer luck” (p. 363). Importantly, these students used the concept of fairness to address the extent to which an assessment truly values and reflects the work and effort that they invested.

In order to increase the fairness of the assessment process, participants C and D recommended that class performance should be counted toward the final university

admission decision. They also felt that this would increase the motivation for students to work hard during their classes because their efforts would actually be rewarded. Participants A and B presented an additional means of increasing test fairness, by suggesting that test-takers should be offered two chances to take the CAEL.

A-(int.): "The university should give me a second chance to do the CAEL. If I cannot pass then, maybe I should do ESL again."

B-(high-int.): "In China if we can't pass the final test our teacher will give us another chance. If you can pass it you can show that maybe you didn't have a good condition."

### Advice

During the last part of the second interview, all participants were asked what advice they would offer to a new student based on their experiences preparing for and taking the CAEL. Although many of the elements of their responses reflected what participants had communicated in previous parts of the study, these statements nevertheless appeared to exemplify an attempt to capture the essence of the factors they considered most relevant in the testing process. Insofar as students are likely to communicate their test experience and advice to others, the following statements seemed to express what had crystallized in participants' minds as some of the most important aspects of preparing for and taking the CAEL.

Two of the participants emphasized that effort and diligence played the most

significant role in their successful completion of the EAP program as well as the CAEL.

D-(low-adv.): “Follow what the ESL teacher asks you to do. Just follow it and work hard and you will be ok. We will spend lots of time on work, but finally it pays off.”

E-(adv.): “Work hard during the class. No sleeping, no lateness, no absences. Write your homework every time and you will do well. One of my friends didn’t get a good grade and he said it’s because the teacher didn’t give us much homework. Because I worked very hard on my study I got English 1020.”

This advice seems reminiscent of central Confucianist principles, in which education and learning are strongly associated with effort. The late 19th century scholar Hsun Tse (cited in Biggs, 1996) strongly expressed the significance of effort in learning by advising students to “sincerely put forth your efforts, and finally you will progress. Study until death and do not stop before.” (p. 32).

Participants A and B focused their advice on practicing listening skills, which had posed a great challenge to them. Also, their recommendations included a reminder of the unexpected difficulty of the reading component and the discord between the practice test and the final CAEL.

A-(int.): “You should learn more vocabulary and when you learn to do the reading you should shorten your time. You should practice listening more; you should listen, listen, listen.”

B-(high-int.): “Do more listening practice and take notes when you listen. You should do more difficult practice because writing this test is difficult. If you think that the test is the same as the practice your confidence will be lower”.

Participant C placed greater emphasis on the importance of immersion in an English environment in developing overall communicative ability. She did not feel that specific test practice or strategies were an important factor in achievement; instead she suggested that communicating with the teacher to learn about areas of weakness would be most useful for improving students’ language abilities.

C: “Try to not speak Chinese. Give yourself an English environment and try to speak more English. Communicate with your ESL teacher.”

Finally, participant F communicated her resentment toward the test in her final advice. She again expressed her disappointment and frustration that the practice tests had misled her to believe that the CAEL was easy.

F-(adv.): “The CAEL exam is not as easy as you think. If it is possible go back to China and do the TOEFL.”

## Summary

Studies on washback have generally focused on how a *test* affects the educational context, teachers, and learners; however, the question posed in this study reverses this focus, by exploring how *learners* and *contextual* factors surrounding the test shape the washback process. The goal of the current study was to explore washback on learners that may be the result of important *contextual* factors through an exploration of learners' perceptions toward a high-stakes test. The factors that emerged during the analysis phase of this study were categorized into *test-specific* washback factors and *contextual* factors that shape washback on learners. As the following summative discussion of results shall illustrate, although these factors may be addressed as specific categories, they also have some overlap.

The factors identified as *test-specific* washback appeared to be closely connected to the test properties and practice materials of the CAEL Assessment. These factors included test knowledge, test format, difficulty level, test perception, and practice-test mismatch. With regard to test knowledge, participants noted that they had never heard of the CAEL before joining their EAP program; in contrast, most of them were familiar with the TOEFL. Their comments illustrated that test knowledge and familiarity were closely connected to their test comfort and preferences. Also, the length of time that participants required to develop familiarity with the CAEL appeared to differ depending on their previous test experience; those having extensively prepared for the TOEFL prior to joining the EAP program indicating that they did not know enough about the CAEL to feel comfortable.

Participants' previous test experience was also evidenced in their descriptions of the CAEL format, which were based on comparisons to the TOEFL or other familiar tests. These juxtapositions served to illustrate which parts of the CAEL they perceived as new and different from their previous testing experiences. The test features that participants most emphasized were the topic-based and integrated skills format of the CAEL. Furthermore, participants also used comparisons between the tests to assess the difficulty level of the CAEL. The perceived difficulty level of the test varied depending on how participants viewed the test tasks and their requirements. For example, those who judged the CAEL as more difficult than the TOEFL explained that one needed to understand the full meaning of the articles and could not simply guess like in multiple-choice tests; whereas those who judged the CAEL to be easier than the TOEFL stated that one did not really need to understand the articles but was simply required to scan for key words. However, participants did not necessarily prefer an "easier" test, but instead placed greater emphasis on the meaningfulness of the test result. This meaningfulness was in turn expressed in participants' overall positive attitude toward the test format, which was asserted in their descriptions of the CAEL tasks as "true and useful learning", as well as viewing the CAEL indeed as a "fire test" because of the connection they perceived between the test tasks and the university setting that they were preparing to enter.

Although participants initially perceived the test format of the CAEL as novel and unfamiliar, they generally expressed positive attitudes toward the test. However, a perceived mismatch between the difficulty level and question format of the final versus practice versions of the CAEL affected participants' attitudes negatively. In particular, they stated that the practice tests were considerably easier than the final CAEL and had

instilled them with a false sense of confidence. Importantly, participants appeared not as much concerned about the overall difficulty level of the final CAEL as they were about the *differences* in difficulty compared to the practice tests. They felt that they would have been able to prepare better, had the practice tests matched the difficulty level of the final CAEL. In addition, participants expressed concerns about the unfamiliar question format they encountered in the final CAEL. They maintained that the reading comprehension questions in the practice tests could generally be answered simply by using skimming and scanning strategies. However, some of the questions in the reading section of the final CAEL required them to summarize information in their own words, rather than copying key phrases from the articles. This finding indicates that a narrow focus on certain strategies in test practice may cause test-takers difficulties in completing varied tasks and, importantly, may lead to interference in test performance as well as possibly inaccurate assessment.

The variables identified as *contextual* factors in the washback process were extra-test factors that did not appear to be closely connected to the CAEL, but instead could conceivably influence the washback effect of any given high-stakes assessment. These factors included: attitudes toward English, the role of tests in achieving goals, attitudes toward the language learning environment, test practice in and outside the classroom, teacher role in test preparation, parental expectations, affective reactions, as well as test result consequences and judgments of fairness.

Overall, participants held a positive attitude toward learning English, emphasizing not only the significance of mastering English to achieve their individual goals, but also the importance of the language for their society. Participants also expressed positive views of

tests in reaching their language learning goals, because they provided a clear aim to which they could direct their efforts, acted as a motivator, and also served as evidence for the hard work they had invested in their studies. At the same time, however, language proficiency tests represented an obstacle preventing the participants from reaching their actual goal of gaining admission to university. This barrier frustrated participants because they experienced time pressure to begin their university studies; this frustration indeed evolved into anger for students who had repeatedly failed the test. These emotionally negative reactions represent a concern, because anger expressed among learners toward the test may in turn adversely affect and possibly result in negative washback on the overall learning environment.

Participants generally held positive attitudes toward the EAP program, regarding the abilities and skills developed in the classes as essential for performing well in their university courses. Furthermore, they perceived an overlap between their class curriculum and “test practice”, while nevertheless holding different views on the form and appropriate amount of classroom test practice. Participants also reported notable variations in “teacher talk” and the overall role of the teacher in preparing for the CAEL.

Throughout the duration of the EAP program, most participants did not perceive the need to engage in additional test practice outside the classroom, but focused rather on more “general” language learning activities. In the days leading up to the test, participants did engage in additional test preparation activities at home which, importantly, appeared to vary according to participants’ general learning preferences and past test experiences, rather than being shaped by the format of the test they were preparing for. In this preparation phase for the test, common sources of anxiety that participants experienced

included uncertainty about their current level of English proficiency, how they would perform on the test, and whether the final test would match their expectations. In addition, parental expectations and pressures caused participants considerable stress, especially in the days before taking the assessment.

Most participants expressed somewhat negative attitudes toward the test-taking experience, because the difficulty of the test and the reading comprehension questions did not match their expectations. Importantly, this mismatch of expectations caused *increased* anxiety for participants and may in turn have led to lower than normal performance in the reading section of the test. In addition, participants expressed the need for more time, not simply to complete questions, but rather to think and consider their responses carefully. Concerns about possible failure of the test led participants to devise plans in case they would not be granted admission to the university. Although they regarded the EAP program as valuable and held a generally positive attitude toward the CAEL, in their plans of action most participants did not elect to repeat the EAP program, and all included attending a different or additional test that they could possibly pass more easily.

