The Development of CELBAN

(The Canadian English Language Benchmark Assessment for Nurses):

A Nursing - Specific Language Assessment Tool

September 2003

Centre for Canadian Language Benchmarks

Acknowledgements

- newly-practising nurses (both internationally-educated and Canadian-educated) who participated in pilot-testing
- individuals who assisting in setting up focus groups in 4 provinces
- stakeholders who participated in focus groups
- actors who participated in the Listening Test Video/Audio
- CCLB Board Members, CCLB Nursing Committee: Rob Boldt (chair), Jim Jones, Peggy Frederikse, Carolyn Dieleman, Margaret Pidlaski, Pauline McNaughton (CCLB Executive Director), and Marianne Kayed (CCLB Project Manager)

Contents

7 - CELBAN Test Development and Administrative Reference Documents 48

8 - Appendices 49

Glossary of Terms 63

Tables

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

1 - Executive Summary

A growing shortage of nursing professionals in Canada is projected over the next ten years. Internationally-educated nurses entering the profession in Canada could ease this projected shortage. However, one of the issues involved in licensing these nurses is language proficiency, and how it is measured. Stakeholders have indicated the need for a nursing-specific assessment tool to facilitate integration of nurses into the profession. Based on this need, the Centre for Canadian Language Benchmarks (CCLB) has initiated a project, *Benchmarking the Nursing Profession Across Canada*. The process involves the following four steps:

1. A Feasibility Study (completed in 2001)

2. Phase I: *An Analysis of the Language Demands of the Nursing Profession Across Canada* (completed in 2002)

3. Phase II: *The Development of a Nursing-Specific Language Assessment Tool* (CELBAN)

4. Phase III: Implementation of the CELBAN

The results of Phase II, *The Development of a Nursing-Specific English Language Assessment Tool*, are presented in this report. This assessment tool is named the Canadian English Language Benchmarks Assessment for Nurses (CELBAN). The test was developed by a team of consultants from a wide range of backgrounds and locations across Canada. The CELBAN was developed in six stages.

Stage One: Planning

Stage Two: Development of Draft One of the Assessment Tool

Stage Three: Piloting of Draft One of the Assessment Tool

Stage Four: Revision of Draft One/Development of Draft Two

Stage Five: Piloting of Draft Two of the Assessment Tool

Stage Six: Development of Final Assessment Tool and Final Reports

---

Using the results of Phase I, *An Analysis of the Language Demands of the Nursing Profession across Canada*, the assessment tool was developed with content that reflects the language demands of the nursing profession. The CELBAN measures the English language proficiency of nurses in four separate areas: Speaking, Listening, Reading and Writing. The desired outcome for candidates is to achieve

scores that reflect the Canadian Language Benchmark (CLB) Levels necessary for entry into the nursing profession in Canada as determined in Phase I of the project:

SKILL	CLB Level
Speaking	8
Listening	9
Reading	8
Writing	7

The CELBAN was pilot tested with internationally-educated nurses and nursing students at six locations across Canada. Altogether, 270 candidates were tested. Statistics were analyzed and feedback was considered carefully to ensure the reliability and validity of the test. The process of test development was comprehensive, rigorous, and inclusive, utilizing extremely valuable assistance and feedback from an excellent test development team, and a wide range of stakeholders. The result is a thorough assessment of the English proficiency of the candidates in Speaking, Listening, Reading and Writing articulated in terms of Canadian Language Benchmarks (CLB) Levels. Not only do candidates receive their scores in the four skill areas, they also receive written feedback from the assessors regarding their strengths and weaknesses in the productive skills, Speaking and Writing.

With Phase II now completed, the test development team recommends that the CCLB proceed with Phase III, Implementation of the CELBAN across Canada.

2 • 1 - Executive Summary

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

2 - Introduction

Project Background

The Centre for Canadian Language Benchmarks (CCLB)[1] has initiated this project, *The Development of the Canadian English Language Benchmarks Assessment for Nurses (CELBAN).* Following an extensive analysis of the language demands of the nursing profession (Phase I, *Benchmarking the Nursing Profession Across Canada*), and with feedback provided from a wide range of stakeholders, the recommendation was made to the CCLB to develop an English language assessment tool specifically for internationally-educated nurses.

The project is designed to address the critical shortage of nursing professionals in Canada. Statistics Canada anticipates that over the next five years, a large percentage of nurses will retire. A recent report, *Planning for the Future: Nursing Human Resource Projections* (Canadian Nursing Association, 2002) stated that putting together the demand and supply figures, projections suggest that there will be a shortage of 78,000 RNs in 2001 and 113,000 RNs by 2016. The Ontario Ministry of Health and Long Term Care, in a previous report, *Ensuring the Care Will be There* (Registered Nursing Association of Ontario with Registered Practical Nurses Association of Ontario, 2000)*,* stated that unless solutions are found, and found soon, the country's health care system will suffer significantly. This view extends across the country among members of the health care profession and members of the general public.

