

**Colleges Integrating Immigrants to
Employment Project (CIITE)
Phase 2 Final Report**

**Language Proficiency (Assessment)
Project Report
August 31, 2007**

Revised - November 20, 2007

Table of Contents

1. Preface	1
Use of the Canadian Language Benchmarks.....	3
2. Introduction	4
2.1 Overview of CIITE	4
2.2 Overview of Project Activity Areas	4
3. Recommendations	6
4. Program Benchmarking Pilot	8
4.1 Research Design.....	8
4.2 Results of the Program Benchmarking Pilot.....	9
5. CLBPT Pilot	12
5.1 Research Design.....	12
5.2 Results of CLBPT Pilot	13
6. Test Analysis	17
6.1 Research Design	17
6.2 Findings of Test Analysis	18
6.3 Discussion and Analysis of Test Analysis Findings.....	19
6.4 Aligning Tests with the CLB	20
6.5 Use of IELTS in Ontario Colleges	21
7. Integration with Other CIITE Projects	22
7.1 Admissions and Language Proficiency	22
7.2 Advising and Language Proficiency	22
7.3 Employment Preparation and Language Proficiency	22
7.4 Credential Assessment & Advanced Standing and Language Proficiency	23
7.5 Competency Assessment and Language Proficiency	23
7.6 Data Collection in CE & PT Studies and Language Proficiency.....	23
8. Acknowledgements	24

Appendices

Appendix 1. Template for Colleges to Reflect on the Usefulness of the CLBPT at their Institution	25
Appendix 2. The Impact of CAT/2, CAAT-D, Accuplacer and CAEL on L2 Learners in the College System	31
Appendix 3: Relationship between Program Enrolment, Language, and Performance ..	36

1. Preface

Reading This Report

This report is highly technical and focuses on language assessment issues. It describes the research design, methodology, and process of the Colleges Integrating Immigrants to Employment Language Proficiency (Assessment) Project. This preface may assist those reading the report that do not have expertise in linguistics, language training, language acquisition, or test design.

An Overview of Language Assessment Practices at Ontario Colleges

Assessing language skills is a highly complex endeavour since using language involves the continuous application of multiple sets of symbols, e.g. aural language symbols (words, intonation etc.), written language symbols, and non-verbal cues. Assessment, testing, and examination are integral components of all college programs. Faculty and staff are well aware of the complexities involved in assessing individual abilities to perform particular tasks, or to demonstrate knowledge of particular theories, practices, and processes. Language assessment, however, is likely the most complex assessment process at colleges since communication is an imprecise, multi-layered process, and language engages multiple functions and modes.

Ontario colleges work in English and/or French. Language admissions requirements for all full-time programs are set by colleges and ask applicants to demonstrate academic readiness through the presentation of high school marks in English or French. Requirements may differ among colleges, and programs may require different levels of language proficiency.

For individuals educated overseas in non English/French-speaking countries, and who have not completed high school in Canada, colleges require that academic readiness in English or French be demonstrated through the presentation of marks from approved tests. As with high school marks, colleges set admission requirement levels in relation to tests such as the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), etc. It is with respect to individuals who apply to colleges or seek advisement with neither high school marks nor international test scores that language assessment issues occur and barriers are created.

For those who apply without high school marks or English as a Second Language (ESL)/French as a Second Language (FSL) test scores, colleges have their own testing procedures. These generally involve the use of tests that have been developed by the

college or purchased from a third party. As there is no province-wide policy covering this issue, a multitude of tests have been adopted. Most colleges use mature student tests that assess the skills of native speakers such as the Canadian Adult Achievement Tests (CAAT). In instances where colleges have adopted ESL/FSL tests to administer locally, which very few have done, it has been in an ad-hoc manner. Not surprisingly, this lack of consistency and coherence presents a problem for second language immigrants without test scores or high school marks. It is the barriers created for this population, that the Language Proficiency (Assessment) Project set out to address.

Scope of the CIITE Language Project

CIITE Phase 1 identified critical barriers in the assessment of language proficiency in English and/or French at Ontario colleges. Concern with language issues is prevalent across the country; a recent Statistics Canada report stated that immigrants who have been in Canada for four years indicated that the two main difficulties they have faced in Canada were finding an appropriate job and dealing with language barriers.¹

The CIITE Phase 1 report highlighted numerous barriers with respect to issues of language at Ontario colleges. These include the confusion caused by the multitude of language assessment tools used, as well as the perceived inappropriateness of some of these assessment tools for L2² (ESL/FSL) learners. Barriers also included the inability to consistently define the levels of language proficiency required to be successful in each program of study, and the lack of available, appropriate, and effective language training programs corresponding to learner needs. However, to define a manageable scope of activities, it was decided that the Phase 2 language project would focus on assessment. The exclusion of language training reflects the fact that a great deal of work is already taking place in this area.

Three activities were identified in order to encourage consistent, fair, and accurate language assessment practices that utilize a common framework and appropriately assess for the requirements of college programs. First, “language benchmarking”, which had been successfully adopted by colleges in other provinces, was piloted. Language benchmarking is a process whereby the delivery of a post-secondary program is examined to determine the level of language proficiency required to be successful in the program. Second, the Project piloted the Canadian Language Benchmarks Placement Test (CLBPT) to determine whether it could be adopted as a common test for all

¹ Statistics Canada. *Knowledge of Official Languages Among New Immigrants: How Important is it in the Labour Market?* 2005. 89-624-XWE.

² In this document, L2 refers to a person for whom English is not their native language.

colleges. Third, it was determined that a common framework (the Canadian Language Benchmarks) be used to analyze tests, thus employing the same descriptive scale so that a common means of aligning multiple tests could be adopted.

Use of the Canadian Language Benchmarks

Similar to many post-secondary institutions, Ontario colleges do not use a common framework to measure, describe, and assess language issues. In CIITE Phase 1 it was recognized that a common framework could remove some of the barriers created by the lack of a common frame of reference. Given this reality, CIITE investigated the applicability of introducing the Canadian Language Benchmarks as the common framework for colleges. The Canadian Language Benchmarks (CLBs) were chosen because they are being adopted across the country as a national set of standards for the description of language ability. There are two documents which describe the Canadian Language Benchmarks: The *Canadian Language Benchmarks 2000 (CLB 2000)*, and the *Niveaux de compétence linguistique canadiens 2006(NCLC)*. Prior to the NCLC, a French translation of the *CLB 2000, Standards linguistique canadiens (SLC)* was the only available French CLB tool. It is important to note that during the CIITE project the NCLC were still under development. Without a proper French tool, this project examined the use of the CLB in Ontario colleges primarily through the *CLB 2000* and English-language assessment tools.

The Centre for Canadian Language Benchmarks (CCLB)³ describes the *CLB 2000* as “a descriptive scale of communicative proficiency in English as a Second Language, expressed as benchmarks or reference points. They cover four skill areas: reading, writing, speaking and listening, and use real life language tasks to measure language skills.” Benchmarks provide a standard way of describing what a learner can “do” with language.

Project Purpose

Underpinning the Language Proficiency (Assessment) Project was the assumption that the purpose of assessment is to determine the ability of the person tested to function in a program of study and that the language demands of the program of study therefore need to be determined. The dual focuses of the project were thus assessment tests and determining the language demands of particular programs.

³ The CCLB is the national body which sets standards for the Canadian Language Benchmarks. The CCLB supports and promotes the use of these national standards in education, community, and workplace settings. Please visit the CCLB’s website for further details: www.language.ca.