Judgments of test fairness were composed of several factors including: the concordance of participants' achievement expectations and their test results; compatibility of test results with perceived ability levels; fairness of others' test results; and overall fairness of a "single-instance" testing situation. In discussing participants' test outcome, the importance of the concordance between their result expectations and the actual test result was evidenced in differing reactions to the same test result; specifically, some welcomed it with relief while others rejected it as a failure. Although participants, overall, felt that the test fairly assessed their abilities, all of them considered their reading abilities

to be higher than indicated by their test result. This perceived underestimation of reading abilities may in turn reflect participants' reports of increased anxiety and strategy mismatch in this section of the test. In addition, judgments of fairness were influenced by the perceived compatibility of class and test performance as well as participants' result comparisons to other test-takers. Finally, the "single-instance" testing situation was perceived as unfair because external conditions (i.e. stress, nervousness, illness, etc.) could have affected learners negatively and because it did not value the work and effort that participants had invested in their English learning throughout the term.

As indicated previously, the quotes and themes discussed are not evidence that a certain type of washback *is* occurring, but rather provide information about what participants found significant in their experience in a high-stakes assessment environment, and what researchers may therefore want to take into account in formulating hypotheses about what factors are important in mediating washback on learners.

### **Introducing the Contextual Washback Model**

The present study identified two broad categories that may influence washback on learners, namely, *test-specific* and *contextual* factors. In particular, *test-specific* factors appeared to reflect participants' perceptions largely on the basis of specific properties of the test (e.g., test format). In contrast, *contextual* factors appeared to reflect participants' perceptions largely on the basis of extra-test factors (e.g., perceptions of fairness).

The distinction between *test-specific* factors and *contextual* factors is fuzzy. On the one hand, this distinction may be described as arbitrary in the sense that, because both sets

of factors appear to shape learner perceptions, they operate through the psychology of the individual. That is, *test-specific* factors may influence learners by the way they view or experience a test, and *contextual* factors may influence learners by the way they view or experience extra-test factors surrounding the assessment. However, the distinction is not arbitrary in the sense that whether we focus on factors that are *test-specific* versus *contextual* factors can shape our approach to conceptualizing and researching washback. For instance, those researchers interested in test factors may wish to look closely at the structure of a test to understand its influence on learners' approaches to language learning; as such, research and conceptualizations based on the study of the TOEFL may differ widely from those based on the CAEL. In contrast, those researchers interested in the role of context in shaping washback may wish to focus on, for example, the psychology, culture, learning environment, and the peers of learners in high-stakes language learning contexts, without any interest in the specific test involved. These findings would, in theory, apply potentially to any specific test.

The *Contextual Washback Model* (see Figure 5) illustrates the *test-specific* factors and *contextual* factors that affect the washback on an individual learner in a high-stakes language assessment learning environment. It is important to note that this model, in contrast to Bailey's basic model of washback (see Figure 3, Chapter 2) and Henrichsen's Hybrid model (see Figure 4, Chapter 2), focuses on washback that is shaped and experienced by an individual, rather than the washback effect that a test has on the level of the learning environment. In particular, because the current study focused on participants' perceptions of a high-stakes assessment environment, rather than on the structure of a given test, the delineated model uniquely reflects the social-psychology of washback. This

altered focus implies that a test cannot be assumed to have a specific overall washback effect, but instead may differ from learner to learner (see Alderson and Wall, 1993; Watanabe, 2002).

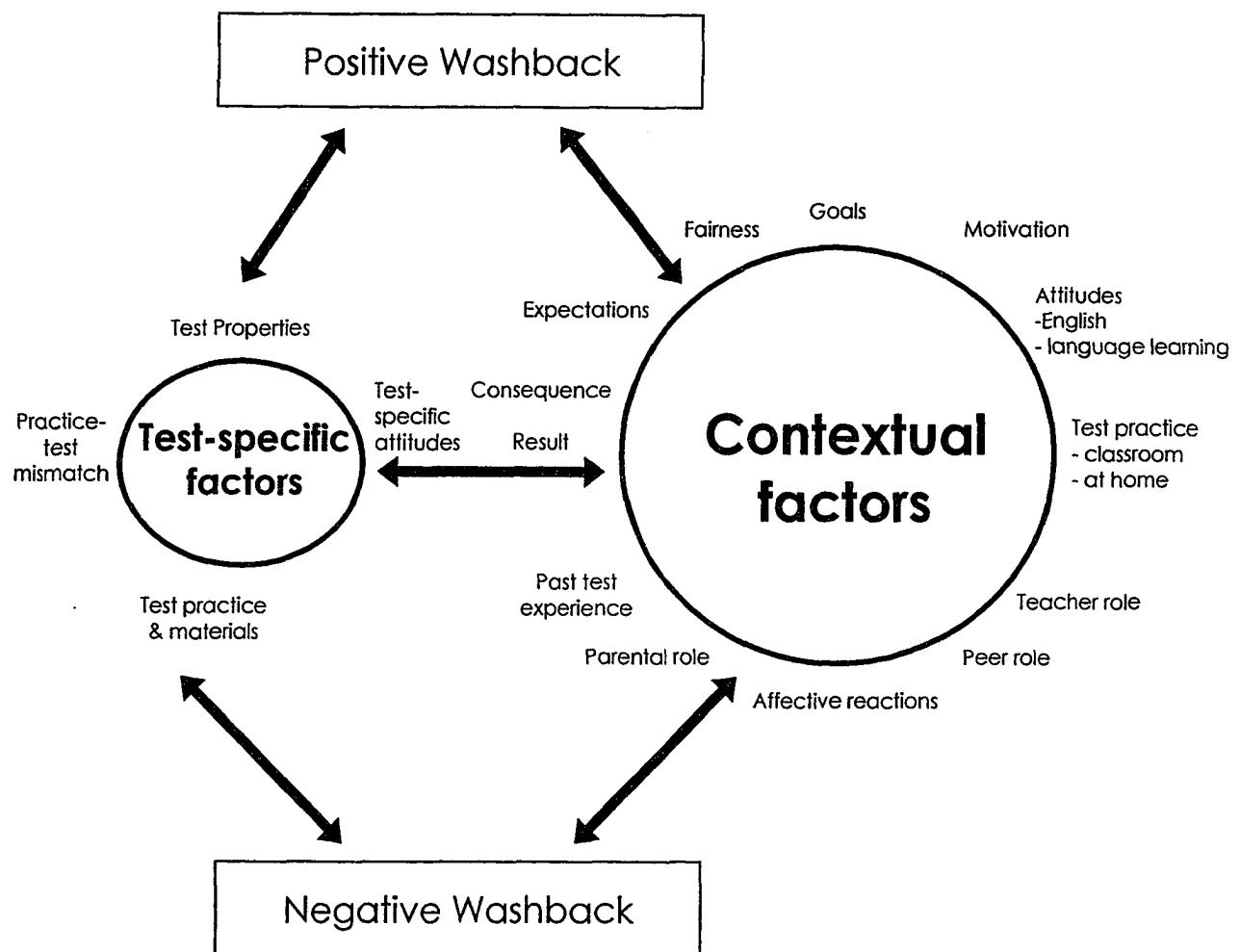


Figure 5. *Contextual Washback Model*. Factors affecting washback on the individual learner in a high-stakes assessment language learning environment.

The *test-specific* factors relating to the CAEL Assessment that were identified in the current study included test properties, test practice and materials, test-specific attitudes, and practice-test mismatch. These factors appear to be specific to a given high-stakes language assessment (e.g., test properties) and, consequently, the content of these factors would vary if this model were applied to another test. For example, as the CAEL is a topic-based assessment, the factor of “test properties” would differ for a test such as the TOEFL, which is based on a non-integrated, multiple-choice format. Variables comprising *contextual* factors are not specific to a given test, but rather may affect the washback of any high-stakes assessment.

The circles surrounded by the *test-specific* and *contextual* factors represent the *interactional space*, where various factors may influence each other. The potential complexity of this *interactional space* may be characterized as “dendritic”. Technically, the term dendritic refers to the branching process of a nerve cell, which can be treelike, tightly interconnected with other cells, and whose specific function may be difficult to disentangle. With respect to the current model, for example, the *contextual* factors of expectation, result, and fairness may share complex interconnections. In particular, because participant F had been one of the best students in her advanced EAP class, she expected to achieve a minimum band score of 60 in the CAEL (expectation). However, she received an overall band score of 50 and was outperformed by classmates who had lower class grades than her (result). Therefore, participant F felt that the CAEL result was unfair and also questioned whether the test could accurately measure her language abilities (fairness). Another example of possible connections among contextual factors was parental role and affective reactions. For instance, participant A reported that her father’s

encouragement (parental role) provided her with comfort and confidence (affective reaction) during the last days of test preparation. In contrast, participant C did not speak with her parents before the test because she feared that their expectations (parental role) would cause her to experience increased pressure and anxiety (affective reaction). The various possible connections among factors represent an important dimension of washback because they may provide insight into the actual processes that shape the positive or negative quality of washback.

