

# Development of an Occupation Specific Language Assessment for Health Science Professions

*Based on the Canadian Language Benchmarks*

## An Enhanced Language Training Initiative Project



## Final Report

This ELT project was a partnership between the Access & Options for International Health Professionals and LCRT Consulting.

Report Prepared by



©2005 LCRT Consulting and Access & Options for International Health Professionals  
For additional information on this project visit [www.lcrtconsulting.com](http://www.lcrtconsulting.com)

# Contents

Contents .....	1
Preface.....	2
Acknowledgements .....	3
Project Overview.....	4
Project Methodology.....	7
Project Outcomes .....	11
Benchmarking of Project Participants (Phase I) .....	12
Participant and Program Data Collection (Phase II).....	13
Benchmarking Programs and Placements (Phase III).....	15
Development of Assessment Tool (Phase IV) .....	19
Evaluation of Assessment Tool (Phase V).....	23
Production and Reporting (Phase VI) .....	28
Project Learnings & Recommendations .....	29

## Preface

This report summarizes a research and development project conducted through the Enhanced Language Training Initiative (ELT), a national initiative to improve immigrant access to the labour market which allocates funding for the development and delivery of labour market language training, including profession-specific language assessment and training.

The Michener Institute for Applied Health Sciences, Access & Options for International Health Professionals (Access & Options), submitted a proposal for the development of an occupation-specific language assessment tool for health science professionals in partnership with LCRT Consulting, a group of language training and testing specialists.

The project received approval in January 2005. The ELT funds support and enhance the Access & Options bridging project for international health professionals already underway at the Michener Institute by adding the much needed development of a customized language assessment.

Core partners, Access & Options and LCRT Consulting, worked over a period of 15 months in consultation with a group of content experts, members of the professional community, project participants, and the Michener Institute staff and management.

Through this project we have succeeded in conducting an analysis of the language needs of internationally trained medical health professionals within both academic and practical settings, developing a language assessment tool that targets the communicative competencies for professional practice in these health care occupations, and pilot testing the newly developed tool. The work was based on the Canadian Language Benchmarks.

Results indicate that the language assessment meets the needs of the Access & Options project to place applicants within their program. There is interest on the part of the Michener Institute in exploring the use of the tool for placement of international students into the program, as well as from regulators affiliated with the program who have identified the need for a communicative-based assessment of language proficiency.

Although we have accomplished our goals for this project, we have identified the potential application of the tool in a number of other contexts including other medical technology programs, programs across Canada, and offshore medical technology programs. In order for this to be possible, it will be necessary to develop parallel versions of the assessment as well as assessment materials for additional medical health technology occupations.

We look forward to building upon our successes to date.

*LCRT Consulting and Access & Options*  
April 2005

# Acknowledgements

## **A&O**

Cecil Canteenwalla  
Director  
Access & Options for International Health  
Professionals

Brendan Corr, DRS(R), A(Ont.)  
Program Manager, Radiological Technology  
Access & Options for International Health  
Professionals

Joyce Clitheroe, MLT, BHA  
Program Manager, Medical Laboratory Sciences  
Access & Options for International Health  
Professionals

Kathleen Olden-Powell, BHA, RRT, RRCT, CAE  
Program Manager, Respiratory Therapy  
Access & Options for International Health  
Professionals

Tim Chipman  
Production Coordinator, Audio Visual Services  
Michener Institute for Applied Health Sciences

## **LCRT Consulting**

Peggie Shek  
Financial Administrator and Test Invigilator

Bruce Russell  
Researcher & Test Developer

Andrea Strachan  
Researcher & Test Developer

## **Content Experts**

Brian Robert Gordon, MRT(R)  
University Health Network

Joanne Salines, MLT  
The Scarborough Hospital  
General Division

Kevin Taylor, RRT  
Professional Practice Leader  
Respiratory Therapy  
St. Michael's Hospital

Michelle Murray, MRT(R)  
University Health Network

Patrick Nellis RRT  
Staff Respiratory Therapist  
University Health Network

Petra Sheldrake, MLT  
Senior Technologist in Clinical  
Chemistry  
Scarborough Hospital, General Division.

Valerie Anzil, MRT(R)  
Clinical Instructor  
Joint Department of Medical Imaging  
University Health Network

## Project Overview

*The Development of an Occupation Specific Language Assessment for Health Science Professions Based on the Canadian Language Benchmarks* is a project of the Michener Institute for Applied Health Sciences – Access & Options for International Health Professionals in partnership with LCRT Consulting. The partnership brings together a training institution that offers a bridging program for internationally trained health science professionals and adult second language teaching and testing specialists. The project's objective is to develop a language assessment tool that is standardized and which allows candidates to demonstrate their language competence in communicative tasks specific to their professions.

This project is part of the federal government's Enhanced Language Training Initiative (ELT), which recognizes the need to enhance existing language training supports for immigrants to improve their access to the labour market.

The rationale for this project is two-fold. First, research conducted by government and non-governmental organizations has identified the lack of occupation-specific language testing and training as a barrier to the successful integration of immigrants to Canada. Secondly, there are current and predicted labour market shortages in Ontario in the medical health professions, including Medical Laboratory Technology, Diagnostic Cytology, Magnetic Resonance Imaging (MRI) Technology, Radiological Technology and Respiratory Therapy.

Demonstration of language proficiency is a key step in the registration process for internationally trained health professionals, and often a barrier. Regulators have traditionally accepted results from standardized tests (often based on University entrance requirements), which are inappropriate for experienced professionals seeking re-entry into an occupation for which they have training and experience because the language requirements of practice differ greatly from those in a first year university program.

Through this project we researched the language demands for the health professions of the Access and Options program and developed a language assessment tool that targets the actual communication skills required for success in the professional registration process. As such, this project enhances the upgrading courses and placements provided through the Access & Options Project.

The Access & Options project started as a Bridging Project funded by the Ontario Ministry of Training Colleges and Universities, which was initiated by The Michener Institute in 2002 to address these skill shortages and labour market needs in Medical Laboratory Technology, Diagnostic Cytology, MRI Technology, Radiological Technology and Respiratory Therapy. The goal of the Bridging Project is to assist internationally educated health professionals become qualified to work in Canada without duplicating their existing learning.

The goal of this Enhance Language Training project is to support candidate success in the Access & Options project by offering accurate and specific assessment of communication skills before a candidate enters into the program.

The key objective of this language assessment is the demonstration of language skills within a profession-specific context.

Partners in The Michener Institute's Access & Options Bridging Program (A&O) have observed that communication breakdown in work placements and difficulty with competency-based examinations point to language difficulties that arise despite success in currently used language proficiency tests<sup>1</sup>. The reasons for this are varied and to date not documented formally.

This new assessment tool provides an alternate language testing process that is now ready to be studied through a longitudinal analysis to determine whether results in a communicative-based occupation-specific language tests are more successful. The occupation-specific language assessment tool developed through this project can verify whether candidates have the language skills required for successful completion of the A&O project and also acts as a diagnostic tool to support candidate counseling, referrals, and the development of targeted individualized learning plans. The assessment results can also be used as a needs assessment to inform the development of language and communication curricula.

As the Access & Options project moves towards a sustainability phase, the focus has been on implementing program components that have longevity and which have an impact on the success of applicants. One such component is a language assessment tool to help determine an applicant's readiness to succeed in academic and technical upgrading, as well as in the clinical practice. This is essential from the perspective of the regulators, the Michener Institute for Applied Health Sciences and the Access & Options for International Health professionals and the program participants.

## **Project Objectives and Outcomes**

This project focuses on the development of a communicative sector-specific language assessment tool, based on Canadian Language Benchmarks 7-10, for internationally trained immigrants trying to access professions in health sciences. There are three professions involved in the Access & Options project: Medical Laboratory Sciences (Medical Laboratory Technology, Diagnostic Cytology), Radiological Technologies (Radiological Technology {X-Ray}, Magnetic Resonance Imaging {MRI}), and Respiratory Therapy.

A research component of the project helped to determine the benchmarks of the target language use in these health science professions. This research will benefit others doing work with internationally trained professionals in this sector across Canada.

The development component of this project has resulted in a language assessment tool that has been pilot tested and is ready for use within the A&O program.