2. Introduction

2.1 Overview of CIITE

Over 80,000 immigrants come to Ontario annually with post-secondary education or other forms of training. These new Canadians often seek bridge training or additional education or training opportunities in order to prepare to enter the workforce. Ontario colleges provide programs and services to many recently arrived internationally trained immigrants (ITIs). Conservative estimates indicate that 15,000 ITIs apply to colleges annually, with approximately 60,000 currently registered as college students. As immigration to Ontario increases these numbers will rise exponentially.

As the numbers of immigrants accessing colleges has grown, the barriers they face on the path to employment through Ontario colleges have become increasingly apparent. The recognition of these barriers, and a commitment by colleges to create effective pathways to employment for internationally trained immigrants, resulted in the establishment of CIITE.

CIITE is a three phase project. Phase 1 (2005) was the consultative and analytical phase. Phase 2, the current phase (2006-2008), engages colleges in the development, design, and testing of new systems to remove or address barriers for immigrants. Phase 3, the final phase (2008-2011), will implement the new systems, tools, and processes that were piloted in Phase 2.

2.2 Overview of Project Activity Areas

The three key activities of the Language Proficiency (Assessment) Project were:

1. **Program Benchmarking:** The anticipated outcome was to “establish college program benchmarking for selected programs and a protocol for implementing benchmarks in conjunction with CCLB.”
2. **CLBPT Pilot:** The *Canadian Language Benchmarks Placement Test* (CLBPT) was piloted to “determine if and how CLB assessment tools can be incorporated into college assessment practices as a reliable, accurate, efficient, and cost-effective tool.”

3. **Test Analysis:** CIITE engaged in a comparative analysis of the selected CLB-aligned language test (CLBPT) with current language assessment tools used for admission to Ontario colleges. This activity was intended to “develop a comparative scoring chart that correlates the CLB assessment tool with other language assessment tools commonly used by colleges.”

This report is divided into these three areas; however, the findings from these activities are interdependent since they are all integral to supporting a new approach to assessment that involves pairing language assessments with program requirements. It should be noted that although these three activities were undertaken in isolation of the other CIITE projects, they are elements in a continuum dealing with barriers for ITIs and therefore will integrate with other CIITE projects in the *Phase 2 Final Report* (see Section 7, page 22, for further details).

3. Recommendations

The following recommendations emanate from the Language Proficiency (Assessment) Project findings and reflect the interdependency and overlap of the three project activity areas. All the recommendations focus on increasing the capacity of Ontario colleges to ensure the success of ITIs in college programs and the workplace.

1. The Canadian Language Benchmarks (CLB) be adopted in Ontario colleges as a language proficiency framework.

The results of all three activity areas support this recommendation. Colleges would benefit from using a common language proficiency framework and the CLB provides colleges with consistent national standards by which to refer to language proficiency. If adopted, the benchmarking of college programs, combined with CLB-aligned assessment tools, would result in a fairer assessment and admissions process for ITIs and all learners whose first language is not English (hereafter referred to as L2).

2. Colleges adopt program benchmarking as an essential instrument for both assessment and program design/delivery. Program benchmarking should include:

- 1. Full-time and part-time programs that have a high concentration of L2s and/or ITIs.**
- 2. Analyzing representative semesters at the beginning, middle, and end of a program, including any work placements or practicums.**

3. The college system should not adopt the CLBPT as the post-secondary assessment tool for entry into college programs.

By piloting the CLBPT to a cohort of ITIs in six colleges, it was determined that the CLBPT is not suitable for implementation across the Ontario college system as the post-secondary language assessment tool. Please see Section 4 of this report for the rationale for this recommendation. It is important to note that it is the decision of each college whether it is appropriate to adopt the CLBPT as a tool for placement into ESL courses at their institution.

With respect to using a common assessment tool, it was not within the scope of the CIITE project to make recommendations that span beyond Ontario colleges. However, the Language Proficiency Project Team believes there would be significant benefits in supporting a national initiative to develop a CLB-based test that could be used in post-secondary educational institutions across Canada.

4. Colleges should use a suite of language tests to assess language sufficiency in order to enter a college program or a college language course. Commonly used language assessments should be aligned with the CLB in order for colleges to base decisions on a student's ability to function in the program (see Section 6.4 for an explanation of aligning tests with the CLB).

Colleges have existing language tests that can be used for specific purposes. To best assess a student's language proficiency for success in college programs, colleges should use a suite of language tests. This is because tests are designed for specific purposes, and could be used by colleges for three different functions: determining program eligibility, placement into language courses, or identifying support and service needs. With a suite of language tests, colleges will be able to test appropriately both for the function, and based on the requirements of particular programs.

It is important to note that CIITE is currently benchmarking the International English Language Testing System (IELTS). Since Citizenship and Immigration Canada (CIC) accepts only IELTS results when assessing immigration applications, this test will be analyzed and considered as one of the tests to be included in the suite of language tests used by Ontario colleges.

4. Program Benchmarking Pilot

4.1 Research Design

The Program Benchmarking pilot used the CLB as a tool to analyze the language demands within selected college programs. The pilot activities were conducted by seven colleges: Algonquin, Centennial, George Brown, La Cité, Mohawk, Seneca, and Sheridan, in collaboration with the Centre for Canadian Language Benchmarks (CCLB).

Thirty-one post-secondary programs and eight language courses were benchmarked between January and June 2006. The seven colleges provided a sound test sample to pilot a methodology within different organizational structures and varying populations of ITIs.

The post-secondary programs benchmarked are listed below. The numbers beside the programs indicate the number of colleges that benchmarked the program, if more than one.

Automotive Technician	Business Admin
Biological Sciences	Common Business
Community Worker	Culinary Management
Early Childhood Education	Educational Assistant
Environmental Protection Technician	Graphics Technology
Financial Services Underwriting	Human Resources
Hotel and Restaurant Management (2)	Manufacturing Management
Motive Power Technician	Opticianry
Office Administration (2)	Accounting (3)
Practical Nursing (3)	Respiratory Therapy
Social Service Worker – Gerontology	Transportation Engineering
Quality Assurance	Pre-Health Sciences
Wireless Technology	

Based on a methodology developed by Red River College (Winnipeg, Manitoba), the pilot involved research teams at each pilot college using a detailed approach to data collection which included:

- selecting a cross-section of academic programs
- conducting extensive classroom and lab observation
- analyzing a multitude of course materials (including textbooks, handouts, websites, and other supplemental resources)
- interviewing instructors, students, and key program administrators.

Red River College provided a comprehensive set of data collection tools, which were supplemented by tools provided by the CIITE Project Office. The Project Office also provided college teams with a final reporting template. Using these tools, each benchmarking team had the discretion to organize and conduct their research within the given timeframe (January to June 2006), according to the individual context of their institution.

After the data collection activities, the *Canadian Language Benchmarks 2000* was used as a framework for assigning appropriate benchmarks in each of the four language skill areas: reading, writing, listening, and speaking.

In order to identify common themes and differences amongst the participating colleges, two roundtables were held: the first during the research phase (May 2006) and the second after the process was completed (November 2006). The latter roundtable was useful in determining which lessons learned were common across pilot colleges and which were unique to a particular college. Both roundtables provided a forum for researchers from different institutions to work in small groups and share their knowledge and experience. As each college has unique needs, the objective of these sessions was not to build consensus amongst Ontario colleges. Rather, the roundtables allowed researchers to gain insight into the benchmarking process and develop strategies to deal with some of the more subjective and ambiguous aspects of this research.