The two categories of *test-specific* and *contextual* washback factors also appear to influence each other, exemplifying the overlap between the concepts. Importantly, this influence may be bi-directional rather than emanating only from the test. That is, *contextual* factors (e.g., past test-experience) may shape the way a learner views *test-specific* factors (e.g., test format). For example, participant E had practiced extensively for the TOEFL test at home, attended TOEFL test preparation classes, and had taken the test (past test-experience). He had thus developed familiarity with the TOEFL and felt that the CAEL was more difficult because it was “new” and “different” (test format). On the other hand, *test-specific* factors (e.g., practice-test mismatch) may influence a learner’s context (e.g., affective reactions). For example, participant F perceived a great difference between the reading tasks in the CAEL practice versus final test versions (practice-test mismatch). She reported that this task difference caused her confusion and increased her nervousness during the test (affective reactions). The possibility that *test-specific* and *contextual* factors influence each other bi-directionally may be an important consideration with regard to the evaluation of positive or negative test washback. In particular, this possibility would suggest that the quality of test washback would need to be investigated in conjunction

with *contextual* factors in order to evaluate the overall washback of a test. This possibility is supported by Cheng (1997), who suggests that a close examination of the educational context surrounding a test is essential in determining the influence of a specific test on teaching and learning (p.52).

Importantly, both *test-specific* factors and *contextual* factors could conceivably contribute to positive and/ or negative test washback. The degree of positive or negative washback would therefore be determined by the combined impact of all factors. For illustrative purposes, the elements of participant E's experience may be hypothesized as tending toward positive or negative washback in reference to *test-specific* or *contextual* factors. For instance, with regard to positive washback, participant E held an overall positive attitude toward English learning and the EAP program (*contextual*); he was highly motivated by tests and generally had positive experiences in taking tests (*contextual*); he liked that the CAEL format was similar to studying in university classes (*test-specific*); his test result met his expectations and he thus considered it fair (*contextual*). With regard to negative washback, participant E perceived the CAEL as new and unfamiliar to him and he felt more comfortable with the TOEFL (*test-specific*); he experienced considerable stress because of high parental expectations (*contextual*).

Admittedly, the extent to which a learner's reported positive or negative perceptions map directly onto positive and negative washback, can only be hypothesized but not directly determined, based on the data in this study. In addition, the findings of the current study do not indicate the relative importance of individual factors. For example, a participant may have reported overall positive perceptions and attitudes toward the test, but be most affected by a single negative factor (e.g., perceptions of test fairness).

Nevertheless, the current model provides a conceptual framework to better research the variety of washback factors that may affect individual learners in a high-stakes assessment environment.

The *Contextual Washback Model* emphasizes the findings of the current study that extra-test factors appear to play an important role in shaping washback on learners. The themes comprising *contextual* factors that emerged from the analysis of the data focused on participants' affective experiences and immediate surroundings. This finding is supported by Alderson and Wall's (1993) claim that the washback of a test can differ from student to student depending on their experience, circumstance and perception of fairness. The authors also note that the complex nature of washback is likely shaped by forces such as society, education, and schools; it is thus important to keep in mind that the limited number of individual factors represented in the *Contextual Washback Model* are surrounded by a wider contextual field of the educational system, culture, and society.

One of the key suggestions of the findings in this thesis is that washback researchers should focus greater attention on *contextual* factors (or more precisely on learner perceptions that may be shaped by both *test-specific* and *contextual* factors), rather than on the impact of the test alone, in order to better understand washback on learners and in order to understand how to foster optimal language learning in high-stakes test environments. Alderson and Hamp-Lyons (1996) offered a similar proposition in stating that it could be tempting to infer that the test alone does not cause washback but rather that the stakeholders themselves cause the washback. Importantly, this shift in attention away from the test itself may offer insights into the complexity of washback findings to date. Bailey (1996) noted that the paucity of empirical research on washback could in part

be attributed not only to the complexity of the phenomenon itself, but also to the difficulty of deducing what can be evidentially linked to the test. The findings represented in the *Contextual Washback Model* suggest that the complexity of the phenomenon may be greatly linked to the context surrounding any given high-stakes assessment, and that the *specific* test itself may not play a central role. This possibility, in turn, carries important implications for linking washback to test validity (Hughes, 1988; Frederiksen and Collins, 1989). In particular, the role of extra-test factors in shaping washback is important in relation to Messick's (1996) suggestion that increasing authenticity and directness of the test construct will minimize factors that produce negative washback. Furthermore, the role of *contextual* factors in shaping washback appears to suggest that the quality of the test may be independent of the positive or negative nature of the washback effect (see Alderson and Wall, 1993).

## **Chapter 5: Conclusions and Implications**

Many research studies have emphasized the importance of tests in mediating the washback process of high-stakes language learning environments. Although these investigations have identified a variety of factors that may contribute to washback, few have examined the role of learners and the test context. The goal of the current study was to explore washback on learners that may be the result of important *contextual* factors, such as learner attitudes toward language learning, affective reactions, past test experience, and other variables that shape learner perceptions within a high-stakes testing situation. The factors that emerged during the analysis phase of this study were categorized into *test-specific* and *contextual* factors. In particular, *test-specific* factors which were explored through participant perceptions reflected specific properties of the test such as: test knowledge, format, difficulty, test-specific attitudes, as well as test-taking experience. In contrast, *contextual* factors which were explored through participants' perceptions reflected extra-test factors such as: attitudes toward English and the language learning environment, the role of English in achieving goals, test practice in and outside the classroom, teacher role in test preparation, parental expectations, affective reactions, as well as test result consequences and perceptions of fairness. Importantly, the findings appear to suggest that *contextual* factors may play a significant, complex role in mediating washback. Based on the findings of this study, a *Contextual Washback Model* was delineated to broadly characterize the combined influence of *test-specific* and *contextual* factors on the quality (positive or negative) of test washback. Although this model does not include an exhaustive representation of possible washback factors, it serves to expand

current understandings of the nature and scope of washback and may inform future research.

The findings of the current study must be considered in relation to certain limitations regarding the participants, context, and nature of the investigation. The findings discussed are based on a small number of participants who volunteered to share their test experiences and may thus not fully represent the perceptions of learners who were unwilling to participate in the study. Furthermore, all volunteers were recruited from different classes within the same EAP program and therefore shared a similar learning context. Importantly, the EAP program situated the CAEL Assessment in a particular context, where test results alone, and not class grades, determined acceptance to the university.

As in many qualitative investigations, the possibility of biased responding on the part of the participants must be considered. A central question in the present study is whether participants were motivated to please the researcher, who was also coordinator of the EAP program. This possibility is mitigated by at least two reasons. First, participants expressed a balance of both positive and negative views on several aspects such as the EAP program, preparing for and taking the CAEL, as well as the results and consequences of the test. Second, throughout the interviews there were spontaneous expressions of emotion associated with many responses, suggesting a genuine engagement with the issues being discussed.

The findings hold implications for the use of the CAEL Assessment within the EAP program considered in this study. Participants' perceived mismatch between practice compared to final test versions, suggests a need for better test preparation that would

include practice tests with higher difficulty levels. Furthermore, the curriculum of the different class levels is not matched directly to the levels of the test. Consequently, there is no fixed progression of difficulty within the program that the CAEL levels reflect. Thus, what is learned and what is tested are somewhat different; the CAEL is thus used as crude proficiency test rather than an achievement test that would measure improvement as reflected by class activity. Therefore, a policy recommendation for the EAP program would be to coordinate class levels and curricula with CAEL band levels, allowing learners to prove their capabilities through class performance (e.g., achieving a minimum of B- in class) instead of a single test performance. Learners would thus be able to satisfy language requirements through class performance and would take the CAEL only in case of failure as a secondary avenue to gain acceptance to university. Alternatively, class and test performance could be combined in order to reflect the work and effort that learners invested throughout the term of study.

The findings of the current study also hold implications for future washback research. First, this study represents a preliminary exploration based on a single methodology and a limited number of subjects. It is likely therefore that the factors identified in this study represent a subset of the potential number of factors that actually influence washback. Future research using varied methodologies (including observation, interviews, questionnaires, as well as quantitative methods) and larger and more varied participant samples would be useful in more exhaustively identifying those factors relevant to washback. Second, given the probability that certain factors influence washback more strongly than others, research should also identify the relative importance of respective factors in shaping the overall positive or negative quality of washback. Finally, research

should further investigate the links among the various factors, which may offer insight into actual washback processes. Importantly, an in depth examination of the connections between *contextual* and *test-specific* factors may address Wall's (2000) challenge to develop frameworks for predicting or controlling washback.