## **The Target Group**

The target group for this project, and the resulting assessment tool is the client group of the Access & Options program. Access & Options clients have satisfied the preliminary professional requirements for application to one of these regulated health professions and are aiming to complete the licensure process. This process requires clinical practice and professional practice exams.

---

<sup>1</sup> Requirements for the occupations listed in this report include: Test of English as a Foreign Language (TOEFL) with a minimum Band score of 550; CanTEST with a minimum Band score level four (4); Test of English as a Foreign Language (TOEFL) with a score of at least 500 for the paper based test or 173 for the computer-based test; Test of Spoken English (TSE) with a score of at least 40.

The Access & Options client group includes both men and women, typically from mid-twenties to early forties. They come from a variety of national and linguistic backgrounds from Europe, Asia and Africa but at present mainly from the Middle East and Southeast Asia.

These include internationally trained health professionals who have experience in the following areas, or in areas that have a transferable skills set:

- Medical Laboratory Sciences (Medical Laboratory Technology, Diagnostic Cytology)
- Radiological Technologies (Radiological Technology {X-Ray}, Magnetic Resonance Imaging {MRI})
- Respiratory Therapy

# Project Methodology

## Theoretical Framework

The central language framework for this test development project was the *Canadian Language Benchmarks 2000: English as a Second Language for Adults*<sup>2</sup> (CLB). The CLB is a descriptive scale of language proficiency that describes English as a second language ability from beginner (CLB 1) to near native (CLB 12). There are benchmarks for four skill areas (listening, speaking, reading, and writing) and a detailed description of the performance conditions and competency outcomes for each.

This comprehensive benchmark system is based on current research in language acquisition, curriculum development, language testing, and communicative competence in second language learners. The theoretical foundation of the CLB is described in the CCLB document *Canadian Language Benchmarks 2000: Theoretical Framework*.<sup>3</sup>, a document that provided the fundamentals for the rating systems set up for this language assessment. The team also referred to the *Canadian Language Benchmarks 2000: Additional Sample Task Ideas*<sup>4</sup> when determining the test tasks.

In addition to the above noted documents from the Centre for Canadian Language Benchmarks, the team referred to Bachman and Palmer's *Language Testing in Practice: Designing and Developing Useful Language Tests*<sup>5</sup>, which describes language test development procedures, and Dan Douglas' *Assessing Language for Specific Purposes*<sup>6</sup>, which discusses how to integrate occupation-specific tasks in test development.

The team also tapped into member's experience and knowledge in the field of English as a second language for adults and occupation-specific language training<sup>7</sup>.

## Benchmarking

After studying the principles in the above-noted references, and conducting a brief literature review, the test development team planned the work based on the recommended procedures for test development. This included a detailed analysis of the target language situations and target language use (the types of communication used in these professions in a clinical setting, and, because this is also an academic program, also in an academic setting). The team analysed these communicative competencies in terms of the benchmark descriptors and determined ideal benchmarks for each stream of the program.

<sup>2</sup> Pawlikowska-Smith, Grazyna (2000) *Canadian Language Benchmarks 2000: English as a second Language for Adults* Ottawa: Center for Canadian Language Benchmarks

<sup>3</sup> Pawlikowska-Smith, Grazyna (2002) *Canadian Language Benchmarks 2000: Theoretical Framework* Ottawa: Center for Canadian Language Benchmarks

<sup>4</sup> Pawlikowska-Smith, Grazyna (2002) *Canadian Language Benchmarks 2000: Additional Sample task Ideas* Ottawa: Center for Canadian Language Benchmarks

<sup>5</sup> Bachman, Lyle and Palmer, Adrian (1996) *Language Testing in Practice: Designing and Developing Useful Language Tests* Oxford University Press

<sup>6</sup> Douglas, Dan (2000) *Assessing Language for Specific Purposes* Cambridge University Press

<sup>7</sup> For biographies and list of projects visit [www.lcrtconsluting.com](http://www.lcrtconsluting.com)

The team concurrently used the Canadian Language Benchmarks Placement Test (CLBPT), a CLB-based language assessment tool for placement in ESL programs ranging from CLB 1-8, to benchmark participant language levels. This done, it was possible to analyse the minimum competency requirements for the program and for clinical placements.

A number of issues arising from the use of the CLBPT are discussed in the section “Benchmarking Participants”. The CLBPT tool limited the assessment results to CLB 8, and tested general communicative competencies as opposed to those relating to health care (i.e. the language needed for patient care).

## **Test Development**

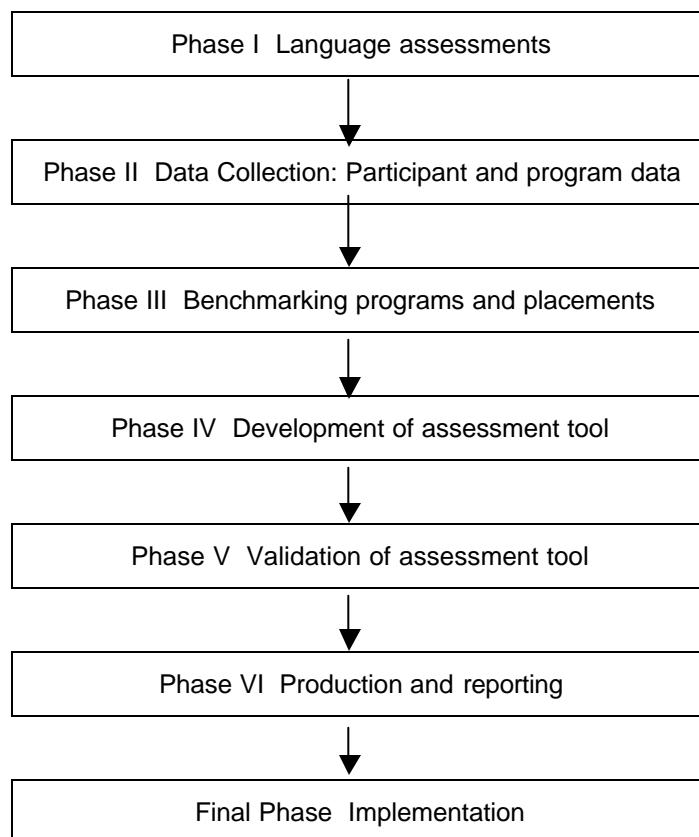
Once the benchmark range for the program and clinical practice was established, the team selected sample tasks with the help of content experts (professionals working in the identified professions). These sample tasks represented the actual workplace situations. The collaboration between the test developers and the content experts ensured the authenticity of the test tasks.

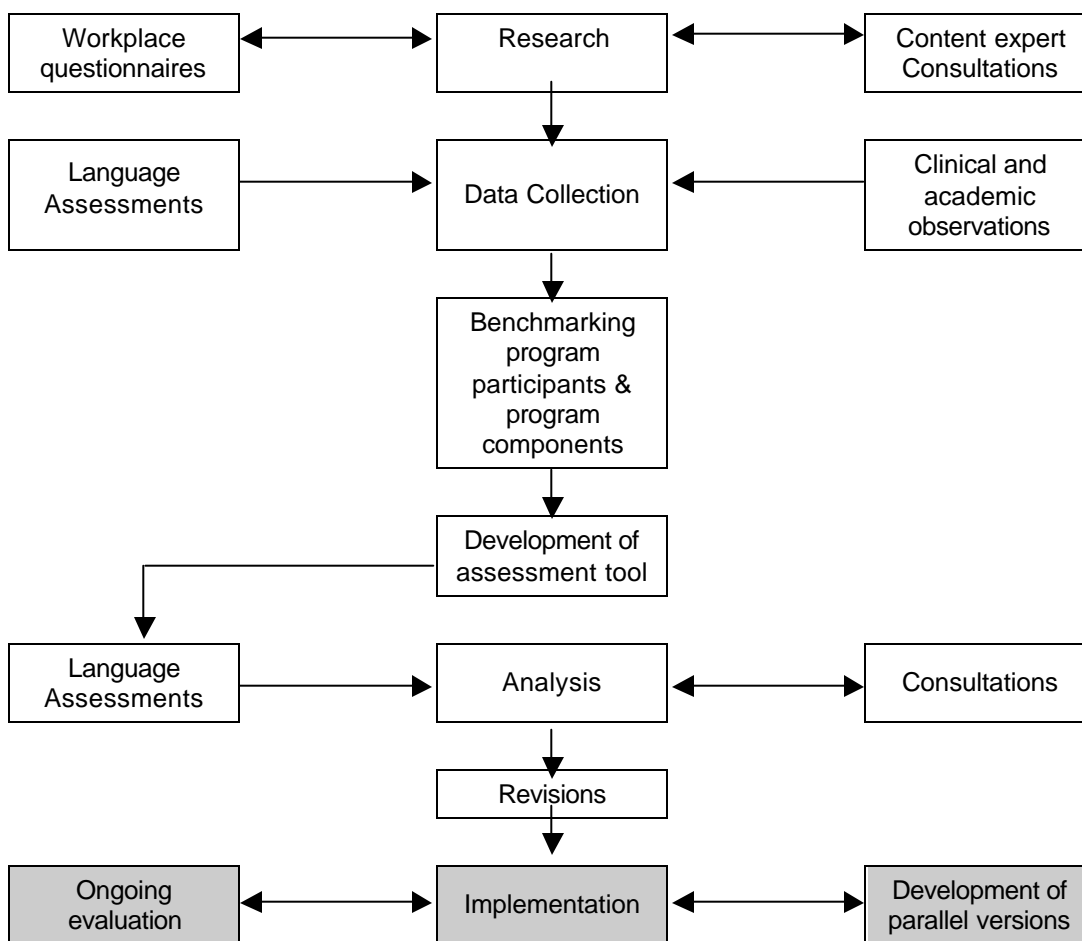
Test specifications were then developed based on the target language use and the CLB framework. This was followed by the development of tasks and rating scales that were pilot tested with programs participants. Once a draft version of the test tasks was ready, a pilot test was conducted. Data from the pilot test helped to determine test task accuracy, inter-rater reliability, and concurrent validity with the CLBPT.