4.2 Results of the Program Benchmarking Pilot

4.2.1 Development of Protocol Guide

One of the project deliverables was to establish a protocol for colleges to implement benchmarks in conjunction with CCLB. To achieve consistency in program benchmarking, the project has developed the *Language Benchmarking at Ontario Colleges: A Guide to Program Benchmarking Using the Canadian Language Benchmarks*. For a copy of this guide please refer to the CIITE website: www.collegeconnect.on.ca/ciite

4.2.2 Scope Considerations for Benchmarking

The Program Benchmarking Pilot focused only on the first and second semester of a program. The results of this project indicate that benchmarking a college program should be expanded for the following reasons:

- The search for ITIs to participate in the CLBPT Pilot activities identified that many ITIs enter programs in a semester other than the first. This means that for program benchmarking to be useful, colleges should examine language demands in representative semesters at the beginning, middle, and end of a program.
- Processes for the recognition of previous education (i.e. Advanced Standing, Transfer Credit) can be supported by analyzing any semesters of entry for students who receive advanced standing.
- The Program Benchmarking Pilot found that enhanced skill levels in all four areas (reading, writing, listening, and speaking) were often required for practicums or work placements. This suggests that benchmarking the language used in workplace settings could enhance the benefits of benchmarking for student success and program renewal.

4.2.3 Program Benchmarking Benefits

Several benefits of program benchmarking for Ontario colleges have been identified.

1. Benchmarking provides a common structure for colleges to understand and describe the language used in college programs. It enhances the recognition of language as a key element in education and provides a framework for colleges to consider how language usage is employed to convey course content. The activity of benchmarking includes looking at how students are expected to access the information taught within vocational programming (e.g. reading, lectures) and demonstrate comprehension (e.g. through essays, oral presentations).
2. Benchmarking describes the types of language used in a program; for instance, formal, informal, academic, and/or colloquial. It also describes the extent to which the four language skills (listening, speaking, reading, and writing) are used in a given program. Understanding the nature of language used, as well as the prevalence of each of the four language skills can result in more informed language testing for placement and admissions.

3. In addition to the formal objectives of the benchmarking process, program benchmarking promotes communication between faculty, departments, and L2 students. It may also inform program administrators and staff of useful teaching and learning strategies.

4.2.4 Program Benchmarking Uses

This pilot determined that program benchmarking supports student success by providing colleges with a comprehensive picture of the language skills required throughout a program. Specifically, program benchmarking in Ontario colleges can serve a number of functions within the college system, including:

1. **Establish/validate admission requirements:** Provided that colleges have appropriate CLB-aligned assessment tools, program benchmarks can be used to establish or verify admissions requirements for programs. In applying benchmarks for this purpose, it is important not to use benchmarks as a barrier to entry since language is not the only determinant of academic success. Successful admission also depends upon background in the subject matter, level of support available, and effectiveness of study habits.
2. **Determine placement in language courses:** If a college has preparatory and remedial language training programs that are aligned with the CLB, it could use benchmarks to determine placement in these types of language courses.
3. **Design remedial and preparatory language training:** The language used in post-secondary college programs is specific to the industry sector (field of study) and is also academic. The benchmarking process can assist colleges in improving the alignment of pre-post-secondary language training with the actual classroom demands of post-secondary college programs.
4. **Inform program renewal:** By benchmarking more than the first semester of a program, colleges would receive a report on the progression of language demands within that program. Semester specific benchmarks can provide useful observations about the demands in the four language skill areas. These, in turn, could be used to inform program renewal processes by providing information on language demands across semesters.

5. CLBPT Pilot

5.1 Research Design

5.1.1 Description of Pilot Activities

One of the CIITE Language Proficiency Project deliverables was to “determine how existing CLB assessment tools will be incorporated into college assessment practices as reliable, accurate, efficient and cost-effective tools, where necessary.” This deliverable was addressed through the pilot of the *Canadian Language Benchmarks Placement Test* (CLBPT) conducted at six Ontario colleges: Algonquin, Centennial, George Brown, Mohawk, Seneca, and Sheridan.

Two pilots were set up at each college in the Fall and Winter semesters of the 2006/07 academic year. In total, 260 ITIs completed all components of the CLBPT pilots. In order to administer the test, 58 faculty from the participating colleges were trained and certified as CLBPT assessors by the CCLB. Participants from the relevant benchmarked semester from the Program Benchmarking Pilot were then selected according to criteria set out by the Project Team. Each participant completed the CLBPT and a demographic questionnaire. A final questionnaire was also administered following the receipt of grades in which participants were asked to evaluate the sufficiency of their language skills for the completion of their semester.

The CLBPT assessments were scheduled in college locations as available and included classrooms, boardrooms, meeting rooms, private suites in offices, assessment centres, program interview rooms, and computer laboratories. The academic performance of each participant was tracked according to the participant's final term Grade Point Average (GPA). Instructor feedback was also gathered for every course a participant failed.

Finally, each college was provided with a template (*Appendix 1*) that enabled them to examine the usefulness of the CLBPT at their institution. Following a predetermined process, each college Project Coordinator, in collaboration with researchers participating in the Test Analysis project, filled in the template. They analyzed six key qualities of the CLBPT in order to decide on its usefulness. These qualities were Reliability, Construct Validity, Authenticity, Interactiveness, Impact, and Practicality.

All data was entered into an online database designed and developed by the Language Proficiency Project Team and external Information Technology consultant. Each college compiled college-specific findings into a comprehensive *Pilot Report* which was then submitted to the Project Office. An external Data Analyst was also hired to analyze the data from all six pilot sites as a single data set.

The Project Office then compiled these results for a one-day meeting (June 2007) in order for colleges to meet and share their findings. The Project Team used the results of the CLBPT, as well as the experience of the Test Analysis project, to draw conclusions about the usefulness of the CLBPT for Ontario colleges.

5.1.2 Limitations of the Data Collection

Identifying potential candidates for the two CLBPT Pilots proved challenging for the following three reasons.

1. Current data collected in college systems (e.g. OCAS/SIS) does not deliver definitive results that identify ITIs. Electronic searches based on citizenship status are not able to identify ITIs who are citizens. Not all individuals submit foreign credentials with an application and foreign credentials that are submitted are paper-based and require manual searches.
2. The continuing education database does not identify ITI status.
3. It was discovered that many ITIs do not enter the program in first semester, thus it was necessary to broaden the search to include students outside the first semester.

5.2 Results of CLBPT Pilot

Based on the pilot findings, it was determined that CIITE cannot recommend CLBPT as an appropriate assessment tool to be adopted across the college system. The following sections outline the rationale for this determination.

5.2.1 Cost Effectiveness and Efficiency of the CLBPT

The CLBPT has an interview component that requires a one-on-one assessment conducted by a trained assessor in a small room conducive to this type of test. During the pilot, assessors were able to complete approximately six to eight interviews daily with each interview taking approximately two hours.

Activity	Time Required
Test taker completing reading subtest	30 minutes
Test taker completing writing subtest	30 minutes
Test taker completing speaking/listening subtest	20-25 minutes
Assessment of the speaking and writing subtests	30 minutes
Total Time for Assessment	1 Hr, 55 minutes

Due to the specialized room requirements and the time required to complete the assessment, it was often difficult to schedule tests at times convenient for students. While a full cost analysis is still being compiled, it is the finding that the CLBPT is not cost effective and efficient with respect to the use of college space and human resources, especially for colleges with significant numbers of ITIs.

5.2.2 Reliability of CLBPT Results

With reference to language assessment, reliability is commonly defined as the consistency of measurement across testing situations, sets of tasks, forms or versions, and raters (individuals who mark the tests). Observations during the CLBPT Pilot indicated that reliability was compromised due to several factors, including:

- The variance possible during the speaking and listening component
- Topic variance across the forms of the CLBPT⁴
- Possible inconsistency between raters

Therefore, the CLBPT is considered unreliable for decisions such as entry into college programs. However, if used as a placement⁵ tool into ESL courses the CLBPT could be used in colleges.