## **Bibliography**

- Alderson, J.C., & Clapham, C. (1992). Applied linguistics and language testing: A case study of the ELTS test. *Applied Linguistics*, 13 (2), 149-167.
- Alderson, J.C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14 (2), 115-129.
- Alderson, J.C., & Hamp-Lyons, L. (1996). TOEFL test preparation classes: A study of washback. *Language Testing* (13), 280-297.
- Alderson, J.C., & Banerjee, J. (2001). Impact and washback research in language testing. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, et al. (Eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies*. Cambridge: Cambridge University Press.
- Allwright, D., & Bailey, K.M. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge, England: Cambridge University Press.
- Andrews, S. (1994). The washback effect of examinations: Its impact upon curriculum innovation in English language teaching. *Curriculum Forum* (1), 44-58.
- Andrews, S., & Fullilove, J. (1994). Assessing spoken English in public examinations- why and how? In J. Boyle & P. Favley (Eds.), *English language testing in Hong Kong* (pp. 57-86). Hong Kong: Chinese University Press.
- Andrews, S. (1995). Washback or washout? The relationship between examination reform and curriculum innovation. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bringing about change in language education* (pp. 67-81). Hong Kong: University of Hong Kong.
- Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback - a case study. *System* (30), 207-223.
- Bachman, L. (1991). What does language testing have to offer? *TESOL Quarterly* 25 (4), 671-704.
- Bachman, L., & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. (2001). Designing and developing useful language tests. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, et al. (Eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies*. Cambridge: Cambridge University Press.

- Bailey, K.M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing* (13), 257-278.
- Bailey, K.M. (1999). Washback in language testing. *TOEFL Monograph Series*, 15. Princeton, NJ: Educational Testing Service.
- Blue, G.M., Milton, J., & Saville, J. (Eds.) (2000). *Assessing English for academic purposes*. Bern: Peter Lang AG.
- Bradshaw, J. (1990). Test-takers reactions to a placement test. *Language Testing*, 7 (1), 13-30.
- Brindley, G. (Ed.) (1995). *Language assessment in action*. Sydney: National Centre for English Language Testing and Research.
- Brown, H.D. (1990). M&Ms for language classrooms? Another look at motivation. In Alatis, J. (Ed.) *Georgetown Round Table on Languages and Linguistics 1990*. Washington D.C.: Georgetown University Press.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Brown, J.D. (1997). Do tests washback to the classroom? *TESOLANZ Journal* (5), 63-80.
- Brown, J.D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly* 32 (4), 653-675.
- Brown, J.D., & Hudson, T. (2002). Criterion-referenced language testing. Cambridge: Cambridge University Press.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal* (10), 12-42.
- Burrows, C. (1999). Adopters, adapters, and resisters - did the assessment of the CSWE change teaching in the AMEP? Paper delivered at the Language Testing Research Colloquium 99 Conference. Tsukuba, Japan.
- Chapham, D.W., & Snyder, C.W. (2000). Can high-stakes national testing improve instruction? *International Journal of Educational Development* (20), 457-474.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education* 11 (1), 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 (3), 279-301.

- Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education* 15, 253-271.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, & Y. Watanabe (Eds.), *Washback in language testing: Research contexts and methods* (pp.147-170). Mahwah, NJ: Lawrence Earlbaum Associates.
- Cheng, L., & Curtis A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, & Y. Watanabe (Eds.), *Washback in language testing: Research contexts and methods* (pp.3-17). Mahwah, NJ: Lawrence Earlbaum Associates.
- Cheng, L., & Watanabe, Y. (Eds.)(2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Earlbaum Associates.
- Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.), *Issues in language testing research* (pp.108-129). Rowley, MA: Newbury House.
- Deci E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum Press.
- Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge: Cambridge University Press.
- Eisenhart, M.A., & Howe, K.R. (1992). Validity in educational research. In M.D. Lecompte, W.L. Millroy, & J. Preissle (Eds.), *Handbook of qualitative research in education* (pp. 643-680). San Diego, CA: Academic Press.
- Elson, N. (1992). The failure of tests: Language tests and post-secondary admissions of ESL students. In B. Burnaby & A. Cumming (Eds.), *Socio-political aspects of ESL in Canada* (pp.110-121). Toronto: OISE Press.
- Flowerdew, J. & Peacock, M. (2001). *Research perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press.
- Fox, J. (Ed.) (2000). *CAEL Assessment test score and user's guide*. Language Assessment and Testing Research Unit. Ottawa, ON: Carleton University Press.
- Fox, J. (2004). Test decisions over time: tracking validity. *Language Testing*, 21 (4), 437-465.
- Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher* 18 (9), 27-32.

- Fullilove, J. (1992). The tail that wags. *Institute of Language in Education Journal*, 9, 131-147.
- Gan, Z., Humphreys, G., & Hamp-Lyons, L. (2004). Understanding successful and unsuccessful EFL students in Chinese universities. *The Modern Language Journal*, 88 (2), 229-244.
- Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology*, 13, 266-272.
- Gates, S. (1995). Exploiting washback from standardized tests. In J.D. Brown, & S.O. Yamashita (Eds.), *Language testing in Japan* (pp. 101-106). Tokyo: Japanese Association for Language Teaching.
- Geertz, C. (1973). *The interpretation of culture*. New York: Basic Books.
- Gilette, B. (1994). The role of learner goals in L2 success. In Lantolf, J.P., & Appel, G. (Eds.), *Vygotskian approaches to second language research* (pp.195-213). Norwood, N.J: Ablex.
- Glesne, C. (1999). *Becoming qualitative researchers: an introduction*.(2nd Ed.). Longman.
- Gray, K. (2000). Assessment in EAP: Moving away from traditional models. In Blue, G.M., Milton, J., & J. Saville (Eds.) *Assessing English for academic purposes* (pp. 191-203). Bern: Peter Lang AG.
- Green, D.Z. (1985). Developing measures of communicative proficiency: A test for French immersion students in grades 9 and 10. In P.C. Hauptman, R. Leblanc, & M.B. Wesche (Eds.) *Second language performance testing* (pp.215-227). Ottawa, Canada: University of Ottawa Press.
- Hamp-Lyons, L. (1997). Washback, impact, and validity: Ethical concerns. *Language Testing*, 14, 295-303.
- Hamp-Lyons, L. (1998). Ethical test preparation practice: The case of the TOEFL. *TESOL Quarterly* 32 (2), 329-337.
- Heinrichsen, L.E. (1989). *Diffusion of innovations in English language teaching: The ELEC effort in Japan, 1956-1968*. New York: Greenwood Press.
- Hughes, A. (1988). Introducing a needs-based test of English language proficiency into an English-medium university in Turkey. In A. Hughes (Ed.), *Testing English for university study*. ELT Document # 127. Modern English Publications.

- Hughes, A. (1993). Backwash and TOEFL 2000. Unpublished manuscript, University of Reading.
- Jang, E., & Ryan, K. (2003). Bridging gaps among curriculum, teaching and learning, and assessment. *Journal of Curriculum Studies*, 35 (4), 499-512.
- Jennings, M., Fox, J., Graves, B., & Shohamy, E. (1999). The test-takers' choice: an investigation of the effect of topic on language-test performance. *Language Testing*, 16 (4), 426-456.
- Khaniya, T.R. (1990). Examinations as instruments for educational change: Investigating the washback effect of the Nepalese English exams. Unpublished Ph.D. thesis. University of Edinburgh.
- LoCastro, V. (2001). Individual differences in second language acquisition: attitudes, learner subjectivity, and L2 pragmatic norms. *System*, 29, 69-89.
- Locke, E.A. & Latham, G.P. (1984). *Goal setting*. Englewood Cliffs, NJ: Prentice Hall.
- Lynch, B.K., & Davidson, F. (1994). Criterion-referenced language test development: Linking curricula, teachers, and tests. *TESOL Quarterly* 28 (4), 727-743.
- Madaus, G.F. (1988). The influence of testing on the curriculum. In L.N. Tanner (Ed.), *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the Study of Education* (pp. 83-121). Chicago: University of Chicago Press.
- Masgoret, A.M., & Gardner, R.C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53 (1), 123-163.
- McNamara, T. (1996). *Measuring second language performance*. London: Longman.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18 (2), 5-11.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing* (13), 241-257.
- Morrow, K. (1991). Evaluating communicative tests. In S. Anivan (Ed.), *Current developments in language testing* (pp. 111-118). Singapore: Regional Language Centre.

- Noble , A.J., & Smith, N.L. (1994).*Measurement-driven reform: Research on policy, practice, repercussion* (Tech.Rep. 381). Tempe, AZ.: Arizona State University, Center for the Study of Evaluation.
- Paltridge, B. (1992). EAP Placement Testing: An integrated approach. *English for Specific Purposes*, 11, 243-268.
- Patton, M. (1990). *Qualitative evaluation and research methods*, 2d ed. Newbury Park, CA: Sage Publications.
- Pearson, I. (1988). Tests as a lever for change. In D. Chamberlain & R.J. Baumgardner (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98-107). London: Modern English.
- Pierce, B.N. (1995). The theory of methodology in qualitative research. *TESOL Quarterly*, 29, 569-576.
- Popham, W.J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng, & Y. Watanabe (Eds.), *Washback in language testing: Research contexts and methods* (pp. 171-190). Mahwah, NJ: Lawrence Earlbaum Associates.
- Raimes, A. (1990). The TOEFL Test of Written English: Causes for concern. *TESOL Quarterly*, 24, 427-442.
- Sambell, K., McDowell, L., & Brown, S. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23 (4), pp. 349-371.
- Sharkey, J., & Layzer, C. (2000). Whose definition of success? Identifying factors that affect English language learners' access to academic success and resources. *TESOL Quarterly* 34 (2), 352-368.
- Shohamy, E. (1982). Affective considerations in language testing. *Modern Language Journal*, 66, 13-17.
- Shohamy, E. (1990). Language testing priorities: A different perspective. *Foreign Language Annals* 23 (5), 385-394.
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal* 76 (4), 513-521.