## **Work Plan**

Although the project can be simply described in terms of “phases” these activities were certainly not linear. One of the learnings we can identify through this project is that rather than thinking of the various components of the work in terms of phases, as depicted in the chart above, the project is more accurately illustrated in a flow chart form, which demonstrates an ongoing and spiralling effect of each project activity:

The following chart describes the phases of the project:





The shaded boxes indicate future initiatives.

## Project Outcomes

In the original proposal to Citizenship and Immigration Canada (CIC) – Enhanced Language Training Initiative, project proponents set out expected outcomes based on the various stages of the project. Although the work plan did not follow a linear pattern (which caused a challenge in terms of reporting in the proposed timeline), the objectives for each stage were achieved nevertheless.

In this part of the report we will describe the activities and results of each of these phases. This chart summarizes the information to follow:

### ***Project Evaluation Framework: Expected & Actual Outcomes***

<b>Phases</b>	<b>Expected Outcomes</b>	<b>Actual Outcomes</b>
Phase I Language Assessments	60 Access & Options participants' language skills benchmarked.	Over the course of the project LCRT consulting assessed 147 applicants to the A&O program using the CLBPT; 24 of these were assessed a second time using the new assessment tool.
Phase II Data Collection	Participant and program data is collected and structured for analysis.	Besides the CLB scores for 147 project applicants the project team gathered information from 47 practitioners through an online survey.
Phase III Benchmarking	Clear description of communicative tasks for academic and clinical practice in Michener Programs.	An analysis of the 75 samples of course materials that were collected and 7 hospital observations combined with the practitioner survey results yielded a recommended minimum CLB for each stream of the A&O program.
Phase IV Development	CLB Occupation-specific Assessment Tool	With input from content experts and the project managers, the test development team prepared test specifications. Once proposed test tasks were confirmed as authentic, the team built a 2.5 hours test that included reading, writing, listening and speaking occupation-specific tasks.
Phase V Validation	40 Access & Options students benchmarked.	41 participants took part in the pilot test, 12 from core Michener programs. Pilot test results indicated a high level of satisfaction on the part of the test takers, good inter-rater reliability, and reasonable concurrent validity with the CLBPT.
Phase VI Production and Reporting	Final reports for general circulation	During the course of the project three formal presentations were delivered. These form the basis of a template for use in future presentations. Additionally, test materials have been produced to a professional standard.

## Benchmarking of Project Participants (Phase I)

### **Objective**

In Phase I project participants' language ability was described in terms of the Canadian Language Benchmarks through the use of the Canadian Language Benchmarks Language Placement Tool (CLBPT). The purpose of benchmarking project participants was to establish a baseline benchmark for each candidate that would then be used to:

1. Compare with Michener students in full time programs who use English as a second language;
2. Track participant improvement over the course of the program; and
3. Compare CLBs with participant success in the program.

### **Methodology**

The CLBPT was used to benchmark Access & Options students' communicative performance in five different testing sessions (September 2003, January, April, September and November 2004). Each session represented an intake of a new cohort into the Access & Options program.

The CLBPT is one of two currently available CLB assessment tools endorsed by the Centre for Canadian Language Benchmarks (CCLB). The use of the tool is licensed by the CCLB; only certified CLB assessors conduct the assessments. LCRT Consulting holds a license to administer the CLBPT. The Centre for Canadian Language Benchmarks (CCLB) granted LCRT permission to use the CLBPT for the purposes of this project.

The CLBPT was selected for this project because it gives reliable descriptions for language benchmarks of the test takers and it is streamlined resulting in efficiency in terms of time and cost. However, the CLBPT is a placement test, and as such the results indicate readiness to enter an English as a second language program rather than an occupation-specific bridging program.

The CLBPT yields results in speaking, listening, reading and writing, as described in the Canadian Language Benchmarks. Participants were asked to complete a pre-test questionnaire, and to sign a consent form that described the purpose of the assessment. A reporting form with CLBPT results and a referral list were also prepared for clients.

The consent form that test takers were asked to complete acknowledged the use of their data for a research project. The questionnaire requested information regarding first language, experience with English in the workplace, level of language proficiency as demonstrated in standardized English language proficiency exams, and participant's own perception of their language skills.

### **Outcome**

A total of 147 individuals were assessed using the CLBPT between August 2003 and August 2004. Most test takers demonstrated language skills in CLB Stage II (CLB 5-8).

The results of the CLBPT identified that 25 % of A&O applicants demonstrated fluent intermediate proficiency or higher<sup>8</sup> (CLB 8), 32% demonstrated adequate intermediate proficiency (CLB 7), 33% demonstrated developing intermediate language proficiency (CLB 6), and 8% demonstrated initial intermediate language proficiency. Based on the benchmarking analysis conducted in Phase III of this project, this data would indicate that 41% of the applicants

---

<sup>8</sup> The CLBPT measures proficiency in Canadian Language Benchmarks 1 (initial basic proficiency) to 8 (fluent intermediate proficiency).

to the program do not have the language skills required to succeed, whereas 59% of the applicants demonstrated language skills conducive to success in the program.

Another interesting finding relates to self-awareness. Applicants were asked to identify their own level of language proficiency based on a chart ranging from beginner to advanced, and to indicate this on a scale of 1-12. Based on the CLBPT results, about 70 percent of participants overestimated their ability. The number climbs even higher for the listening test alone. This would indicate that a self-assessment or full language assessment is necessary upon entry into a program that requires a minimum level of language proficiency.

### ***Impact***

The CLBPT allows for a maximum score of CLB 8 on the 12-point CLB scale. It is probable that many of the tests takers in the test sessions conducted for this project have skills that reach into CLB Stage III (CLB 9-12).

The CLBPT was designed to measure language proficiency for the purposes of placement into English as a Second Language (ESL) programs. Because its purpose is *not* to determine readiness for professional upgrading programs and work placements, the tool was used in this project as a way to establish benchmarks of participants for research purposes only. Once research results regarding the benchmark requirements for the program emerged out of the work conducted in Phase III, program staff referred to the CLBPT results for counselling and referrals purposes. The assessment team prepared a list of ESL programs in the area for referral purposes.

Based on research conducted during Phase III, which established a minimum CLB for participation in the program to be CLB 7, the participant benchmarking results would indicate that 41% of the participants in the A&O have language skills adequate for the program in which they have enrolled. As a result of this finding, the project began to refer applicants with CLB 5/6 to language upgrading with a conditional acceptance in the upcoming session pending proof of participation in an ESL program.