The CLBPT Pilot found that individual performance could be affected by the test taker's background knowledge on the given topic. Also, the limited variations of the test would make it difficult to ensure test security if implemented across the Ontario college system.

⁴ The CLBPT has three possible forms, with limited test topics and varying levels of language use.

⁵ Placement is referred to as a "low-stakes" decision; that is, the decision has lower consequences on a test-takers life. Tests for low-stakes decisions are usually less expensive to design and administer since they do not require the same amount of resources to comply with reliability and validity standards. The CLBPT was designed as a placement tool. Admissions is referred to as a "high-stakes" decision since the consequences for a test-taker of being inaccurately assessed are higher.

5.2.3 Accuracy of CLBPT Results

For a CLB-based test to accurately assess a student's preparedness for a college program, the tasks and language used in the test would need to be similar to the tasks and language used in the college program. The chart below provides a comparison between the information used for CLBPT and what is normally used in a college program.

Area	CLBPT Tasks and Language Used	College Tasks and Language Used (also applicable to benchmarking programs)
Reading	Text length 1 to 4 paragraphs, focused on topics of general interest	Text length 20 to 30 pages, focused on technical and academic content
Writing	Paragraph about personal experience, brief essay containing personal opinion	Reports, summaries, case studies, business correspondence
Listening	30 seconds to 1 minute, speaker describing general topic	Lectures 2-3 hours with PowerPoint describing technical, academic content
Speaking	Answer personal questions, describe an everyday activity from photo story, state a point of view and provide support	Content specific activities such as small group discussion, presentations, group work

Therefore, the accuracy of the CLBPT results, when compared to a CLB-based program benchmark, may not provide a comprehensive indication of the participants ability to engage in all of the language tasks required in the semester.

5.2.4 CLBPT and Different Language Types

Ideally, the language assessment tool(s) adopted by the college system need to test different types of language whether formal, informal, academic, or colloquial. The CLBPT tests informal, everyday language but does not test the use of formal, academic types of language that are often required in post-secondary programs.

5.2.5 CLBPT as a Placement Test

The CLBPT is a language proficiency test designed for placement in language courses. Several colleges that participated in the CIITE Language Proficiency (Assessment) Project currently do not deliver language at varying levels (e.g. College English/French, remedial English/French, and developmental English/French). As a result, the CLBPT

as a placement test is not appropriate for colleges that do not offer different types of language training.

5.2.6 Use of CLBPT with Postsecondary Benchmarked Programs

Of the 31 postsecondary programs benchmarked during the program benchmarking pilot, 20 programs had benchmarks above Level 8. However, CLBPT only tests the four skills (writing, reading, listening, and speaking) up to benchmark Level 8. Therefore, the CLBPT should not be used to test the skills required for success in post-secondary college programs since a large percentage of the programs may be above Level 8.

5.2.7 Relationship between Program Enrolment, Language, and Performance

The data gathered in the CLBPT pilot provides a strong indication that in the case of ITIs (given their professional and educational background) language should be assessed in association with other elements. There is strong indication, for example, that enrollment in a program related to previous education and/or profession impacts performance. In general, the data suggests that strong language skills, in association with enrolment in a related field (see Q1/Q2 graph in *Appendix 3*), constitute the best case scenario for ITIs' success in a post-secondary institution.

The impact of being in a related program was observed across the participants and most of the language skills (see graph of Q6/Q7 and Q8/Q9 in *Appendix 3*). For students tested at a BM lower than the program BM, the data indicates less failure and more success among those enrolled in a program related to their previous education/profession, as compared to students enrolled in a dissimilar program. Listening and speaking results are not clear and further study is therefore necessary.

6. Test Analysis

6.1 Research Design

The purpose of the test analysis activity was to develop a comparative scoring chart which aligns a CLB assessment tool with selected language tests currently used in the Ontario college system. The CLB assessment tool chosen was the Canadian Language Benchmark Placement Test (CLBPT) and the language tests selected for analysis were:

1. The Canadian Achievement Test 2nd Edition (CAT/2 – Level 19)
2. The Canadian Adult Achievement Test level D (CAAT-D)
3. Accuplacer
4. Canadian Academic English Language test (CAEL)⁶

The tests analyzed comprise several subtests. For the purposes of the project, only subtests currently being used in Ontario colleges were analyzed. Thirteen researchers from six different colleges were divided into two groups, coordinated by a Research Coordinator from the CIITE Project Office. The assignment of activities was as follows:

Test	Group 1: Mohawk, Seneca, and Sheridan	Group 2: Algonquin, Centennial, and George Brown
CAT/2:		Test Specifications Reading Vocabulary, Reading Comprehension, and Language Expression
CAAT-D	Test Specifications Reading Comprehension, Reading Vocabulary, English, and Spelling Subtests	
ACCUPLACER	Test Specifications Reading Comprehension (score point 51) subtest Sentence Skills (score point 53) subtest	Test Specifications Reading Comprehension (score point 80) subtest Sentence Skills (score point 86) subtest
CAEL	Test Specifications Listening and Writing subtests	Test Specifications Reading and Speaking subtests

In October 2006, the researchers participated in a one-day training session developed and delivered by the CIITE Project Office. The methodology for the test analysis was developed by the CIITE Project Office. Using an 'expert-judgment method',⁷ a panel of researchers from six different colleges determined test specifications, analyzed test

⁶ Although CAEL is not amongst the most used tests by colleges, it was analyzed here because the colleges involved expressed the desire to better understand it.

⁷ Experts in the area work individually and blindly, compare results, and discuss issues to reach consensus.

items in relation to the CLB, and built consensus for the final results. Researchers also observed the main similarities and differences between these tests and the CLBPT. Once inconsistencies were resolved, a compiled version of the findings was prepared by the research coordinator and sent to the researchers.

In a two-day, face-to-face meeting (April 2007), researchers reflected on the impact of the analyzed tests on L2 college applicants/students and college programs. In the same meeting discussed above, researchers reviewed all the materials developed during the test analysis activity and in small groups discussed how comparable the selected tests were with the CLBPT.

6.2 Findings of Test Analysis

The original deliverable of the Test Analysis activity was to develop a comparative chart that aligns CLBPT benchmarks and scores of selected tests. The assumption was that these tests were comparable with the CLBPT. However, after analyzing and benchmarking the fundamentals of the tests, the researchers concluded that the differences between the CLBPT and the other language tests make them incommensurable.

When researchers compared the CLBPT with the CAAT-D, CAT/2, and ACCUPLACER, they found that the subtests require the test taker to be able to recognize the correct form of discrete items in the areas of mechanics, grammar, and sentence structure in a multiple choice format. In contrast, grammar, sentence structure, and mechanics are not dealt with in this discrete manner in the *CLB 2000*. The CLBPT, which is based on the *CLB 2000*, is designed for L2 learners, covers the four skill areas, and is task-based and communicative.

When researchers compared the CLBPT with CAEL, they found that while both tests are task-based, communicative, and designed for L2 learners, the different types of language (everyday communicative language versus language for educational purposes), construct (low to upper middle range of proficiency versus middle to upper range of proficiency), and targeted audience (L2 immigrants versus L2 university applicants) make them incommensurable. Aligning CLBPT benchmarks and CAEL band scores would be misleading.

6.3 Discussion and Analysis of Test Analysis Findings

At the final meeting, the team reflected on the impact of the tests on L2 learners and the college system. This was not part of the initial test analysis activity. The inclusion of this task stemmed from the recognition that the theoretical knowledge gained during the project directly related to situations in colleges and across the college system. As college faculty, researchers were in a unique position to engage in this type of analysis since they hold expertise both in test analysis and the college system. Many of the test analysis researchers had also participated in the program benchmarking pilot and therefore had recently engaged in an analysis of the language used in college programs.