Internationally-educated nurses entering the profession in Canada could help to ease the projected shortage. However, one of the issues involved in licensing these nurses is language competence and how it is measured. In 2000-2001, with funding from the Ontario Government, the CCLB completed a feasibility study entitled, *Benchmarking the Nursing Profession and Developing an Occupation Specific Assessment Instrument*. This study included a survey of over 50 professional nursing stakeholder organizations across Canada, who were contacted to explore interest in a nursing-specific English language assessment instrument. This survey focused on occupational groups and regulatory bodies. One of the key questions posed was whether stakeholders believed a nursing language assessment tool would be of

---

[1] The CCLB is a national, not-for-profit organization, primarily serving the adult English as a Second Language (ESL) community in Canada including learners, teachers, program administrators, and materials, curriculum and test developers. A Canada-wide combination of language training specialists, assessment service providers and both federal and provincial government members forms the CCLB Board of Directors and staff of the CCLB are committed to maintaining and promoting language proficiency standards based on the Canadian Language Benchmarks.

benefit. Ninety-two percent (92%) of respondents responded YES, which led the CCLB to undertake this project.

Further results of the feasibility study were even more encouraging. Stakeholders generally indicated that existing assessment instruments such as the Test of English as a Foreign Language (TOEFL) and the Test of Spoken English (TSE) were too general to adequately evaluate the ability of internationally-educated nurses to communicate effectively in the profession in Canada. The benefits of a nursing-specific assessment instrument expressed in Canadian Language Benchmarks, as cited by stakeholders, are summarized as follows:

- To help internationally-educated nurses who are not presently working as nurses in Canada to enter the profession.

- To provide a standard means of assessing English language competence for internationally-educated nurses.

- To alleviate the need for a national centre for the assessment of applicants to the nursing profession educated outside Canada.

The results of this feasibility study clearly indicated the need for a nursing-specific assessment tool. However, before such a tool could be developed, it was necessary to do an in-depth analysis of the English language demands of the nursing profession in Canada.

Based on the results of the feasibility study, the CCLB undertook the next stage of the project, *Benchmarking the English Language Demands of the Nursing Profession Across Canada* (Phase I). That project was completed in 2002. More information on this project is available in the report available at the CCLB website, www.language.ca . The Canadian Language Benchmarks (CLB) was used as the standard to measure the language demands of the nursing profession.

What are the Canadian Language Benchmarks?

The CLB is a descriptive scale of communicative proficiency in ESL, expressed as benchmarks or reference points. They provide a framework of reference for learning, teaching, programming and assessing adult ESL in Canada, and a national standard for planning second language curricula for a variety of contexts, a common "yardstick" for assessing the outcomes. The CLB descriptors are available in the document, *Canadian Language Benchmarks 2000,* which can be ordered at the Website www.language.ca at no charge. The CLB provides descriptors for four language skills: speaking, listening, reading and writing on a scale from CLB Level 1 to CLB Level 12. These twelve levels are divided into three stages: Stage I, Basic Proficiency (Levels 1 to 4); Stage II, Intermediate Proficiency (Levels 5 to 8); and Stage III, Advanced Proficiency (Levels 9 to 12).

The CLB was developed in response to a 1992 consultation undertaken by the Government of Canada through the department now called Citizenship and Immigration Canada (CIC). This consultation with experts in second language teaching and training, testing and measurement confirmed that no one instrument, tool or set of "benchmarks" was widely used or appropriate to Canadian newcomers' needs. A national working group on language benchmarks was established by CIC in 1993 to oversee and guide the development of benchmarks. Field-testing of a draft document was implemented in 1995, and in 1996 the CLB Working Document was ready for distribution and use in English. In 1999 revisions were made to the CLB Working Document based on feedback from stakeholders across the country. Based on this feedback, the Canadian Language Benchmarks 2000 was published. The CLB is presently used in English language training programs across the country to determine content and curricula of ESL programs. According to the CLB 2000 (p. IX),

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

> …the CLB standards can help to articulate ESL needs, practices and accomplishments. They can also facilitate clear communication throughout the ESL community, and between it and other community/national organizations and agendas (e.g., instructors, learners, educational programs, assessors and counsellors, language education funding bodies, labour market associations, licensing bodies, and employers).

Because the CLB provides a common language to discuss levels of language proficiency, it has the potential to be useful for a wider range of applications. It has been used to benchmark academic programs, occupations, and assessment tools. Two assessment tools, Canadian Language Benchmarks Assessment (CLBA) and Canadian Language Benchmarks Placement Test (CLBPT) have previously been developed to measure English language proficiency based on the CLB. These tools are used primarily to assist in placing ESL learners in appropriate ESL classes, and to provide learners with a scale to describe their language proficiency.