This finding illustrated the need for language assessments to be conducted upon entry into a bridging program, particularly if a financial investment is required for the participant. The language assessment can provide some idea of the participant's readiness to succeed. It is important for applicants to receive a diagnostic report and appropriate referrals so that they can focus their language learning and be better prepared for the next intake session.

The CLBPT scores were also used as a point of comparison with test results with the new assessment tool (see discussion in the test development section). It is recommended that a thorough longitudinal study be conducted which will scientifically determine whether this hypothesis is accurate.

## **Participant and Program Data Collection (Phase II)**

### ***Objectives***

During phase II the research team collected participant and program data with the aim to analyse the information in terms of the communicative competencies required within the Access & Options program as well as requirements to complete the professional registration process (professional practice examinations, application process and interviews); and communication observed in clinical practice.

## **Methodology**

During Phase II the following information was collected:

- Participant data collected during CLBPT assessment process;
- Sample documents for the registration process for medical professionals in this project (i.e. application forms, information on regulations and legislation, examination handbooks, etc.);
- Descriptions of the registration procedures for medical professionals in this project (i.e. fact sheets, interviews with regulators on the process and difficulties faced during this process);
- Any research on professional competencies, with a focus on the communication requirements, for each occupation represented in the project (i.e. Regulatory body competency statements; HRDC Essential Skills Profiles);
- Observation of courses offered to the bridging program participants;
- Informal conversations with instructors and students;
- Review of sample materials (i.e. lecture handouts, PowerPoint notes, etc.);
- Observation of work placements; and
- Documentation from the workplace.

The research team prepared a data collection form that allowed for a methodical record keeping of the information discovered through the analysis of the above-noted documents and activities. Checklists for reading, writing, speaking/listening tasks observed in the workplace were also developed. A consent form was completed for each workplace observation.

These forms and procedures were designed to collect:

- Participant data, including proficiency tests already taken, participant evaluation of their own language skills, and information on work experience in Ontario;
- Language requirements for the certification process (including proof of language proficiency and levels of complexity in documentation required for the process);
- Language requirements of the A&O Program (academic and work placement) articulated in terms of the Canadian Language Benchmarks (CLB); and
- Language competencies identified by practitioners through a survey.

Access & Options program managers provided samples of materials used in the program and facilitated access to courses for the purpose of observation and discussion with students and instructors. Researchers accessed information on professional competencies from regulatory bodies.

Following the observation and review of materials, the researchers developed a survey for practitioners which was made available online. The survey was based on research into the communicative competencies required for each of the occupations and the equivalent CLB descriptors. Program consultants helped to identify professionals in the medical community who would be willing to offer their insight into language use in the workplace. Forty-seven individuals completed this detailed questionnaire.

The survey responses from practitioners were key in determining what competencies were most frequently required in each of the occupations so that the team could:

- a) Verify that the test tasks under development were authentic; and
- b) Establish common communicative competencies among the occupations so that a streamlined, and therefore more efficient test could be developed.

### **Outcome**

Approximately 75 samples were collected and prepared for an analysis based on the Canadian Language Benchmarks. Researchers organized the hard copies of samples and observation notes into binders. The data was collated and organized for the analysis to be conducted in Phase III.

### **Impact**

The information collected in Phase II was used not only as representative samples for the benchmarking process, but also as a way to identify target language use<sup>9</sup> for the medical professions and then as a starting point for the development of occupation-specific test items.

In terms of establishing benchmarks for the professions, the broadest range of target language use situations should be observed. Given the time and resources available, this project's scope was fairly limited. A longitudinal study, which would compare the new test results with success in the program, is recommended.

## **Benchmarking Programs and Placements (Phase III)**

### **Objectives**

The objective of Phase III was to analyse the data collected in Phase II, and articulate the language competencies required for successful participation in the program in terms of the Canadian Language Benchmarks. It was expected that this analysis would allow for generalizations to be made in regards to the benchmarks required for each medical health profession in the project.

Overall, the information collected in Phase II was used for two main purposes:

1. Each "sample" collected in Phase II was benchmarked in Phase III using an expert judgment standard-setting approach.
2. "Benchmark" the occupations based on the benchmarking analysis.

The benchmarking process helped to establish a minimum benchmark competency statement for the Access & Options program and the "samples" helped identify target language use for the medical professions in the project as a starting point for test development.

### **Methodology**

The benchmarking process involved describing language use in different situations (academic upgrading and clinical workplace) and establishing a connection between these communication competencies and the competency standards in the different levels of the Canadian Language Benchmarks.

---

<sup>9</sup> "Target language use" is the phrase used by test developers to describe the language that would be typically required in the situations for which test takers are being tested. In other words, if we are testing to determine whether A&O applicants will be able to succeed in their certification process, we need to test whether they are able to perform in those situations (i.e. upgrading seminars, clinical placements, test-taking).

Each course observation, workplace observation, interview, and sample material was entered into the data collection table. Specific language competencies for each activity were then described and compared to the benchmark descriptors. The descriptors were analysed against the CLB competency standards using an expert judgement standard setting method where the two researchers benchmarked the samples independently, then compared results and discussed any differences until arriving at an agreement. Seventy-five samples of communication tasks were benchmarked in this manner.

In addition to observations and review of program materials, professional competency charts were prepared based on key competencies for each profession. The research team then conducted a set of consultations with Access & Options program managers who are registered health professionals and provided much of the expert advice for the project..

The first set of test tasks was designed based on this list of communicative competencies. The list was also essential in helping the team identify communicative competencies that are common to the three medical health professions in the Access & Options project. This enabled the team to design a test that would efficiently assess language skills in three different sets of professionals.

### ***Challenges in the Benchmarking Process***

A number of challenges presented themselves throughout the benchmarking process. The first being the limited time and resources made available for the benchmarking process, and for the research component of the project. In hindsight, and based on CCLB recommendations, a minimum of six months should be allocated to the benchmarking process alone<sup>10</sup>. Secondly, a much larger sample of workplace activities should be reviewed, and a larger research team would enrich the standard setting process.

Secondly, it was not easy to get access to the workplace, even through the work placements organized by Access & Options program. This is due partly because the timeline of the project did not always coincide with student placements.

Another challenge related to the fact that the CLB competency standards are based on four skill areas (speaking, listening, reading, writing), and in a number of our samples, some skills were represented, and others not. For example, when analysing a textbook, the focus was on reading, but we did not necessarily participate in the lecture related to that reading so we were unable to document the task beyond reading comprehension. Clearly when observing workplace practice, all four skill areas were observed and recorded. A participant-research style of research in the classroom and clinical session, where the researcher follows a group through a two-week period, might yield broader results. Within the Access & Options program this would require a longer commitment based on the part-time nature of the program.

While the skills-based approach of the benchmark descriptors did sometimes pose a barrier to the benchmark analysis, the integrated nature of some of the communicative competencies described in the benchmarks are clearly congruent with the reality of communicating in the workplace.

The research indicated almost immediately that the academic language demands of the project are very different, and often higher, than those required in clinical settings. It is evident that the academic requirements of these occupations are more rigorous than workplace language use. This is due mostly to the complexity of academic reading texts (reading comprehension skills) and the amount of information being presented verbally through long, technical lectures (listening

---

<sup>10</sup> Centre for Canadian Language Benchmarks (2004) *Developing an Occupation-specific Language Assessment Tool Using the Canadian Language Benchmarks* p. 16

comprehension skills). The benchmarking of the occupations will be influenced by the higher benchmarks required for the academic work. For this reason we reported each area (workplace and academic) separately.

### **Outcomes**

The research and benchmarking findings of the project identified that although the three occupations vary greatly in their demands, it is clear that language skills in the CLB 7-9 range are required for full participation in the program. Strictly speaking, CLB 9 would be ideal for the academic portion of the program. However, evidence of success among Access & Options participants with CLB 7-8 indicates otherwise. An important factor to consider is content knowledge and each individual's medical experience, which often have a greater impact on the learning and performance in the program than language skills alone. The motivation of the adult learner can often overcome other barriers.

This is another reason to conduct occupation-specific language testing. This occupation specific language test allows experienced, knowledgeable adults to demonstrate their communication skill in a relevant, respectful manner. The increased comfort with the context of the test will allow for a more authentic demonstration of communicative performance on the part of the test takers. Participants in the pilot performed better and attested to feeling more satisfied with their performance on a language assessment that reflected tasks with which they are familiar.