6.3.1 Informing on Language used in a College Program

For this activity, researchers first reflected upon the extent to which the experience of taking the CLBPT might inform test takers of the language used in college programs in terms of language for communication, areas of language knowledge, and metacognitive strategies. Researchers determined that, depending on the program:

- The experience might suggest that the language demands are lower than they actually are.
- The experience might suggest that explicit knowledge of language mechanics is not required.
- The experience might misleadingly inform the test taker that the metacognitive skills required for success are not as complex as they actually are.
- The Listening/Speaking component might be somewhat informative of the conversation skills necessary to manage an interview.
- The Listening component might not be informative of the listening format and content in the academic context.
- The Reading component might be somewhat informative, but the content is unrelated to college work.
- Two of the Writing tasks in the CLBPT are unrelated to college work. The last task might give a minimal indication of written exercises in colleges. There is some comparability with essay type questions.

6.3.2 Consequences for Ontario College System

Finally, researchers examined the potential consequences for the Ontario college system of using this test with L2 students.

Positive Consequences	Negative Consequences
Communicative language is tested in the four skill areas.	Since the CLBPT's highest BM is 8, students at BM above 8 (BM 9 -12), are not identified.
	Since the CLBPT's highest BM possible is 8, it cannot be used for programs whose language is more demanding.
	Since CLBPT tasks focus on everyday communicative tasks, the test does not properly reflect the language and tasks used for college programs.

For the impact of CAT/2, CAAT-D, Accuplacer, and CAEL on L2 learners and the college system please see *Appendix 2*.

6.4 Aligning Tests with the CLB

It is possible to analyze tests using the CLB document. Benchmarking tests means determining at which benchmark test takers should be in order to obtain specific test scores. It is important to note that when different tests are benchmarked, such as the CLBPT and TOEFL, it could be determined that specific scores of each test operate at the same benchmark, but actually test different language abilities and tasks. This is due to the different purposes of language tests and the nature of the CLB, as explained below.

The benchmarks in the CLB reflect successive levels of achievement on the continuum of English language acquisition. Benchmarks build upon each other and provide examples of language used for communication in the community, in studies, and in the workplace. That is, each benchmark covers a broad range of communicative tasks. Therefore, elements of any of the CLB benchmarks could be found in most language tests. For example, the CLBPT tests an individual's ability to communicate in the community, and language tasks related to everyday life can be found on the CLB continuum. On the other hand, the TOEFL tests language for communication in an educational environment, and many of these types of communicative tasks can also be found on the CLB continuum. Thus, it is possible for tests to operate at the same benchmark level, but test different language abilities. This is why a suite of language tests that can be used for specific purposes is valuable for Ontario colleges.

6.5 Use of IELTS in Ontario Colleges

All colleges require evidence of English proficiency from applicants for whom English is a second language. While most accept TOEFL scores, some Ontario colleges have established IELTS scores acceptable for college admissions. In order to remove barriers for ITIs, colleges should work to accept test scores that individuals may already have, rather than require that individuals pay for additional testing. IELTS is the only language test used by CIC when assessing immigration applications therefore, many ITIs would have IELTS test scores available. Thus, colleges should consider the inclusion of IELTS within the suite of language tests used for assessment.

7. Integration with Other CIITE Projects

Although the seven CIITE projects are being conducted separately, each project focuses on a point in an ITI's journey through the college system to employment. As such, each project will impact the others. This section summarizes the integration issues and opportunities identified to date.

7.1 Admissions and Language Proficiency

CLB-aligned language tests, combined with program benchmarks, can help inform admissions decisions. Registrar Services, Assessment Centres, and benchmark researchers will need to work together to set up practices for using program benchmarks and test scores that ensure these scores do not become a barrier to entry. For example, the Language Training Centre at Red River College has the expertise and capacity to interview students whose CLB scores fall short in any one skill area to determine if any additional testing is warranted and to make recommendations about program readiness or referral to adult EAL (English as a Second Language) classes. In larger institutions this type of relationship would not be feasible. One possible solution would be to set up a range of benchmarks for Admissions criteria.

7.2 Advising and Language Proficiency

Program benchmarks can assist ITI Advisors in providing guidance on a client's readiness for college programs. Information that aligns existing tools with the CLB will allow Advisors to inform clients on the usefulness of common test scores they may already have.

7.3 Employment Preparation and Language Proficiency

Both employers and ITIs consistently rank communication skills as one of the top barriers to employability.⁸ The Language Proficiency (Assessment) Project found that benchmarking the language used in the workplace is an important part of setting students up for success both in the program, and consequently, their preparation for employment. Types of language assessment and their use in Ontario colleges will impact the advice, referrals, and training recommended in the Employment Preparation Project.

⁸ Immigrant job seeker data from *Statistics Canada Longitudinal Survey of Immigrants to Canada*, Catalogue no. 89-611-XIE; Managers data from Canadian Labour and Business Centre's Viewpoints 2002 Survey.

7.4 Credential Assessment & Advanced Standing and Language Proficiency

A decision to grant credit for previous learning relies upon the assurance that an applicant has met the language proficiency requirements of the institution. The Credential Assessment & Advanced Standing Project has created a *Standardized Credential Assessment (SCA) Report* that includes a module for language proficiency. Currently, the OCAS application form only includes a space for TOEFL scores. This limits an applicant's ability to demonstrate their language proficiency and for institutions to include language proficiency scores in one form.

Program benchmarking can also aid the assessment of credentials for advanced standing by providing benchmarks throughout a program. It is recognized that students will be required to have varying levels of language proficiency depending on the semester, and the technical vocabulary of the program. Program benchmarks are one way of providing staff that assess for advanced standing with an indication of a student's ability to succeed in the relevant semester. CLB-aligned tests would allow colleges to provide credit for previous learning with the assurance of both knowledge of program content as well as language proficiency.

7.5 Competency Assessment and Language Proficiency

The CIITE Competency Assessment Project is currently developing a model for competency assessment. The project has noted that language assessment is often a necessary component of competency assessment since this type of assessment requires the ability of an applicant to communicate their own knowledge in ways other than a formal credential. This may include an objective interview process, and/or demonstration of skills. This demonstration may include the ability of the applicant to describe the tools of the profession and their proper use.

7.6 Data Collection in CE & PT Studies and Language Proficiency

Only one of the programs benchmarked in the Program Benchmarking Pilot was in Continuing Education. The CLBPT Pilot found a significant ratio of ITIs in this program. Thus, in order to serve ITIs most effectively, Continuing Education and Part-Time studies with high numbers of ITIs in the program should be benchmarked. In order to support these program benchmarking decisions, further data collection in CE/PT studies is warranted.

8. Acknowledgements

The activities described in this report were made possible thanks to the enthusiasm and commitment of various people and institutions.

The Language Proficiency (Assessment) Project Team would like to recognize the contributions of the Benchmarking Researchers, the Centre for Canadian Language Benchmarks (CCLB), Red River College, the students, instructors, and college departments who participated in the research for the Program Benchmarking activity and the CLBPT Pilots.

For their assistance with the CLBPT Pilot data collection and analysis, CIITE would like to thank James Brouwer (Herculean Educational Solutions), and Duncan MacDuff (Niagara College).