- Shohamy, E. (1993). The use of language tests for power and control. In J. Alatis, (Ed.), *Georgetown University round table on language and linguistics* (pp. 57-72). Washington, DC: Georgetown University Press.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: washback effect over time. *Language Testing*, 13 (3), 298-317.
- Shohamy, E. (1997). Testing methods, testing consequences. *Language Testing*, 14 (3), 340-349.
- Shohamy, E. (1998). Critical language testing and beyond. *Studies in Educational Evaluation* 24 (4). 331-345.
- Shohamy, E. (2000). The relationship between language testing and second language acquisition, revisited. *System*, 28, 541-553.
- Shohamy, E. (2001). *The power of tests. A critical perspective on the use of language tests*. London: Longman.
- Simner, M.L. (1999). Postscript to the Canadian Psychological Association's position statement on the TOEFL. Published by the Canadian Psychological Association, Professional Affairs Committee Working Group on Test Publishing Industry Safeguards.
- Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. *Educational Research Journal*, 28, 521-542.
- Spolsky, B. (1997). The ethics of gatekeeping tests. *Language Testing*, 14 (3), 242-247.
- Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research. Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage.
- Swain, M. (1985). Large-scale communicative testing. In Y.P. Lee, C.Y.Y. Fok, R. Lord, & G. Low (Eds.), *New directions in language testing*. Hong Kong: Pergamon Press.
- Turner, C. (2001). The need for impact studies of L2 performance testing and rating: Identifying areas of potential consequences at all levels of the testing cycle. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, et al. (Eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies* (pp. 138- 149). Cambridge: Cambridge University Press.
- van Lier, L. (1988). *The classroom and the language learner: Ethnography and second-language classroom research*. London: Longman.

- Wall, D., & Alderson, J.C. (1993). Examining washback: the Sri Lankan impact study. *Language Testing* 10 (1), 41-69.
- Wall, D., Clapham C., & Alderson, J.C. (1994). Evaluating a placement test. *Language Testing*, 11 (3). 321-344.
- Wall, D. (1996). Introducing new tests into traditional systems. *Language Testing* 13(3), 334-354.
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: can this be predicted or controlled? *System* (28), 499-509.
- Watanabe, Y. (1996). Does grammar-translation come for entrance examinations? Preliminary findings from classroom-based research. *Language Testing* 13 (3), 318-333.
- Watanabe, Y. (2002). Three theories of motivation and washback to the learner. Presentation delivered at the AILA 2002 Symposium Interdisciplinary research approaches to language testing and learning.
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, & Y. Watanabe (Eds.), *Washback in language testing: Research contexts and methods* (pp. 19-36). Mahwah, NJ: Lawrence Earlbauim Associates.
- Watkins, D., & Biggs, J. (Eds.) (1996). *The Chinese Learner: Cultural, Psychological, and Contextual Influences*. Hong Kong: Comparative Education Research Centre.
- Weir, C.J. (1990). *Communicative language testing*. Hennel Hempstand, UK: Prentice Hall.
- Wesche, M.B. (1987). Second language performance testing: The Ontario Test of ESL as an example. *Language Testing* 4 (1), 28-47.
- Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making and classroom practice*. Cambridge: Cambridge University Press.

## **APPENDIX A**

### **Student Interview 1: Questions**

#### **1 Establishing Background**

- ◆ Language learning experience.
- ◆ Past experience with testing:
  - What kinds of language tests have you taken before? Circumstance.
  - Has testing ever affected you negatively/ positively?

#### **2 Motivation for Learning English**

- ◆ Describe the relevance of English in achieving your goals.
- ◆ Do you feel that language tests motivate you to learn?
- ◆ How would you language learning be different if you did not have to achieve a good language test result?

#### **3 Attitude about ESL**

- ◆ Describe what you do in class here.
  - How is it similar/ different to how you studied English before?
- ◆ Perceptions of IEP program/ teachers.
- ◆ Do you feel that the courses are preparing you well for university? Why or why not?
- ◆ What aspects of the course do you find most/ least interesting?

#### **4 Test Preparation Practices**

- ◆ Can you describe what you know about the CAEL?
- ◆ Are you engaging in test practice during class ? Too much/ too little/ enough.
- ◆ Does your test practice for the CAEL differ from other tests that you have prepared for? Describe.
- ◆ Do you engage in test practice outside of the classroom? Describe.
- ◆ If you are repeating the same test, have you changed your preparation practices? Describe.
- ◆ Have you ever stayed home in order to prepare for an important test?

#### **5 Perception of the High-Stakes Test**

- ◆ Do you feel that you can prepare for the test – perception of control.
- ◆ Do you think that the CAEL is fair? Explain.
- ◆ What is your opinion of the CAEL compared to other high-stakes tests such as the TOEFL? - face validity / reputation.
- ◆ What do your peers say about the CAEL?
- ◆ Do you think the CAEL is a good measure of your ability to use English?
- ◆ With what type of test format do you feel most comfortable (multiple choice, essay, short answer response)?

#### **6 Overall Testing Questions**

- ◆ Do you think we need tests? Why or why not?
- ◆ Do you think such a test is necessary for your success in university?
- ◆ If you could change the way that international students are tested for university entrance - what would you do differently?

Overall do you feel that the CAEL is affecting you positively / negatively?

## **APPENDIX B**

### **Student Interview 2: Questions**

#### **1 Attitude about ESL**

- ◆ After your test taking experience with the CAEL, do you feel that the activities that you were doing in class were similar/ not similar to the CAEL assessment?
- ◆ Was there anything you encountered during the test that you did not feel prepared for? Describe.
- ◆ Looking back at your ESL term, is there anything that you wish would have been different about your classes?

#### **2. Test Preparation Practices**

- ◆ Did you try to prepare for the test in the days leading up to the exam? Describe.
- ◆ If you were preparing for a different test, such as the TOEFL, would you prepare differently? Explain.
- ◆ Did you speak with your friends or parents on the day before the test? Describe.
- ◆ Do you feel that your preparation was effective or would you change something if you had to prepare for the exam again?

#### **3. The Test**

- ◆ Can you describe your feelings on the morning of the test?
- ◆ What strategies do you remember using during the test? Did you find them useful/ not useful?
- ◆ What do you feel was the most difficult aspect of the test?
- ◆ How did you feel about your performance of the test?
- ◆ What comments did you and other students make after the test?

#### **4. Test Results**

- ◆ What was your test result?
  - is this the result that you had expected? Explain.
  - are you satisfied with the placement you received?
- ◆ Do you feel that your result was fair? Explain.
- ◆ Do you feel that your result reflects your true English ability? Explain.
- ◆ What have other students said about their test results and placement?
- ◆ What would you have done if you received a placement of ESL + 1 credit course?

#### **5. Reflecting on the Test**

- ◆ Thinking back to the beginning of the semester, what situation would have been better for your English studies:
  1. to pass the TOEFL and enter university quickly
  2. to study ESL for one term and take the CAEL
- ◆ What was your overall experience preparing for the CAEL?
- ◆ If you had to give a new ESL student advice, what would you tell him/ her?
- ◆ Other comments?

## **APPENDIX C**

### **The Canadian Academic English Language (CAEL) Assessment DEFORESTATION**

#### **Sample Version**

In the CAEL Assessment, you will be asked to do the things you would be expected to do in a university course. You will read two articles which will provide you with background information about a topic and answer questions about them. You will listen to part of a university lecture which will deepen your understanding of the topic. The final task will be to write a response to a question about the topic using information from the readings and the lecture.

#### **Summary**

<b>Time</b>	<b>Task</b>	<b>Points</b>
30 minutes	Reading 1	15 points
30 minutes	Reading 2	27 points
15 minutes	Listening	19 points
45 minutes	Writing	Level placement

**Total time: 2 hours**

All the readings and the lecture in this sample CAEL Assessment are on the topic of deforestation. It is important to understand that at this point in the test you do not need to know anything about the topic of deforestation. All of the information you need to respond to the Writing Task is provided in the lecture and the readings.

You will be able to use information from the lecture and the readings to complete the **Writing Task**.

At the end of the test, you will write an essay agreeing or disagreeing with the following statement:

**Canada's largest natural resource, its woodlands, is in danger of being exhausted. Although part of the fault is due to natural causes, the main culprits are the forest industry itself and the government.**

**Reading 1: "Canada's Wasted Woodlands" (30 minutes/15 points)**

READ THE QUESTIONS BEFORE YOU READ THE ARTICLE. Now, read the article in order to answer the questions.