### **Recommended CLB**

The following chart represents the team's findings for Phase III. It is important to keep in mind that only a random sampling of classes were observed, and these may not be representative of the entire program. Additionally, researchers had limited exposure to the workplace. That being said, the thorough review of the materials collected, the observations conducted, and the benchmark analysis led the research team to the following benchmarks. The benchmarks listed here reflect an average for each situation.

		<b>CLB Speaking</b>	<b>CLB Listening</b>	<b>CLB Reading</b>	<b>CLB Writing</b>
Medical Laboratory Technology	Academic	7	10	9	7
	Clinical	7	8	7	7
	Overall	7	9	8	7
Radiological Sciences	Academic	7	9	8	7
	Clinical	9	7/9	8	8
	Overall	8	8	8	7
Respiratory Therapy	Academic	7/8	9	7/9	8/9
	Clinical	9	9	8	8
	Overall	8	9	8	8

Overall, the researchers have concluded that although the three occupations vary somewhat in the communicative demands, it is clear that language skills in the CLB 7-9 range are required. For

those entering the program with a CLB 7, it is recommended that additional supports be offered as they prepare for the demands of the professional practice exams and professional practice itself.

### **Impact**

CLBPT assessments indicated that the average CLBPT scores for participants were 6.9 for speaking, 6.2 for listening, 6.7 for reading, and 6.6 for writing (See CLBPT Assessment Report for more information). According to the benchmark analysis, the benchmarks required for the program activities lie in the CLB 7-9 range. The benchmark scores and the benchmark analysis would indicate that the overall language abilities of the applicant group fall below the language abilities required for participation in the program, in clinical placements, and eventual certification and employment. The impact of these findings is that additional supports should be put in place to optimize the participants' chances for success. It also indicates that with language screening the program will likely experience a higher level of success.

Another important finding which has an impact on the program is that there is a clear difference between the academic and the workplace language required in the occupations analysed which is probably typical of many professions. This creates a challenge in terms of setting a minimum language level for entry into the Access & Options program if the minimum language level required for clinical practice is lower than that for the academic component of the program. It is recommended that a longitudinal study of project participants be conducted to determine whether those applicants scoring CLB 7 on the benchmarks test are not as successful as those scoring CLB 8 or 9.

The analysis also indicates that the training methods within the program should continue to emphasize clinical laboratory practice and professional practice exam preparation, but also that technical courses currently being presented in lecture format might be delivered in an alternative format (i.e. online with study groups or quiz sessions) so that comprehension is facilitated for those at CLB 7. Supports such as tutoring and other study skills and strategies could also help these individuals.

In terms of test development, based on the benchmarking of the participants and the benchmarking of the program and professional clinical practice, the team decided to design a test that would measure CLB 6-10. The test would include tasks that represent both practical and academic tasks, and task performance expectations would be set at four different levels (CLB 6-10) of difficulty to allow test takers to demonstrate their ability at each level. This system allows for each test taker to be placed within a benchmark, and for a diagnostic statement of their language "level" to be delivered. An added advantage of the benchmarked test results will be the possibility to design communication courses that address the specific needs of different cohorts.

And finally, based on the research in the benchmarking phase of the project, a cut score (pass mark) must be recommended. Although a CLB 8 would seem appropriate based on the overall benchmark results (see table below), the results of the CLBPT and MILA (Michener Institute Language Assessment) testing indicate that there are a number of participants in the program who have scored in the CLB 7 range and have succeeded.

The following table illustrates the recommended minimum CLB for entry into the program:

	CLB Speaking	CLB Listening	CLB Reading	CLB Writing
Medical Laboratory Technology	7	8	8	7
Radiological Sciences	8	8	8	7
Respiratory Therapy	8	8	8	8

### Development of Assessment Tool (Phase IV)

The information collected in Phase II and analysed in Phase III provided a clear statement of target language use for each profession, which included situations and tasks that exemplify what a person would be required to do in the program and in clinical placements. This has made it possible to identify common language competencies among the three professions.

The test development process involved creating test tasks based on these authentic communication competencies required for practice in these professions.

#### **Methodology**

The test development phase comprised three major steps: the development of test specifications, consultation with content experts, and test task design.

##### *Develop Test Specifications*

The first step in the test development process was to write a set of Draft Test Specifications. The specifications documents provided a “roadmap” for the test developers to follow in creating tests tasks and rating scales. The Test Specifications remained a “living document” throughout the development phase, and were revised thoroughly after the pilot test and test revisions. The final test specifications provide a blueprint for the development of future versions of the test.

In developing the test specifications, test developers referred primarily to the following documents:

Centre for Canadian Language benchmarks (2000) *Canadian Language Benchmarks 2000: English as Second Language for Adults*

Centre for Canadian Language benchmarks (2002) *Canadian Language Benchmarks 2000: Theoretical Framework*

Centre for Canadian Language benchmarks (2002) *Canadian Language Benchmarks 2000: Additional Sample Task Ideas*

Bachman, L and Palmer, A (1996) *Language Testing in Practice: Designing and Developing Useful Language Tests*

Douglas, D (2000) *Assessing Languages for Specific Purposes*

The test specifications document includes:

1. *A Statement of the Test Purpose.* This statement describes the intended purpose of the test. This is the foundation of the test design, and test administration. The test will be part of the Access & Options Program at the Michener Institute for Applied Health Sciences. It will be

used to determine the applicants' Canadian Language Benchmarks (CLB). The test will provide diagnostic information on language proficiency levels of these applicants. A CLB score determined by the A & O program will determine whether or not a candidate has the language fluency requirements to participate in the program. The program is a bridge to full certification, therefore the test should predict whether a candidate has the minimum required language proficiency to complete the certification process.

*2. Characteristics of the Language Users and Test Takers:* Every test is designed for a specific audience. The characteristics of this target audience must be clearly explained as a way to ensure that the test tasks are suitable. For this test, the target audience is: internationally trained medical technology professionals in three areas: radiology, medical laboratory technology, and respiratory therapy. Having already satisfied the professional requirements, this group is poised to begin the licensure process, which will require clinical practice and professional practice exams. The group includes both men and women, typically from mid-twenties to early forties. They come from a variety of national and linguistic backgrounds from Europe, Asia and Africa but at present mainly from the Middle East and Southeast Asia. Test takers will be at various stages in the medical health technology licensing process and demonstrate varied levels of communicative competence.

*3. A Description of Target Language Situation and Target Language Tasks:* This part of the test specification was initially informed by the description of professional competencies as discovered through research in the initial stage of the project. As the project unfolded, classroom and workplace observations, as well as consultations with practitioners helped the team define authentic target language situations and tasks. Phases II and III were essential in helping us establish what language use situations and tasks to represent in the test.

*4. Constructs to Measure:* Based on analysis of target language use, and on the CLB framework, the aspects of the language that need to be measured were listed. These included: grammatical knowledge, textual knowledge; functional knowledge, socio-linguistic knowledge and background knowledge.

*4. Test Content:* This section of the test specifications describes the test itself, how it is organized, the time required, the scoring criteria and the scoring procedures for the test.

For reasons of test security the specific contents of the test specifications reside with the *Michener International Language Assessment Test Package*. More information can be obtained through the Access & Options program or LCRT Consulting.

#### *Consultations with Experts*

The test tasks were developed in consultations with Access & Options staff and other Michener Institute content experts. A core "focus group of experts" was established. These experts agreed to participate in consultations either individually or as a group, and offered to provide authentic workplace materials.

Three focus group meetings were held with all content experts. During the first meeting content experts received an introduction to the test purpose and a clear description of their roles. Content experts were also asked during this first meeting to identify reading materials and authentic workplace scenarios for the test task development. The second and third meetings involved the review of the proposed test tasks which the test developers prepared based on the authentic workplace reading materials, writing tasks and communication scenarios submitted by the focus group members.

Content experts greatly facilitate the design of test tasks. In this project content experts can be seen posing in picture stories, acting in workplace interaction on video, and as lecturers in the test's lecturette. Their contribution was invaluable and ensured authenticity of the test task content as well as the task itself.

The rating scales for the tasks were based on performance descriptors and indicators in the Canadian Language Benchmarks. This process was much more complex than it seemed initially as the team tried to reconcile the benchmark descriptors with the actual tasks, and in the amount of effort it took to establish the precise break from one benchmark level to another. In test development the time required to set up rating systems should not be underestimated.