For their cooperation in the Test Analysis activity, CIITE would like to thank the developers of ACCUPLACER (The College Board), CAEL (Carlton University), CAT (Canadian Test Centre), CAAT (The Psychological Corporation), and CLBPT (Centre for the Canadian Language Benchmarks) for providing material for the analysis and answering any questions from the researchers. We are equally grateful to the researchers who undertook the challenge of analyzing these tests, in many cases juggling research and classroom duties.

Finally, CIITE would like to extend sincere gratitude to the Language Proficiency (Assessment) Project Team for their contributions and dedication: Fiona Bain-Greenwood (Seneca College), Rebecca Carnevale (CIITE), Janet Gambrell (Sheridan College), Maria Glass (CIITE), Anna Gris (Mohawk College), Diane Hall (Sheridan College), Angela Jacewicz (Seneca College), Leslie James (CIITE), Bernice Klassen (Algonquin College), Diane Lussier-Bouchard (La Cité Collegiale), Pat Meek (CCLB), Lyne Mychaud (La Cité Collegiale), Joanna Pesaro (Centennial College), William Radford (CIITE), Denise Richard (George Brown College), Leah Teklemariam (CIITE), and Adrianna Wilson (George Brown College).

Appendix 1. Template for Colleges to Reflect on the Usefulness of the CLBPT at their Institution

Reflecting on the usefulness⁹ of the CLBPT

According to Bachman & Palmer (1996, 18), test usefulness can be described as a function of six interrelated qualities -- Reliability, Construct Validity, Authenticity, Interactiveness, Impact, and Practicality. The questions below are thus divided into these six areas.

Step 1 – Answering the questions

With the assistance of a Test Analysis Researcher, answer the questions below on the usefulness of the CLBPT within your college.

To help you go through this activity, we have included examples of answers (attachments 'Example'). Please note that since these questions have been adapted to fit our context (Ontario colleges), the questions and answers in the examples will be slightly different. In addition, we have included definitions of key words and phrases (page 7). If you have any questions, do not hesitate to contact Maria Glass (mglass@georgebrown.ca).

Step 2 - Reflecting on your answers

Once you have answered the questions, go over your results and decide whether the CLBPT is useful to your college.

Bachman & Palmer warn us that evaluating test usefulness is a subjective activity that involves a great amount of value judgment. There is nothing we can do about this. What we can and should do is to avoid two extremes which, also according to Bachman & Palmer (1996, 134), can compromise the process. They are:

1. “the unrealistic expectation that the highest level possible for every quality must or can be achieved, and
2. the indefensible view that one or more of these qualities can be given priority to the virtual exclusion of the others.”

Thus, when reflecting on the CLBPT usefulness, you should attempt to seek a balance amongst the six qualities.

1. Reliability

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
1. To what extent do characteristics of the <i>test setting</i> vary from one administration of the test to another?		
2. To what extent do		

⁹ Adapted and expanded from the *Checklist for Evaluating Usefulness* provided in Bachman and Palmer's book titled *Language testing in Practice*, 1996, Oxford University Press.

characteristics of the <i>test rubric</i> vary in an unintended way from one part of the test to another, or on different forms of the test?		
3. To what extent do characteristics of the <i>input</i> vary in an unintended way from one part of the test to another, from one task to another, or on different forms of the test?		
4. To what extent do characteristics of the <i>expected response</i> vary in an unintended way from one part of the test to another, or on different forms of the test?		
5. To what extent do characteristics of the <i>relationship between input and response</i> vary in an unintended way from one part of the test to another, or on different forms of the test?		

2. Construct Validity

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
Clarity and appropriateness of the construct definition, and the appropriateness of the task characteristics with respect to the construct definition.		
6. Is the language ability construct for this test clearly and unambiguously defined?		
7. Is the language ability construct for the test relevant to the purpose of the test?		
8. Is the language ability construct for the test relevant to the college programs?		
9. To what extent do the test tasks reflect the construct definition?		
10. To what extent do the scoring procedures reflect the construct definition?		
11. Do the scores obtained from the test help you to make the desired interpretations about the test takers' language ability?		
Possible sources of bias in the task characteristics		
12. What characteristics of the test <i>setting</i> are likely to cause different test takers to perform		

differently?		
13. What characteristics of the test <i>rubric</i> are likely to cause different test takers to perform differently?		
14. What characteristics of the test <i>input</i> are likely to cause different test takers to perform differently?		
15. What characteristics of the test <i>expected response</i> are likely to cause different test takers to perform differently?		
16. What characteristics of the test <i>relationship between input and response</i> are likely to cause different test takers to perform differently?		

3. Authenticity

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
17. To what extent do the characteristics of the test tasks correspond to the language used for the purpose of communication in colleges?		
18. To what extent do the characteristics of the test tasks correspond to the language used for the purpose of teaching and learning (both in content and English courses) in colleges?		

4. Interactiveness

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
Involvement of the test takers' topic knowledge		
19. To what extent do the tasks presuppose a specific area or level of topical knowledge, and to what extent can we expect the test takers to have this area or level of topical knowledge?		
Suitability of test tasks to the personal characteristics of the test takers		
20. To what extent are the personal characteristics of the test takers specified in the design statement?		
21. To what extent are the characteristics of the test tasks suitable for the test takers as they are specified in the design		

statement?		
22. To what extent are the characteristics of the test tasks suitable for college test takers' characteristics?		
Involvement of the test takers' language knowledge		
23. Does the processing required in the test tasks involve a very narrow or a wide range of areas of language knowledge?		
24. Does the processing required in the test tasks involve areas of language knowledge that are relevant to college programs?		
Involvement of functional knowledge in the test tasks		
25. What language functions, other than the simple demonstration of language knowledge, are involved in the input and formulating a response?		
26. To what extent are these language functions relevant to college programs?		
Involvement of the test takers' metacognitive strategies		
27. To what extent are the test tasks interdependent?		
28. How much opportunity for strategy involvement is provided?		
29. To what extent are the strategies required relevant to college programs?		
Involvement of the test takers' affective schemata in responding to the test tasks		
30. Are the test tasks likely to evoke an affective response that would make it relatively easy or difficult for the test takers to perform at their best?		

5. Impact

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
Impact on individuals		
▪ <i>Impact on test takers</i>		
31. To what extent might the experience of taking this test inform test takers of the language used in the college		

programs (in terms of language for communication, areas of language knowledge, and metacognitive strategies)?		
32. To what extent might the experience of taking this test inform test takers of the skill areas they need to improve?		
33. How relevant, complete, and meaningful is the feedback that is provided to test takers?		
▪ <i>Impact on instructors</i>		
34. How consistent is the purpose of this test with the general educational goals of instructors?		
35. How consistent are the areas of language ability measured in this test with those included in college program teaching materials?		
Impact on your college		
36. To what extent can the scores obtained from this test help Admissions to make interpretations about L2 test takers' language ability?		
37. Are the interpretations of the test scores consistent with, or relevant to, the Ontario College Program Standards? ¹⁰		
38. To what extent do the language elements tested coincide or conflict with those deemed necessary for L2 students to function in your college programs?		
39. To what extent do the language elements tested coincide or conflict with those deemed necessary for L2 students to function in the workplace after graduation?		
40. What are the potential consequences, both positive and negative, for your college, of using this test with L2 students?		

6. Practicality

Questions for logical evaluation of usefulness	Extent to which quality is satisfied	Explanation of how quality is satisfied
41. What type and relative amount of resources are		

¹⁰ For information on the Ontario College Program Standards, please visit <http://www.edu.gov.on.ca/eng/general/college/progstan/intro.html#overview>

required for the administration of this test?		
42. Is the amount of resources required realistic for your college?		

Key words/phrases

(Vary in) an unintended way: Any variations across tasks or forms not specified as part of the construct. For example, different forms of input, different topical contents, different sets of instructions, etc.