1. Which paragraph states the main concern of this article? \_\_\_\_\_ (1 point)

2. What is the purpose of the first six paragraphs? (2 points)

---

---

3. According to this article, what are the two main reasons for the current state of the forestry industry? (3 points)

---

---

4. If the forestry industry declines, what will the economic consequences be for Canada? (4 points)

---

---

5. List two solutions to Canada's forestry problems suggested in this article. (2 points)

---

---

6. Match the number of the paragraph in the article to the sentence below which best summarizes the paragraph. The first one has been done for you. (3 points)

- a. 17 Discusses the economic consequences of the industry's failure.
- b. \_\_\_\_\_ Describes the loss of productive forest area through both human and natural causes.
- c. \_\_\_\_\_ Talks about the industry response to the situation.
- d. \_\_\_\_\_ Suggests a possible solution to the political aspect of the forestry problem.

**THIS IS THE END OF READING 1**

## Reading 1

### Canada's Wasted Woodlands

**Because of poor management, its most treasured resource could be gone in a century**

By Douglas Martin

Victoria, British Columbia

1. For one last moment, the towering hemlock stands defiantly, as if unwilling to relinquish the idea of living after nearly three centuries of effort. Then, ever so slowly, it leans over Yellow Creek Mountain, about 80 miles north of here, and falls to earth with a flat thud.
2. It's one of about 130 trees that William Laplante - a French Canadian who has worked as a "faller" of trees for four decades - will saw down this sunny day. For him, the hemlock's death waltz is happily routine.
3. He makes more than \$160 for six and a half hours work each day; the labour is considerably easier than it used to be, and across the Pacific, there are plenty of eager buyers.
4. "The Japanese are crazy for that wood, eh?" Mr. Laplante said, pointing to the tree.
5. On the other side of the same mountain, his son Marcel, a faller for 17 years, is less sanguine.
6. "Pretty soon, all your big wood's going to be gone," he said in an emotionless voice.
7. In accord with many forestry experts, the younger man worries whether enough is being done to conserve Canada's biggest natural resource. But even if the woodlands are preserved, others fear that Canada's forests are proving uneconomical in a changing, more competitive world.
8. Essentially, the problem is that the industry wastes up to half of each tree it cuts; it also fails to replant much of what it harvests. Experts worry that, given present trends, the resource that in the nineteenth-century timber baron thought would last 700 years may disappear in less than a century.
9. At the same time, competitors are cropping up all over the globe - from the forests of Siberia to the jungles of Brazil. And these emerging industry rivals are, in some cases, producing new wood more quickly, more efficiently and more cheaply than their Canadian counterparts.

10. Awareness of the dilemma was first sparked by the migration of the environmental movement north of the 49th parallel; to politicians, expenditures for preservation and replanting, lagging far behind American outlays since Theodore Roosevelt's time, have now become as Canadian as Mounties and moose.

11. But the biggest impetus has come from the timber industry itself, which now finds it is using more costly fuel to carry trees to mills from increasingly remote cutting areas. Many now contend that if replanting had occurred as the trees were harvested, much of the current expense would have been avoided.

12. Moreover, the recession-wracked industry is enduring its fourth bad year in a row, and the effects are readily apparent in the country's unemployment and welfare rolls and in some corporate balance sheets. In fact, more than one-fifth of the timber industry's work force is unemployed; and in British Columbia alone, forest product companies lost \$300 million in 1982.

13. "Our forests are in a shambles, a mess," asserted Jack Walters, director of the University of British Columbia's big research forest: "It's a tragedy of the first dimension." And Michael Innes, manager of forestry for Abitibi-Price Inc., the world's largest producer of newsprint, said: "We're gradually running out of wood, and there's no more wood over the next hill."

14. This assessment is shared by Canada's chief forester, F.L.C. Reed: "Canada will run out of economically priced trees if it doesn't pull up its socks." He acknowledged that, as a result of poor management practices, "The next crop of trees is going to be smaller than the last one, and it's going to take longer to grow."

15. But the problem is not an obvious one - and with good reason. More than 40 percent of Canada is covered by trees - enough to build about 33 million houses. Flying across the country, it is almost impossible to avoid the impression that Canadians, in essence, live in the woods.

16. Indeed, one-tenth of the productive forest area in the world lies in Canada. The trees here on Vancouver Island represent one of the world's greatest exploitable supplies of mature large timber. Elsewhere in Canada, smaller spruce trees make excellent pulp and paper products.

17. The economic consequences are immense. Canada leads the world in forest exports, with the value of shipments in excess of \$12 billion. Some 300,000 Canadians work in the forestry industry, with indirect employment estimated at more than 1 million. Only the former Soviet Union has more trees, but so far has been unable to capitalize on its advantage.

18. But the United States, which now buys large quantities of timber from Canada, faces no threat of a wood shortage; and new supplies are popping up in its Southeast. Meanwhile, a growing number of wood substitutes abound. And Michael Newton, professor of forest ecology at Oregon State University, said that if the latest in forest-management procedures were employed in the United States, "We could export timber like we're exporting wheat."

19. It is Canada that has the chips on the table: forest products represent one-seventh of all manufactured goods, last year accounting for \$23 billion in sales. They make a bigger net contribution to the nation's balance of payments than metals, agriculture, fisheries and the automotive industries combined.

20. In the face of these economic realities, it is obviously a bad time for Canada to be running low on trees, but that is exactly what experts say is happening - albeit slowly. They say that in the past the size of the resource was vastly overestimated, and that Canada may have only half of the productive forest it thought it had.

21. The mounting concern about Canada's timberlands comes against the increasingly clear perception that elsewhere in the world, wood is plentiful. Accordingly, the Forest Industries Advisory Committee, a high-level business and labour panel, recently warned that Canada's forestry industry future hinges on its ability to compete in today's tough marketplace.

22. Even if much of Canada's softwoods are of a higher quality than wood elsewhere, "the world seems to have an infinite capacity for making do, if the price is right," the panel said in a report to the Ottawa government. The panel asked for far more assistance from Canadian taxpayers. Its justification was stark.

23. "If the industry fails, in the future there will be a helluva lot fewer taxpayers," James Buchanan, president of the Canadian Paperworkers Union, said last week.

24. "In Canada, we are no longer a dominant producer of a scarce resource - we're now a large owner of an abundant resource," observed C.C. Knudsen, chairman and chief executive of MacMillan Bloedel Ltd., Canada's largest forest product company and now a unit of Noranda Mines Ltd. "And we are beginning to price ourselves out of the market."

25. Indeed, MacMillan Bloedel Ltd., like many other Canadian companies, has already felt the pinch. Last year the company reported a \$57.3 million dollar loss, compared with a modest profit of \$3.3 million in 1981. But 1983's prospects have brightened a bit, with \$14.1 million in first-half profits, compared with a loss of \$21.5 million in last year's first half.

26. At the same time, Abitibi's fortunes have also declined sharply, as profits slipped to \$80.9 million in 1982, from \$117.6 million in 1981. Although the company reported a \$10.8 million profit in the first six months of 1983, this is well below the \$43.2 million it earned for the first half of 1982.
27. And it seems that the long-term outlook for many forest-product companies - both big and small - remains less than promising. The Science Council of Canada, a quasi-governmental organization that assesses scientific and technological resources, reports that about 6,000 square miles of forest land is destroyed or wasted annually. The council cites insufficient replanting, inability of some lands to regenerate naturally, and the effects of fire, insects, disease and wind.
28. "Natural regeneration is not keeping up with what you need," said Jack Munro, president of the International Woodworkers Union in Western Canada. "We're at a crossroads."
29. Apart from the timberland here in British Columbia - the last stop on the colourful westward trek of the loggers and saloonkeepers and painted ladies who followed them - the virgin forest is clearly gone. What is left is so scattered - or so remote and ecologically fragile - that it is not economical to exploit. New growths, in areas where the first generation of timber was long ago harvested, have fallen short of hopes.
30. Signs of diminishing resources are everywhere. Supplies near mills have been so depleted in some areas that it is necessary to transport logs hundreds of miles. Quebec and Ontario furniture-makers must import maple logs from the United States. And millions of acres of forest land in the Maritime provinces have been devastated by the spruce budworm.
31. The overall Government statistics are equally chilling. Each year Canada loses nearly 40,000 acres of forest land through lack of natural or artificial regeneration, one-fifth of what is harvested each year. And more than 30,000 acres fall prey annually to fires and disease. Still more is lost to parks, farms and subdivisions. The result has left Canada's backlog of wasted forest lands at somewhere between 11.5 million and 70 million acres - with most analysts leaning toward the higher figure.
32. The upshot has been the renewal of a debate as old as the Dominion itself. "We are recklessly destroying the timber of Canada and there is scarcely a possibility of replacing it," declared Sir John A. Macdonald, Canada's first Prime Minister, in 1871.
33. Canadians are also worried about dwindling demand for the wood they have left - although such slackening might conceivably buy some time to replenish the forests. In the 1960s and early 1970s, global wood products orders were growing at nearly 5 percent a year. Analysts now believe the industry will be lucky to see a 3 percent annual growth rate in demand, even with a strong economic recovery.