### *Test Design*

The language assessment design was influenced by a number of factors. The greatest influence on the test design came from the CLB framework. There are four distinct skill areas in the CLB: speaking, listening, reading, and writing. The test developer's first challenge was to design tasks that were reflective of real life tasks, and therefore integrated, yet which yielded a score for a specific skill area.

It was decided to offer each test taker an opportunity to demonstrate their ability to perform four different tasks in each language skill area, each at a higher level of difficulty. In other words, for each skill area, four tasks must be completed, presented at increasing levels of difficulty.

The second challenge faced by the test developers relates to the need for a streamlined and efficient assessment. One of the major problems in language assessment is the cost. Therefore, the test development team set out to design a test that would be both comprehensive, but also efficient in terms of time, human resources, and materials.

It was decided that, although demanding in terms of human resources, an interview is essential to determining competence in speaking. As a result, the test development team designed a 20-minute interview, to be conducted one-on-one. Interviewers record the session and rate a second time, comparing the results to ensure consistency. The other parts of the test can be conducted in a group setting and are not as demanding in terms of test administration.

Another aspect of the efficiency required of this language assessment is the fact that the assessment, although "occupation-specific" in fact encompassed three different occupations. In order to incorporate these three different occupations into one testing process it was necessary to carefully establish common competencies. An online consultation process asked practitioners to identify tasks that were typical for each occupation. The results of these consultations were cross-referenced and converted into set of common competencies. As a result, the test includes both interdisciplinary and profession-specific tasks. The interdisciplinary tasks contain medical technology/therapy content.

The tasks are performance based, so although organized by skill area, one skill is demonstrated by the use of another. For example, listening to telephone messages is an important workplace competency and involves both listening comprehension and intelligible note taking. Likewise, writing in the workplace includes the use of forms as well as the ability to write coherent and accurate paragraphs, as seen in the Incident Report task.

The interview moves from a discussion about personal and familiar topics, to more familiar workplace topics, and finally to an unfamiliar abstract and complex workplace problem following the Canadian Language Benchmarks model of increasing language complexity. Likewise, during the listening test, the taker moves from simple note taking while listening to a monologue, to listening to dialogues, and finally to a more formal mini-lecture. In reading, test takers start by

reading simple forms, and move on to read moderately complex text (formal workplace forms and memos) and finally complex academic texts (textbook excerpt; journal article). Writing tasks reflect the writing of the workplace for the most part: form filling, writing a summary-report, writing a message to a colleague and finally a more complex academic task, a short opinion essay.

Once the test tasks were designed, the test developers organized the performance descriptors into rating scales for each test task. A system was developed and tested during the pilot test.

### **Outcome**

As a result of these different steps taken in Phase IV (Test Design), the team built an assessment process comprising tests for four different skill areas that encompasses many aspects of the academic and work requirements for health care professionals. The test offers each test taker at least four opportunities to demonstrate their skills in each of the four language skill areas. This comprehensive design makes it possible for the test administrator to assign benchmark levels to the test takers with confidence.

The performance conditions for each task are based on the benchmarks, and as in the benchmarks, tasks increase in level of difficulty. This means that it is possible to determine whether the test taker meets the benchmarks requirements of each task resulting in the assigning of a CLB to the performance of the test taker.

The test takes two and a half hours to administer, including the one-on-one interview. This is not onerous for the test taker, and it makes it possible to schedule two sessions in a day, allowing for a greater number of assessments to be conducted.

The following chart summarizes the test format and design:

<b>Speaking</b>	<b>Listening</b>	<b>Reading</b>	<b>Writing</b>
Interview Part 1: personal background & professional experience CLB 6/7	Take workplace phone messages CLB 6/7	Read the description of a health and safety incident in a hospital CLB 6	Take simple to complex telephone messages CLB 6
Interview Part 2: describe workplace process using a photo story CLB 7/8	Receive workplace related instructions over the telephone CLB 7/8	Read a Workplace memo CLB 7	Fill out incident form CLB 7
Interview Part 3: Role-play offering detailed technical instructions in the workplace CLB 8/9	Listen to a workplace conversation CLB 8/9	Read a complex interdisciplinary text CLB 8	Write a short opinion essay CLB 8
Interview Part 4: Problem-solving discussion CLB 9	Professional development lecturette CLB 9	Read a professional journal editorial CLB 9	Summarize lecturette CLB 9

Materials required for test administration include the test booklets, an audio player (laptop and speakers), and an LCD projector. Test booklets must be prepared prior to test administration.

The time required of assessors is 20 minutes for the interview, another 20 minutes to review the recorded interview, and 30-40 minutes for marking. This means that the minimum cost per test would include the fees for 1 hour and the test printing costs (the location is assumed to be provided).

### ***Impact***

During the evaluation of the test, participants commented that they felt the test reflected their professional needs. Through this assessment second language speakers have a chance to demonstrate language skills that are relevant and realistic for their purposes. The test is targeted to a specific purpose, and both test takers and the institutions using the test can be confident that the test measures language performance. In other words, the test measures what and how well a person can perform different tasks in English. Because the tasks have been directly linked to the A&O program and the requirements for licensure, the language test tells us about the likelihood that the test taker will succeed in the program. Test developers feel confident about generalizing from the test scores since each task represents an actual speech act within health science professions.

This project has had a great impact on the Michener Institute. Although the test purpose for this project is very clearly set out in the ELT project outcomes, both the A&O program and the Michener Institute envisage a number of potential applications for an occupation-specific language test. Also, throughout the development of the test the developers and the A&O staff worked on clarifying the test purpose, and discussing the possibilities of developing parallel versions for different purposes (i.e. admission of first years students). There is also interest from regulators in the test as a way for applicants to meet the objective fluency requirement for registration.

## **Evaluation of Assessment Tool (Phase V)**

The test development phase was followed by an evaluation phase during which a pilot test was conducted. In this section we discuss how the pilot test was organized, and what the results told us about the test design.

### ***Pilot Test Objectives***

The pilot test set out to test a number of different aspects of the test. Specifically we wanted to answer the following questions:

1. Are the test tasks authentic and engaging for the test taker? (test validity)
2. Do the test takers feel they were given a fair opportunity to demonstrate their language skills? (test validity)
3. What is the level of difficulty for the target group? (construct validity)
4. How confident are we that the CLB results are accurate? (construct validity)
5. Is there inter-rater reliability? (reliability)
6. Is the test accomplishing what it claims to do? (test validity – long term)

In order to answer these questions, the test development team and the A&O office recruited participants who represented a sample of the target audience to participate in a pilot test and give feedback on the test. Their scores were used to analyze the test items, and their feedback informed the revisions.

## ***Planning the Pilot Test***

### *Participant Recruitment*

Invitations to participate in the pilot test were sent out one month prior to the pilot test dates. Two dates were made available, a Saturday and a Sunday, as well as a choice of morning or afternoon sessions. A \$50 honorarium was offered. The invitations were coordinated through the Access & Options Office. This ensured that the participant group reflected the target audience and that the privacy of A&O applicants was maintained (it would not have been possible to share a mailing list with an outside organization). Of a total of 180 contacts, 41 actually participated in the testing.

### *Scheduling*

Two groups of approximately 20 participants were scheduled to participate in the pilot test. Given the modular nature of the test, it was possible to organize the group into rotating test sessions so that all interviews could be conducted within the 3 hours of testing. This required three interviewers/assessors, three interview rooms, and two testing rooms.

### *Participant Consent*

Pilot test participants received an orientation to the test and the pilot test objectives at the beginning of the testing session. Each participant was asked to sign a consent form, which explained that the data would be used in aggregate format, and that the interviews would be recorded for inter-rater reliability testing. All participants consented.

### *Questionnaire*

Pilot test participants were asked to complete a questionnaire at the end of the pilot testing session. Most participants found the test to be at an adequate level of difficulty, and the content interesting. The group also reported liking the skills-based approach, but from their feedback it is clear that participants had most difficulty with the listening component of the test. Participant feedback was analysed against the test results and considered during the test revisions.

### *Recruiting Assessors/Raters*

The test administration requires a number of assessors to conduct interviews. For the pilot test, we organized the pilot test so that it was necessary to have test invigilators in the reading/writing and listening/writing test rooms, and three interviewers.