Input: “The information contained in a given test task, to which the test taker is expected to respond” (Bachman, 1990, 125).

Expected response: What is attempted to elicit through instructions and input.

Response: The test taker’s actual response.

Relationship between input and response: Can be classified into reciprocal, nonreciprocal, and adaptive (see Bachman, 1990, 148).

Language ability: Includes language knowledge and strategic competence (set of metacognitive strategies).

Functional knowledge: Part of pragmatic knowledge. “How utterances or sentences and texts are related to the communicative goals of language users” (see attachment ‘Functions’ for more details).

Construct: “The specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from this task” (Bachman & Palmer, 1996, 21).

Construct Validity: “The extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s), we want to measure” (Bachman & Palmer, 1996, 21).

Design statement: A document containing the test specifications.

Reference

Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.

Bachman, L. & Palmer, A. S. (1996). *Language Testing in Practice. Designing and Developing Useful Language Tests*. Oxford University Press.

Appendix 2. The Impact of CAT/2, CAAT-D, Accuplacer and CAEL on L2 Learners in the College System

Reflecting on Impact

CAEL

1. *To what extent does the experience of taking this test inform test takers of the language used in the college programs in terms of language for communication, areas of language knowledge, and metacognitive strategies?*

- The degree to which the CAEL Informs test takers of the language used in college program would depend on the program.
- Depending on the program, the experience might suggest that the language demands are higher than they actually are. Generally, the extended type of writing required in the CAEL is not a large component of first year college programs. The college test experience would usually not require the test taker to integrate as much information from as many sources under the time constraint that CAEL presents. At college, students demonstrate knowledge with short answers, multiple choice, online (with PowerPoint notes as guidance from teacher), reports, and in practical hands-on situations e.g. labs. CAEL focuses on academic English skills whereas many college programs combine academic and workplace English/skills.
- Depending on the program, the experience might appropriately inform the metacognitive skills required for success.
- It might inform test takers of test taking strategies in an academic context.

2. *To what extent does the experience of taking this test inform test takers of the skill areas they need to improve?*

CAEL is integrated and provides the opportunity for the test taker to judge his/her ability to listen, read, gather information, and put ideas together – a positive experience of language in use. The speaking subtest allows the test taker to evaluate his/her skill in performing academic tasks: performing for a particular audience, integrating listening/speaking and reading, and contributing to an academic discussion.

3. *How relevant, meaningful, and complete is the feedback for the test taker?*

The feedback provided to the test taker (in each of the four skills) somewhat informs him/her about strengths and weaknesses. Band scores consist of a holistic, three-sentence description.

4. *What are the potential consequences for the Ontario college system of using this test with L2 students?*

Positive	Negative
<ul style="list-style-type: none"> ▪ Might help to select better qualified students. ▪ Communicative language is tested in the four skill areas. 	<ul style="list-style-type: none"> ▪ Might be too high for entry, and therefore might create barriers. ▪ Might be higher than needed for success in colleges. (Possible solution: cut-off scores should take into consideration the different language demands in different programs.) ▪ Since the themes are often based in engineering/computer science, it may prove to be unfair to students entering other areas.

CAT/2

1. *To what extent does the experience of taking this test inform test takers of the language used in the college programs in terms of language for communication, areas of language knowledge, and metacognitive strategies?*

- The experience of taking the CAT/2 provides test takers with an incomplete view of language used in colleges.
- The experience might provide information about some micro-skills involved in processing a reading text, but the content is not representative of the reading contexts in colleges. The test taker does not have to organize information or construct responses, which are tasks s/he will be required to perform in college programs.
- The experience might provide the test taker with some information about academic vocabulary; however, it would not be sufficient for L2s.
- The experience might provide the test taker with some information about discrete knowledge of grammar; however, it would not be sufficient for L2s.
- Since the test provides the individual with no information about listening, speaking, or writing, the experience might suggest that reading is the most important skill, that listening and speaking skills are not necessary, and that no writing is required.
- The experience does not inform the test taker about how language is used for communication in colleges.

2. *To what extent does the experience of taking this test inform test takers of the skill areas they need to improve?*

- The experience of taking the CAT/2 test provides test takers with very little information about the skills they need to improve for success in college.

- The test's only modality is reading and it is performed using multiple choice. There is no real indication of necessary improvements.
- The test provides no information about listening, speaking, or writing.

3. *How relevant, meaningful, complete, and is the feedback for the test taker?*

The CAT/2 feedback is not relevant or meaningful and there is no relation to academic tasks typical of college work. It is very narrow in scope.

4. *What are the potential consequences for the Ontario college system of using this test with L2 students?*

Positive	Negative
<ul style="list-style-type: none"> ▪ It is practical and cost-effective. 	<ul style="list-style-type: none"> ▪ Since the test focuses only on academic achievement, it will not inform the college about the communicative proficiency of L2 students. ▪ It is a high-stakes test, and therefore might be a barrier. ▪ It could be unfair to L2 students because it is culturally-bound and biased.

CAAT-D

1. *To what extent does the experience of taking this test inform test takers of the language used in the college programs in terms of language for communication, areas of language knowledge, and metacognitive strategies?*

- The experience of taking the CAAT-D provides test takers with an incomplete view of language used in colleges.
- The experience might provide information about micro-skills involved in processing a reading text, but the content is not representative of the reading contexts in college. The test taker does not have to organize information or construct responses, which are tasks s/he will be required to perform in college programs.
- The experience might provide the test taker with some information about academic vocabulary; however, it would not be sufficient for L2s.
- The experience might provide the test taker with some information about their knowledge of discrete grammar ; however, it would not be sufficient for L2s.
- Since the test provides the individual with no information about listening, speaking, or writing, the experience might suggest that reading is the most important skill, that listening and speaking skills are not necessary, and that no writing is required.
- The experience does not inform the test taker about how language is used for communication in colleges.

2. *To what extent does the experience of taking this test inform test takers of the skill areas they need to improve?*
 - The experience of taking the CAAT-D test would provide test takers with very little information about the skills they need to improve for success in college.
 - The test's only modality is reading and it is performed using multiple choice. There is no real indication of necessary improvements.
 - The test provides no information about listening, speaking and writing.
3. *How relevant, meaningful, complete, and is the feedback for the test taker?*

The CAAT-D feedback is not relevant or meaningful and there is no relation to academic tasks typical of college work. It is very narrow in scope.

4. *What are the potential consequences for the Ontario college system of using this test with L2 students?*

Positive	Negative
<ul style="list-style-type: none"> ▪ It is practical and cost-effective. 	<ul style="list-style-type: none"> ▪ Since the test focuses only on academic achievement, it will not inform the college about the communicative proficiency of L2 students. ▪ It is a high-stakes test and might act as barrier. ▪ It could be unfair to L2 students because it is culturally-bound and biased.

ACCUPLACER

1. *To what extent does the experience of taking this test inform test takers of the language used in the college programs in terms of language for communication, areas of language knowledge, and metacognitive strategies?*
 - The experience of taking the ACCUPLACER provides test takers with an incomplete view of language used in the college.
 - The experience might provide information about some micro-skills involved in processing a reading text, but the content is not representative of the reading contexts in college. The test taker does not have to organize information or construct responses, which are tasks s/he will be required to perform in college programs.
 - The experience might provide the test taker with some knowledge of academic vocabulary; however, it would not be sufficient for L2s

- Since the test provides the individual with no information about listening, speaking, or writing, the experience might suggest that reading is the most important skill, that listening and speaking skills are not necessary, and that no writing is required.
- The experience does not inform the test taker about how language is used for communication in colleges.