34. In part, the industry response has been to seek more government support. The advisory committee last week, for example, recommended to the federal government that it assist the timber industry with a mix of tax incentives, trade development and heightened reforestation measures.

35. The Canadian provincial and federal governments, which own more than 90 percent of Canada's forest land and lease logging rights to companies, are addressing the timber resource problem. It is estimated that by 1985, the federal and provincial governments and industry will increase their combined annual forest regeneration spending to \$330 million from \$250 million five years earlier.

36. But much remains to be done. Mr. Newton, for instance, believes that Canada must spend tens of millions of dollars killing the economically worthless stands of hardwood, principally aspen, now blanketing what were formerly rich spruce forests of Canada.

37. Part of the problem, critics say, is that politicians are in the habit of thinking from election to election, rather than in terms of the 60 years or so it may take to grow a mature tree. Another negative effect of the politicization of the forests, said Mr. Walters of the University of British Columbia, is that funds are not channelled to the best growing areas, but rather are apportioned among legislative members with trees in their districts.

38. A solution, Mr. Walters suggests, would be to put Canadian forests in private hands, as roughly three-fourths of the forest in the United States already are. But the Canadian public would probably never relinquish common ownership of its forests, observers say. Most companies believe the new provincial arrangements, which require more of companies but also give them a greater guarantee of tenure, protect long term investments.

39. Growing ecological awareness on the part of the general public has also added momentum. And the continuing spruce budworm epidemic in the East and several severe fires have underlined the limits of the already shrinking resource. Mr. Innes recalled a 300,000-acre fire in 1980 in Abitibi's forests near Thunder Bay, Ontario. He said that the company's board asked him, "If we ever had a second one of those things, what would happen?" His answer: "It would close three mills. Permanently."

40. In forestry, however, the answers are seldom simple. Experts agree that more forest land is better than less. However, proper management could cut in half the time needed to grow a commercial tree.

41. That might be impossible without vast resources of expertise and money. During the five years ending last year, the timber industry had capital outlays of \$15.3 billion, nearly twice the level of the preceding five-year period.

42. Although more money is needed to insure the continued productivity of Canada's forests, the question of how much more, the advisory panel said, may be less important than how new funds are used.

## **Reading 2: "Exhausting the Inexhaustible" (30 minutes/27 points)**

**READ THE QUESTIONS BEFORE YOU READ THE ARTICLE.** Now look for the information in the article.

1. What is the main idea of this article? Put a check mark  beside the correct answer.

(1 point)

- the writings of John A. MacDonald
- the timber trade is Canada's largest industry
- Canada's timber resource is inexhaustible
- Canada is running out of trees

2. What percentage of Canada's land is covered by trees? \_\_\_\_\_ (1 point)

3. Fill in the chart below explaining why this percentage is deceptive. (6 points)

<b>Problem</b>	<b>Details</b>
Quality of trees	• _____ • _____
Availability of suitable trees	• _____ • _____ • _____
Type of mills	• _____ • _____

4. What evidence is there to support the following statement: "Our forests are also shrinking." (1 point)

\_\_\_\_\_

5. What is the purpose of paragraphs 5 and 6 in the article? (1 point)

\_\_\_\_\_

**GO ON TO THE NEXT PAGE**

6. According to the article, in which area could the depletion of Canada's forests have devastating consequences? Put a check mark (✓) beside the correct answer. .

(1 point)

Environment       Employment       Exports

7. Fill in the chart below with information about loss of our forests due to fire and insects. (4 points)

Year	Natural cause	Result
1980		
1981		Eastern Canada
		Quebec

8. According to the article, which one of the following causes the most damage? Put a check mark (✓) beside the correct answer. (1 point)

- Fire  
 Disease  
 Insects  
 Human perceptions

9. List three ways in which the forests are neglected.: (3 points)

- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_

10. According to the article, who is responsible for the failure to manage the forest resource? (1 point)

- \_\_\_\_\_

**GO ON TO THE NEXT PAGE**

11. Which level of government is responsible for each of the following areas of neglect. Use **P** for Provincial and **F** for Federal (Explanation: Federal government is the national government of Canada which passes laws concerning issues that affect all of the provinces. Provincial government is the regional government of each province which passes laws concerning issues affecting only that province. There are ten provinces in Canada). The first one has been done for you.

(4 points)

#### **LEVEL OF GOVERNMENT AREA OF NEGLECT**

- F lack of leadership  
\_\_\_\_\_ has not enforced reforestation  
\_\_\_\_\_ reduced CFS staff  
\_\_\_\_\_ halved available research money  
\_\_\_\_\_ hasn't developed plans for forest industry

12. What is now being done to protect the trees? (3 points)

.

---

.

---

.

---

13. Are the forests still shrinking? Please put a check mark  beside the correct answer.  
(1 point)

\_\_\_\_\_ Yes      \_\_\_\_\_ No

**THIS IS THE END OF READING 2**

## **Reading 2**

### **Exhausting the Inexhaustible**

**A century ago, Canadians thought they could never cut all the trees in the country. Now, as a freelance writer, M.L. Allen reports, our biggest industry is running out of raw material.**

**"The sight of immense masses of timber passing my windows every morning suggests to my mind the absolute necessity there is for looking into the future of this great trade. We are recklessly destroying the timber of Canada and there is scarcely a possibility of replacing it... It occurs to me that the subject should have been looked in the face and some efforts made for the preservation of our timber."**

1. Those words were written by Sir John A. Macdonald in 1871 as he watched giant timber rafts move down the Ottawa River to the sawmills. He was certainly right that something needed to be done. However, it was 50 years before even the first small steps toward reforestation - planting new trees to replace what has been cut - were begun. It was another 50 years before many people started to think there might be a problem. It's only within the last few years that the full reality of the situation has sunk in - Canada is running out of trees.
2. That would seem to be impossible. Trees cover 70% of Canada's land area - 44 million km<sup>2</sup> of forests - an apparently inexhaustible supply. But, those figures are deceptive; the productive forest in Canada is only half as large as it appears to be.
3. A large portion of our trees are not the right kind of quality for harvesting. Of those that are suitable, many are not available; they are protected in parks and reserves, or they are in inaccessible areas such as mountains; or they are so remote as to make harvesting uneconomical. The situation is further complicated by the fact that many mills are built to handle only certain kinds of trees; thus, even if there are still other good trees close by, the mill might not be able to use them, and the mills that could use them might be too far away.
4. Our forests are also shrinking. A Science Council of Canada report states that "one-eighth of the country's productive forest area has deteriorated to the point where huge tracts of land lie devastated, unable to regenerate a saleable crop within the next 60 or 80 years." This amounts to a loss of over 200,000 km<sup>2</sup> each year. In practical terms, this means that in some areas, logs must be hauled hundreds of kilometres to mills; in other areas, even helicopters - at a cost of \$2,500 an hour - are used to get the trees. In some towns - Fort Nelson, B.C.; Dryden, Ontario; Noranda, Quebec; and Chatham, N.B. are examples - there isn't enough

timber to keep the mills operating.

5. The depletion of Canada's forests, however, is far from a merely local problem. The forest industry here is not just big business, it is our biggest business - bigger than oil, bigger than agriculture, bigger than minerals and fisheries. Directly, and indirectly, one in ten Canadian workers is involved in forestry, making it the country's largest employer after government. In 1981, the forest industry produced \$23 billion in goods, and paid more than \$3 billion in taxes. Forest products account for nearly 15% of all goods manufactured in Canada (in New Brunswick it's 30%, and 50% in British Columbia). More than 300 communities across Canada are one-industry towns, solely dependent on forestry, not just for their livelihoods but for their very existence.

6. Internationally, Canada produces 9% of the world's forest products, the third largest producer after the United States and the former Soviet Union. However, our country is the largest exporter of manufactured forest products; 90% of our newsprint, 70% of our lumber, and 60% of our chemical pulp are exported; in all, nearly 30% of the manufactured forest products that enter the world market come from Canada. This export of wood products contributes a net gain of \$12 billion to the Canadian economy, more than the combined exports of farm products, fish, metal and coal.

7. Given the importance of forests to Canada, their depletion could have devastating consequences. It has been estimated that within the next ten years, 250,000 Canadians dependent on forestry could be out of work. At the same time, world demand for wood products is rapidly rising; it's expected to increase by 50% between 1975 and 2000. Thus, not only will existing jobs be lost, but new jobs and new revenue (perhaps as much as \$12 billion a year) will also be lost if Canada is unable to respond to the growing demand.

8. In part, the loss of our forests is due to natural causes. Each year, fire, insects and disease claim about 100 million m of wood. In 1980, for instance, the worst year in five decades for forest fires in Canada, five times more timber burned than was harvested. In 1981, spruce budworm infestation covered 350,000 km<sup>2</sup> of forest in eastern Canada, severely affecting Newfoundland and Cape Breton; in Quebec, losses to the spruce budworm may end up totalling 150 million m of wood.