In the interview, the interviewer is required not only to follow the script, but also to evaluate the speech sample of four different tasks within the interview. For this reason, the administration of the interview requires an in-depth understanding of the Canadian Language Benchmarks to facilitate ease in rating the interviews accurately.

The test development team set out to engage assessors/raters who had this experience. The prerequisite established for assessors/raters was experience in CLB testing, either with the Canadian Language Benchmarks Assessment (CLBA), the Canadian Language Benchmarks Placement Tool (CLBPT) or other CLB-related in-house assessments.

We were able to recruit three CLBPT assessors, two of whom also had experience in administering the linguistic component of the Pharmacy PLA within the International Pharmacy Graduate Program at the University of Toronto, which is also CLB-based.

A two-hour orientation session was conducted. As there were no oral or written samples to review, the orientation consisted of reviewing the interview format and discussing the rating scale.

### ***Pilot Test Methodology***

#### *Number of Participants*

Forty-one participants took part in the testing sessions, which took place over two weekends. Of the total number of participants, 68% were current, previous or future A&O students, and 32% were Michener students.

#### *Evaluating the Interview*

For the pilot test, 40 interviews were conducted (One candidate arrived late and missed the scheduled interview). Assessors for the pilot test have experience with CLB assessments, and received a 2-hour orientation session. All interviews were recorded. Thirteen interviews were rated independently by three separate assessors and scores compared for inter-rater reliability. The scores for the remaining interviews were rated by one assessor, and reviewed by the test developer.

#### *Evaluating the Listening Tasks*

The listening tasks are scored using answer keys that were created based on CLB competencies, performance outcomes and evaluation criteria.. The pilot test was therefore a way to verify the answer keys rather than test inter-rater reliability. The 41 tests were rated based on a set of answer keys, then reviewed by the test development team.

#### *Evaluating the Reading Tasks*

The pilot test results helped us to identify parts of the reading tasks that were unclear, too easy, or too difficult for the test takers, and revisions have been made accordingly. As in the listening task, the reading tasks are scored using answer keys. The pilot test was therefore a way to verify the answer keys rather than test inter-rater reliability. All tests were scored, then reviewed by the test development team. The team tried two different rating procedures before settling on one that best identified the scores in terms of CLB.

#### *Evaluating the Writing Tasks*

For the writing portions of the test it was essential to verify the level of inter-rater reliability given that these tasks are productive tasks and are rated using descriptive bands, which can be subjective unless there is a strong standard setting mechanism among raters. Because the CLB are referenced in our test, it is easier to set this standard among a group of CLB-trained assessors.

### ***Pilot Test Outcomes***

The results overall indicated that the test content was interesting and engaging for the test takers, that the test items are generally sound, and the rating systems are strong enough to result in good reliability among raters.

#### *Interview*

With 52% of the test takers in the CLB 9 and 10 categories, and 48% in the CLB 6-8 category, the test demonstrates the ability to place test takers in Stage II and III of the benchmarks. Given

the nature of the group, the distribution tends towards the higher levels of proficiency. We must observe over the next testing session to identify whether the results are consistently higher with the applicant group as well. However, research indicates that test takers perform better on tests that are relevant to their experience, and this too might explain the higher overall scores.

Assessors reported that the interview script succeeded in eliciting the type of input required for each task. They found the format easy to use, the flow logical and the instructions fairly clear to them and for the test takers. Most assessors were able to complete their interviews in the time recommended, although this is an area where training will prove useful as a way to ensure that the interviews are efficient and conducted in a timely manner.

Participants reported that the interview was relevant and interesting. Participants outside of the A&O program reported difficulty with the photo story and role-play as these assumed previous experience in the field. This nevertheless confirms that the test is meeting its occupation-specific purpose.

Based on the participant questionnaire results, the test takers really enjoyed the narrative (photo-story) and felt the test was related to their professions:

*“I like the interview part.”*

*“I like the picture test most.”*

*“It’s interesting as it is related to my profession. (medical)”*

Inter-rater reliability for the assessment of the speaking tasks was very good. We found a high rate of correlation in the scoring of the speaking tasks: 77%, meaning that 7 out of 10 times the assessors agree. Of the remaining 3%, assessors disagree by one point only.

### *Listening*

As in speaking the distribution spans Stage II and II, with 57% of the test takers scoring CLB 9 or 10, and 43% scoring CLB 5-8. Unlike the CLBPT, listening scores on the MILA tend to be closer to the speaking scores. We feel that this reflects the integrated nature of the test, and listening tasks that are relevant and engaging.

Test takers overall felt the content appropriate and stated they understood it. However, most test takers felt that the task involved too much writing, rather than listening. Few complained about the first task (taking notes on a pre-set form) or the dialogue task (which involved answering multiple choice questions), but many felt the lecture was difficult. This is natural given that the lecture task is a CLB 9 task, and a significant percentage of the test takers demonstrated listening skills in the CLB 6-8 range.

In terms of rating, the answer key provides reliability. A test item analysis indicates that the level of difficulty of the test tasks is adequate for the CLB being tested in each task, including the level of difficulty of the lecture, which is set at CLB 9. We discovered that the audio-visual component of the listening test was not as helpful to test takers as we had originally assumed. In observing the session we noted that many test takers did not look up to view the video while listening. We also found that it was more difficult for test takers to process the lecture along with the Power Point notes on-screen. Test takers found it easier to simply listen to the text and take notes. As a result, the listening test includes a variety of listening inputs as a way to ensure that all learning styles are addressed.

### *Reading*

Two different versions of the interdisciplinary reading passages were tested during the pilot test, and it was possible to determine parallelism between them.

The reading task rating system was designed based on the *CLB 2000: Theoretical Framework*, which outlines a checklist for evaluating reading competencies in understanding both prose texts and formatted texts. The system need some adjustment, but worked well in the subsequent testing session conducted for A&O. Since the marking is based on answer keys, there was no disagreement among raters.

An analysis of the test scores for individual tasks and overall scores indicates that the tasks were effective in determining the benchmark levels of participants. There is a tendency for the test takers to score better on the reading test (69% scored CLB 9 & 10). Our analysis of these results is that if 37 of the 41 test takers in the pilot test group were either in the Michener full time program or in the A&O program, the likelihood of their English level being higher is stronger.

### *Writing*

Twenty-five percent of the writing samples were rated by two different assessors to verify inter-rater reliability. Of the four tests tasks in the test, the lecture summary and the essay were the most difficult to score objectively.

Rating scales for the writing tasks included both a holistic and analytic scores in keeping with the *CLB 2000: Theoretical Framework*.

As in the previous test sections, a large number of test takers scored in the CLB 9-10 range (44%). The writing samples demonstrated a general ability o complete the tasks, but a varying level of accuracy, control, and flexibility in using the language to communicate through writing.

### *Concurrent validity against the CLBPT*

In terms of concurrent validity against the CLBPT, this was difficult to ascertain. Of the 41 test takers, only 24 had completed the CLBPT. Of these, 6 completed the CLBPT in 2003, so enough time has expired to render the results unreliable. Also, of the 24 test takers who have completed a CLBPT test, 13 received a CLB 8 on the CLBPT. Because CLB 8 is the highest score one can receive on the CLBPT, but the Canadian Language Benchmarks scale reaches to CLB 12, it is not possible to know with certainty whether the higher score on the MILA is because of the test design or improvement on the part of the test taker. The remaining 8 items are insufficient for a reliable analysis.

### ***Pilot Test Challenges and Learnings***

A persistent challenge in pilot testing is participant recruitment. Our test group mostly represented the target audience, but was quite different from the group we benchmarked during the research phase. This made any comparisons between the two tests unreliable. Additionally, the outreach plan aimed to recruit 60 participants, but only 41 actually participated in the test. (And this with a \$50 honorarium.)

Conducting the test development and pilot testing of the assessment tool within an already established program was an ideal model for test development. The test development team was able to rely on the project delivery office to reach out to the appropriate target language group, and coordinate the scheduling and room bookings. Having the in-kind contribution of testing space made it possible for the project money to be spent on development.

Another challenge is time, both finding a time suitable to the largest possible group of test takers, and time required to complete the test. We offered the testing sessions on the weekend, and offered morning and afternoon sessions. In order to accommodate a large group of test takers in a short period of time (3 hours) we organize the interviews and the testing so that participants do not have long waiting periods. Three hours is a large investment of time on a weekend for working adults with families.