In Phase I, *Benchmarking the English Language Demands of the Nursing Profession Across Canada*, it was determined that the CLB Levels needed for entry into the nursing profession in Canada were as follows:

SKILL	CLB Level
Speaking	8
Listening	9
Reading	8
Writing	7

This report describes Phase II of the project, *The Development of a Nursing Specific Language Assessment Tool* (CELBAN). Using the results from Phase I, the CELBAN has been developed to assess the English language proficiency of internationally-educated nurses whose native language is not English. The contents of the CELBAN reflect the Canadian nursing context, and the results are reported as CLB Levels in four skill areas: Speaking, Listening, Reading and Writing.

Project Overview

The CCLB is undertaking this project, Phase II, *The Development of a Nursing Specific Language Assessment Tool* (CELBAN)*,* as the third part of a four-step process*.* This process has involved the following steps:

1. A Feasibility Study (completed in 2001)

2. Phase I: *An Analysis of the English Language Demands of the Nursing Profession across Canada* (completed in 2002)

3. **Phase II:** *The Development of a Nursing-Specific English Language Assessment Tool* **(now named the CELBAN) (completed in 2003)**

4. Phase III: *Implementation of the* CELBAN (to be initiated in late 2003)

Phase II, *The Development of the* CELBAN*,* is funded by the governments of Alberta, British Columbia, Ontario and Manitoba with the Ontario Region of Citizenship and Immigration Canada. It has been carried out by the Language Training Centre, Red River College, Winnipeg, Manitoba, together with a team of experts from various locations across Canada.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

2 - Introduction • 5

The project outcome is the development of an occupation-specific English language assessment tool which measures the specific English language proficiency of internationally-educated nurses whose native language is not English. This test provides:

- an occupation-specific English language assessment tool that facilitates access to the nursing profession for internationally-educated nurses.

- a task-based assessment tool that accurately reflects the language demands of the nursing profession in Canada.

- separate scores for each candidate in the language skills of Speaking, Listening, Reading and Writing reported as CLB Levels.

- an analysis of strengths and weaknesses of productive skills (Speaking and Writing) for each candidate.

This assessment tool, the CELBAN, provides a CLB based alternative to language tests that merely evaluate a candidate's general academic knowledge of the English language.

The project is expected to benefit nursing colleges, regulatory bodies, health care employers and internationally-educated nurses. It will offer stakeholders a relevant and accurate picture of the English language proficiency of internationally-educated nurses who apply to practise in the profession in Canada.

There are several reasons for the use of an occupation-specific language assessment tool for the recognition of the language proficiency of internationally-educated professionals. One reason is the growing awareness that the present system for assessing language proficiency lacks relevance and appropriateness. For immigrants whose native language is *not* English, assessment/recognition of language skills often becomes a barrier to accessing their occupation/profession. All stakeholders would agree that language skills must be sufficient to function on the job. Language assessment tools presently used for internationally-educated nurses test academic language skills, and do not in fact measure the language skills needed in a specific occupation. As a result, there are those who can pass the test, but cannot function in the workplace. Others take these language tests and fail, yet at the same time may have the potential to function successfully on the job. For example, in the nursing profession, non-native English speakers are frequently required to take an English test in which they must write an essay on a topic totally unrelated to nursing, a task which does not reflect the type of writing a nurse does on the job. The advantage of an occupation-specific language assessment tool is that the language skills used on the job are tested in a more authentic way, resulting in a more accurate assessment process. It should be kept in mind that all tests have inherent limitations, and at best they can only provide an *indication* of how a person would likely perform in a similar task *in the real world.* However, in the case of the CELBAN, while limitations may still exist, they are significantly *less* than those of other tests being currently used.

Based on assessment research, there is additional justification for the use of occupation-specific language assessment tools. According to Alderson (a prominent researcher in the field of second language test development), when relationship between task characteristics and test-taker characteristics is established, this results in what has been termed "interactiveness". This ability of the test-taker to interact with the text is an essential element in test validity.[2]

In addition to this report, the test

This report describes the process undertaken in the development of the CELBAN (Canadian English Language Benchmarks Assessment for Nurses).

---

[2] Alderson, J. Charles (2000). *Assessing Reading,* Cambridge University Press, 165.

6 ● 2 - Introduction

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

*developers have
designed a "**How-To
Manual**" for other
professions and
occupations interested
in conducting a
similar project.
Information about this
resource is available
by contacting the
CCLB through their
website:
www.language.ca .*

3 - CELBAN Project Methodology

Project Methodology

In developing the CELBAN, preliminary steps included:

- the establishment of a test development team.