9. More than fire and insects, though, the depletion of Canada's forest resources is the result of mistaken attitudes and perceptions. For example, in the middle of the 19th century, a government survey estimated that the pine forests of the Ottawa Valley were sufficient to supply its mills for the next 600 years. Seventy-five years later, the forests were gone and most of the mills had closed down. Still, despite this and other examples to the contrary, the myth of the inexhaustibility of Canada's forests persisted for a very long time.

10. If the forest is thought to be inexhaustible - if you only have to go over the next hill to get more trees - it's not necessary to do much to preserve the woodlands. And that's exactly what's been done to this point - not much. The two levels of government receive huge tax revenues from forestry. However, only about 5 cents on the dollar goes back into the forestry management programs that help to ensure the continued health of the industry. In Canada today, there is one forester for every 800 km<sup>2</sup> of forest; the international standard is one for every 120 km<sup>2</sup>, and in the United States there is one forester for every 80 km<sup>2</sup> of forest. Likewise, in Canada in 1978 only 0.6% of forestry sales were spent on research and development, compared to 1.5% of sales in the United States. By far, though, the most damaging aspect of the neglect has been insufficient reforestation. Canada's forest companies cut about 8,000 km<sup>2</sup> a year, and plant or reseed about a quarter of that area. About 2,000 to 3,000 km<sup>2</sup> of the rest will restock reasonably well on its own. But that still leaves 3,000 to 4,000 km<sup>2</sup> that lie idle, or produce unusable scrub. Hence, the disappearance of the forests. The problem is not so much exploitation by the industry, which has been central to Canada's economy for 150 years. The real failure is that of government, which hasn't made sure that the trees are replaced, that our forest supplies - and the revenue that they generate - will continue into the future.

11. The federal government has not provided much leadership, and even cut the Canadian Forestry Service staff and research funds in half during the 1970s. The provincial governments, which control 90% of Canada's forests, have either ignored the issue of reforestation, or failed to come up with sensible plans for the forest industry.

12. Even where intentions are good, execution has fallen short. In 1977, Premier William Davis promised that Ontario would plant at least two trees for every one cut; in fact, in recent years, only 38% of land cut has been replanted, and foresters think that only one third of the trees planted will survive. What should have been a great renewable resource is starting to look like an unrenewable one. As Les Reed, former assistant deputy minister in charge of the Canadian Forestry Service put it, "We have mined our forest when we should have been managing it."

13. Because trees are such a slow-growing crop, usually taking 60 to 100 years to reach harvestable size, our current problems have been a long time developing. "In this kind of situation, you don't fall off a cliff," Les Reed says, "you slide down a slippery slope." But eventually you'll get to the edge, and for the first time in the history of Canadian forestry, that edge is in sight still distant, perhaps, but very definitely there. "As we know it today, the forest industry is dying," John Walters, a University of British Columbia forestry professor says, "Action is urgently required if the forestry industry is to survive."

14. Blunt talk, but maybe - 113 years after Macdonald suggested it - that's what is needed to get the subject

of preserving Canada's forests "looked in the face." People in the forest industry are speaking out about the problems. The federal government is getting more actively involved. Ontario is making agreements with the forest industry about reforestation programs. Research is developing faster-growing, stronger trees, and new ways to reduce waste and utilize trees that until now have not been useable.

15. So, some things are finally starting to be done. Whether it's enough won't be known for a while. Unfortunately, if it turns out that it's not enough, it will already be too late to make amends. Billions of dollars and hundreds of thousands of jobs are at stake. Meanwhile, the chainsaws are working. In the time it took to read this article, about three hectares of productive forest have been lost.

## **Listening Section (15 minutes/19 points)**

During this part of the test you will hear a taped lecture on deforestation. The tape will be played only once. You must answer the questions while you are listening to the tape. The questions follow the sequence of the lecture. Do not stop to re-write answers during the listening. At the end of the lecture, you will have 5 minutes to go over your answers. You have 3 minutes now to read the questions in this section before you listen to the lecture.

**ANSWER THESE QUESTIONS WHILE LISTENING TO THE LECTURE.**

1. What is this lecture about? Check ✓ one. (1 point)

- Logging in the Amazon rainforests
- Pollution of the earth's atmosphere
- Destruction of the world's forests
- The forest industry since 1950

2. According to the speaker, name 3 reasons why "an all-out war is being waged against the forests of the world"? (3 points)

a) \_\_\_\_\_

b) \_\_\_\_\_

c) \_\_\_\_\_

3. How much of the earth's surface was covered by forests in 1950? \_\_\_\_\_ (1 point)

4. How much is covered now? \_\_\_\_\_ (1 point)

**GO ON TO THE NEXT PAGE**

5. Fill in the chart below with the missing information. (4 points)

Location	Degree of Deforestation
Malaysia	-half the forest logged in the past 20 years -40% forest cut in past 10 years
Amazon forest	
Asia and Africa	
	-more than 1/3 deforested

6. Fill in the chart below with information from the lecture. (4 points)

#### RESULTS OF DEFORESTATION

RESULTS OF DEFORESTATION		
LOCAL	GLOBAL	
erosion		increased carbon dioxide
/ \		
flooding in the wet season	a)	b)
\ /		
produces desert		increased temperature
	/ \	
c)	affects climate	d)

GO ON TO THE NEXT PAGE

**7. The Main Idea (5 points)**

In the space below (approximately twenty lines), summarize, in point form if you like, the main points of this lecture

**THIS IS THE END OF THE LISTENING SECTION**

*Transcript of Lecture — Topic: Deforestation*

At the present time, an all-out war is being waged against the forests of the world. People are using them for a variety of purposes. First, there is the problem of firewood. Nearly half the world uses firewood as its main fuel, and it is partly because of this that forests are being cut down. In poorer countries, the forests are also being logged for their valuable timber, such as teak and mahogany, so that this timber can provide these countries with the foreign exchange that is so badly needed. As well, they're being cleared by slash and burn techniques in the hope of producing the agricultural land needed to feed rapidly growing populations.

It's really not so surprising, then, that the forests are losing the battle. In 1950, about one-quarter of the earth's land surface was covered with forests, and today it is less than one-fifth. The greater part of this loss has occurred in the tropical rain forests, which are continuing to be destroyed at the rate of about 112,000 km<sup>2</sup> a year. This is in an area the size of Nova Scotia and New Brunswick combined.

I'm going to ask you to consider these statistics which are from around the world. In the last twenty years, half of Malaysia's forests have been logged; 40% of Thailand's forest has been cut in the last ten years; the Amazon forest, the greatest rainforest in the world, is now disappearing at a rate of even more than 10,000 km<sup>2</sup> a year. Altogether, about half of Asia and Africa and more than a third of Latin America have been deforested.

The results of such massive destruction are starting to alarm experts, because the consequences are dangerous on both local and global levels. On the local level, the loss of a forest frequently results in erosion, flooding during wet seasons, and water shortages in the dry seasons. Instead of creating agricultural land, deforestation often produces a desert, which is, of course, not an area known for the production of crops.

On a global basis, the loss of these forests could be even more catastrophic. Rain forests have been called the lungs of the world. They take in carbon dioxide and release oxygen. An increase of carbon dioxide in the atmosphere causes what's known as the "greenhouse effect"—a gradual rising of global temperatures that could have a disastrous effect on climate and agriculture throughout the world.

And there is another result which is indirect but extremely serious from this deforestation of the world's tropical forests. The tropical forests are the world's gene pool, containing more than 3.5 million species of plants and animals, of which only approximately 15% have been properly catalogued. The loss of the forest could result in the extinction of a million or more species, and most of these would not even have been discovered, studied, or analyzed for their potential benefits.

Another area where we stand to lose is in the area of medicines. About one-quarter of our medicines are derived from the tropical forests, so the scope of the potential loss is enormous. And every minute, another 20 hectares of tropical forest disappears. The situation is truly alarming.

**Writing Section (45 minutes / level placement) (PLANNING PAGE)**

You have 45 minutes to write an essay agreeing or disagreeing with the following statement:

**Canada's largest natural resource, its woodlands, is in danger of being exhausted. Although part of the fault is due to natural causes, the main culprits are the forest industry itself and the government.**

Use the information from the readings and the lecture to help you organize and plan your essay.

Please write at least 1 page.

**DO NOT COPY DIRECTLY FROM THE READING ARTICLES.** If you wish to use one or two sentences, please use quotation marks ("...").

**You are advised to spend 15 minutes planning your essay and 30 minutes writing. (Space, approximately twenty lines, is provided on the actual test.)**

## **Writing Section Level Placement (TEST RESPONSE PAGE)**

You have 45 minutes to write an essay agreeing or disagreeing with the following statement:

**Canada's largest natural resource, its woodlands, is in danger of being exhausted. Although part of the fault is due to natural causes, the main culprits are the forest industry itself and the government.**