Finding qualified assessors was also a challenge, particularly for pilot testing. Now that the tool has been developed, it is hoped that some funds will be made available to develop an assessor-training module. LCRT hopes to create a pool of qualified interviewers/raters to conduct the test on a regular basis.

The MILA is a comprehensive test, and as such it requires two and one half to three hours to complete. We learned that it is critical to organize the logistics and identify the best scheduling and flow for the test to maximize the use of time and ensure that test takers are not waiting unnecessarily.

One of the greatest challenges in the analysis of the results is the concurrent validity study. This proves to be very difficult given that there are no other CLB tests for testing proficiency above CLB 8. The data we collected through the CLBPT is inadequate for this comparison given that only 24% of the test group had previously taken the CLBPT, and of those, a large number scored CLB 8. The CLB8 scores on the CLBPT are not reliable given that the test takers could in fact be CLB 9 or 10. Another barrier to this analysis is the fact that the MILA is such a unique test, and to compare an occupation-specific test results with a generic language test is somewhat informative, but it is doubtful whether it should tell us with certainty whether there was concurrent validity.

## **Production and Reporting (Phase VI)**

A number of presentations were prepared and delivered to different stakeholders throughout the project. These will be organized into a presentation which test developers will have ready for use in disseminating information about this project. Already the team has confirmed participation on a panel that will be discussing the development of occupation specific language tests at the TESL Ontario Conference in November 2005.

During the test development process a number of audio/visual tools were developed. A lecture was video taped and recorded for use as one of the listening tasks. Two dialogues were staged and acted by content experts helping with the project. A number of listening tracks were scripted and recorded. A series of photographs were taken and organized into photo stories (a sampling of these appear on the front cover of this report). All of these materials have been organized into a testing package on CDROM that contains the testing booklets and the test tasks.

These materials are considered confidential as this first version of the test will be used by the Michener Access and Options Program as part of their intake program in 2005. A second version of the test must be developed soon to ensure that the test remains valid.

## Project Learnings & Recommendations

### ***Time & Resources***

The proposed timeframe spanned September 2003 (pre-ELT project, the ELT project began in January 2004) to March 2005. For a project of this magnitude it is recommended that a two-year span with at least two full time researchers and test developers.

One of the surprises of this project was the fact that the development of rating scales based on the Canadian Language Benchmarks was much more challenging than expected. The integration of the CLB standards with the performance conditions of occupation-specific tasks is a task that requires time and much testing. In hindsight, we feel that more time should have been devoted to the development and testing of test tasks individually before the pilot test.

The development of the *Canadian Language benchmarks Language Assessment for Nurses (CELBAN)* required three years and a much larger research team.

### ***Importance of Partnerships***

Partnerships and other collaborative relationships established during this project were key to making the project work. ESL specialists do not have the insight of workplace culture and processes and must work in conjunction with employers and trainers in the development of occupation-specific training and testing materials.

The collaborative nature of the research and the development process for a project of this kind is paramount. However, this is a process that required adequate time. When working in partnerships, there is much less control over the timelines. Workplaces and training locations have competing priorities that undermine the ability to complete research and consultations in a timely manner. It is recommended that this be recognized, and that projects be planned to allow for the development over time of meaningful relationships between various stakeholders.

Another reason for reconsidering the length of the project was the benchmarking research and analysis. That part of the project alone could easily take up a year. When planning a project that required the participation of many stakeholders, consultations, observations and other activities that are bilateral, the time factor cannot be controlled as easily.

### ***A Basis for Further Research***

The test developed through this project represents a first step in the development of an occupation-specific language assessment for health science professionals. The benefits of the availability of an assessment tool of this kind include a thorough analysis of the professional competencies for these three occupations. These materials would provide an excellent basis for the development of occupation-specific ESL curricula for CLB 5, 6, and 7.

It will also be possible now to conduct a study that looks at performance on an occupation-specific language assessment versus performance on a generic language proficiency test (such as the TOEFL or IELTS). With recent changes that bring tests like TOEFL more in line with the CLB philosophy it becomes more valid to test these assessments against each other.

### ***Longitudinal Study Recommended***

It is imperative that a longitudinal analysis be conducted to validate the construct validity of the test (i.e. determine whether the test is doing what it claims to do: determine an adequate minimum level of language ability required for success in the program). Pilot test results indicate that scores are on average higher than the CLBPT scores, although this statistics is very unreliable given that the test groups differed greatly from the CLBPT testing session and the MILA testing session. No Michener students were tested with the CLBPT so a strong comparison between the two test groups is not possible. It will be important to observe results over the next testing sessions to identify whether the results are consistently higher with the applicant group as well.

It is recommended that A&O and LCRT Consulting explore different sources of funding for ongoing research and further development of this assessment tool.

### ***Computer-Mediated Communication***

Communication is computer mediated in the medical technology field. The use of computer mediated test tasks must be considered in order to make the test even more authentic. Computer literacy is a key communicative competency, but it is absent from the assessment developed here. It is recommended that we explore ways to computerize the various parts of this assessment, and consider the availability of computer labs for testing purposes.

Much of the MILA is in a format that can easily be converted into a computer environment with the proper financial and time investment it is possible to develop a more streamlined computerized model of this assessment tool.

### ***Different Englishes***

One of the most interesting challenges in this project was the question of assessing English language competency in speakers of sub-continental or non-standard English. Applicants from the Philippines and India (among others) demonstrated fluency and confidence in speaking and writing, but the control and accuracy was lacking. The pattern emerging is that these speakers of “other Englishes” use a dialect or version of English that is considered appropriate in a number of settings, but which may cause misunderstanding and confusion in the target workplace.

This issue requires attention. Our approach is to apply the CLB descriptors to all test takers, regardless of their language background. This is the nature of a standardized test. We feel justified in doing so since the performance descriptors are based on how successful the individual is in communicating their message effectively and accurately, characteristics of communication that are valued in the workplace. However, there are many examples of successfully adapted speakers of “other Englishes” in the workplace, so it would be important to discover to what extent the varieties of English interfere with successful workplace performance.

### ***Attention to the Target Audience***

Another learning from the pilot test is the fact that the test is for a very specific target audience. During our pilot test we invited regular Michener students to participate as a control group. The students who volunteered were first year students. The test presumes knowledge and experience in one of the three occupations represented in the program. These volunteers were unable to adequately complete the role-play task, as they did not have background experience, although they are proficient in English.

### **Recommended Cut Points**

Based on the research conducted through this project and the test developer's understanding of the requirements of the A&O program, the following cut points are recommended:

<b>S</b>	<b>L</b>	<b>R</b>	<b>W</b>	<b>Recommendation</b>
CLB 6	CLB 6	CLB 6	CLB 6	A candidate who scores CLB 6 should be referred to a language training program for upgrading prior to entry into the program.
CLB 7	CLB 7	CLB 7	CLB 7	A candidate who scores CLB 7 is minimally competent in completing complex communicative tasks effectively and may experience some difficulty in the A&O program. Recommend that other qualifications be considered when accepting this candidate into the program. The candidate is likely to experience difficulty based on their communication skills.
CLB 8	CLB 8	CLB 8	CLB 8	If a candidate scores all CLB 8 or mostly CLB 8 the candidate would likely succeed in the A&O program.
CLB 9	CLB 9	CLB 9	CLB 9	A candidate who scores CLB 9 will likely not experience difficulty in completing the A&O program based on communication.
CLB 10	CLB 10	CLB 10	CLB 10	A candidate who scores CLB 10 has communicative competence to fully participate and excel in the A&O program.

Candidates scoring three out of the four skill areas should be considered within that benchmarks. For example, an applicant scoring S CLB 8, L CLB 7, R CLB 8 and W CLB 8 should be considered as having CLB 8 overall.

### **Recommended Supports within the A&O Program**

One of the benefits of a comprehensive assessment tool is the ability to diagnose areas of need among the program applicants. Given that a need analysis is possible, it would be also possible to develop targeted language training programs to support those applicants who meet the minimum requirements, or language training programs to bridge those who do not meet the requirements into the program.

It is recommended that A&O and LCRT Consulting explore the possibility of developing an occupation-specific ELT program based on the findings of this project. It would be possible to develop a curriculum for delivery in immigrant agencies across Ontario, to develop a program for the Michener Institute, or to consider the possibility of an online course.