2. *To what extent does the experience of taking this test inform test takers of the skill areas they need to improve?*

- The experience of taking the ACCUPLACER test provides test takers with very little information about the skills they need to improve. The test's only modality is reading and it is performed using multiple choice. There is no real indication of necessary improvements.
- The Sentence Skills subtests might inform the test taker about his/her knowledge of discrete points of grammar and structure, but not about his/her ability to use English for writing.
- The subtests analyzed provide no information about listening, speaking, or writing.

3. *How relevant, meaningful, complete, and is the feedback for the test taker?*

The feedback to the test taker is limited. The test taker can receive a report with three types of scores (total right score, range, and percentile rank). However, the descriptions for skills/abilities are not comprehensive.

4. *What are the potential consequences for the Ontario college system of using this test with L2 students?*

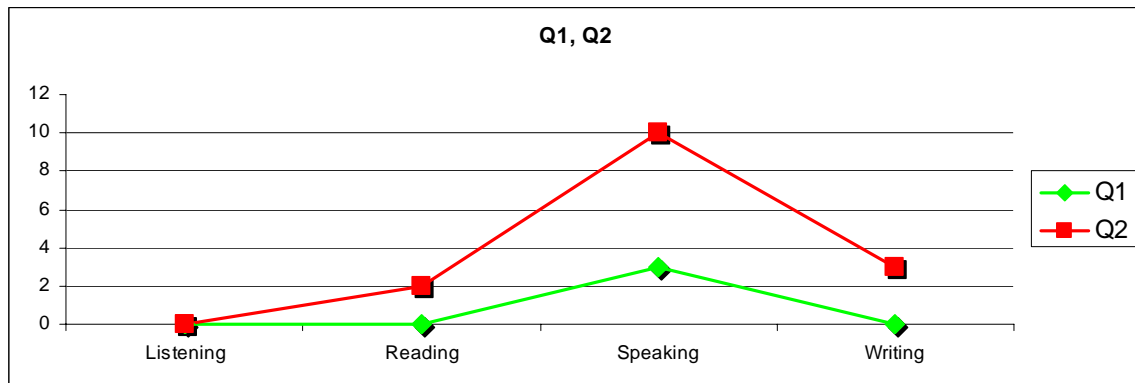
Positive	Negative
<ul style="list-style-type: none"> • If used in conjunction with other tests, it might provide a good snapshot of the students' receptive skills in grammar for placement into differentiated English courses, as well as for language advisement and counseling. • It is practical and cost-effective. • It provides immediate results. 	<ul style="list-style-type: none"> • Since the test focuses only on academic achievement, it will not inform the college about the communicative proficiency of L2 students. • Test takers must be computer literate. • If Sentence Skills is used alone, students may be misplaced in English courses.

Appendix 3: Relationship between Program Enrolment, Language, and Performance

High BM / Related or Different field / Failure

Q1. How many students were tested at the **same or higher** BM as compared to the program benchmark, per skill, and **failed** two or more courses AND had post-secondary in a **related** field?

Q2. How many students were tested at the **same or higher** BM as compared to the program benchmark, per skill, and **failed** two or more courses AND had post-secondary in a **different** field from previous study?



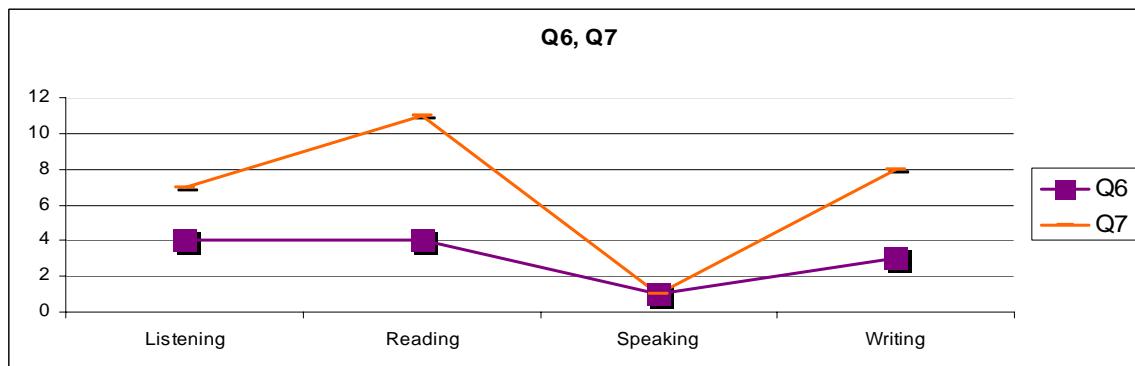
Out of 15 students who failed two or more courses

Observation: There was less failure among students enrolled in a related program.

Possible interpretation: Language, as well as enrolment in a related program, seems to have impact on success.

Q6. How many students were tested at a **lower** BM as compared to the program benchmark, per skill, and **failed** two or more courses AND had post-secondary in a **related** field?

Q7. How many students were tested at a **lower** BM as compared to the program benchmark, per skill, and **failed** two or more courses AND had post-secondary in a **different** field?



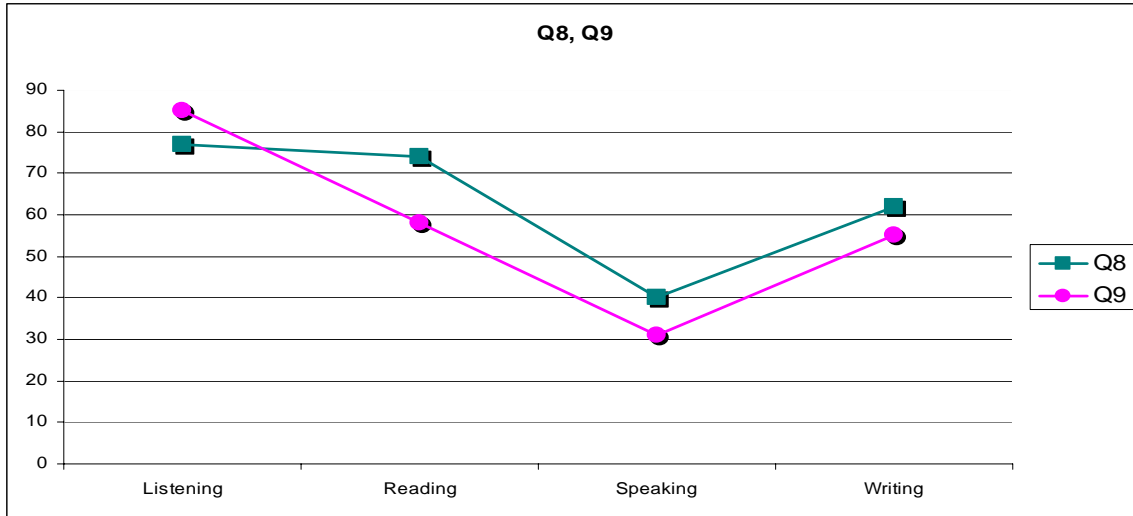
Out of 15 students who failed two or more courses

Observation: There was less failure among students enrolled in a related program despite a low BM (except for speaking).

Possible interpretation: **Enrolment in a related program** seems to have more impact on success than language (except for speaking).

Q8. How many students were tested at a **lower** BM compared to the program benchmark, per skill, and **passed** all courses AND had post-secondary in a related field?

Q9. How many students were tested at a **lower** BM compared to the program benchmark, per skill, and **passed** all courses AND had post-secondary in a **different** field?



Out of 225 students who passed all courses

Observation: There was more success among students enrolled in a related program despite low BM (except for listening).

Possible interpretation: Enrolment in a related program seems to have more impact on success than language (except for listening).