- an application for research /ethics approval.

- the establishment of a time-line for each stage of development.

The **test development team** for the CELBAN was identified in the initial planning stage of the project. The researchers/test developers were second language experts with excellent knowledge of the Canadian Language Benchmarks, and experience in applying the CLB framework to the workplace. A team of consultants in the fields of applied linguistics, test and measurement, statistics, and nursing were contracted to provide input at each stage of the development of the assessment tool. In addition, there were two test reviewers contracted to critique the test before the first piloting, and provide feedback about preliminary revisions.

In total, there were 16 participants on the team including the following: a project manager, 2 researchers/test developers, 3 project consultants, 3 nursing consultants, a statistics consultant, 2 test development consultants, a linguistics consultant, a test & measurement consultant, and 2 test reviewers. These individuals were all experts in their particular fields and assisted from several educational and provincial locations such as the CanTEST Project Office (University of Ottawa), Red River College (Winnipeg), University of Manitoba, Ontario Institute of Studies in Education (University of Toronto), and a private consulting firm in Edmonton.

The test developers applied early for **research/ethics approval** from the Research Approval Committee at Red River College[3]. As is the case at many educational institutions, this committee meets monthly, with a monthly deadline for applications to be considered. Application requires a summary of the proposed research and a detailed description of test development procedures. Samples of materials including consent forms, letters to participants, focus group agenda, pilot testing protocol, and confidentiality agreements were submitted. Also included were details regarding the use and reporting of results and findings.

---

[3] This committee operates according to the Tri-Council Policy Statement on the Ethical Conduct for research involving humans. It ensures that research proposals are coordinated, follow ethical guidelines, and serve the wider purpose of educational knowledge.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

After the creation of a test development team and application for ethics approval, the **time-line for each stage of development** of the CELBAN was designed as follows:

Stage	Description	# of Weeks	Dates
One	Planning	5 weeks	Oct.21/02 – Nov. 22/02
Two	Development of Draft One of the CELBAN	8 weeks	Nov. 25/02 – Jan. 18/03
Three	Pilot Testing Draft One of the CELBAN	7 weeks	Jan. 20/03 – March 7/03
Four	Revision of Draft One/ Development of Draft Two	6 weeks	March 10/03 – Apr.19/03
Five	Pilot Testing of Draft Two of CELBAN	6 weeks	Aprl 21/03 – May 30/03
Six	Development of a Final Draft fo the CELBAN & final reporting	7 weeks	June 2/03 – July 18/03

Stage One: Planning (5 weeks)

Planning for the development of the occupation-specific language tool was a crucial part of the project. It included the following elements:

- literature review

- initial application for ethics approval for research

- initial contact with potential pilot testing sites

- development of a confidentiality agreement

- the establishment of a National Advisory Committee (NAC)

- initial development of test specifications

The test developers did an initial **literature review** to become more up-to-date with current literature related to the development of an occupation-specific assessment tool. This review was an important piece in obtaining background information and establishing a theoretical framework. A bibliography of the resources used for the development of the CELBAN was compiled **(see Appendix A)**.

The National Advisory Council (NAC) was established. The main function of the NAC was to provide feedback at various stages of the project. The NAC members were kept informed of the progress of the project through interim reports. To identify potential NAC members, contacts used during Phase I (*An Analysis of the Language Demands of the Nursing Profession*) were utilized as a starting point. The NAC was composed of a wide range of stakeholders representing as many provinces and territories as possible including registrars, instructors, directors of professional organizations and associations, labour relations representatives, regulatory bodies, policy analysts, and internationally-educated professionals practicing in the field. There were 22 members on the NAC from Newfoundland, Ontario, Manitoba, Saskatchewan, Alberta and British Columbia. The NAC members were:

- Ana Maria Revilla, Counselor and Outreach Coordinator, CARE for Nurses, Toronto, ON

- Jean Barry, Policy Advisor, Canadian Nursing Association

- Brenda Lewis, Registration Consultant, College of Nurses of Ontario

- Bula Ghosh, Past President/Instructor, Immigrant, Refugee, and Visible Minority Women of SK/Cypress Hills Regional College

- Debbie Carry, Acute Care Coordinator, Meadow Lake Hospital, SK

- Paul Fisher, Executive Director/Registrar, Council for Licensed Practical Nurses of NF

- Collin Mercer, Manager, Nursing Consultant, BC Ministry of Community, Aboriginal & Women's Services

- Carolyn Sams, Nursing Consultant, Strategic Leadership Group

- Debbie Stewart, Labour Relations Director, MB Nurses Union

- Heather Hawkins, Director, Regulatory Services, Registered Nurses Association of NF

- Maya Charlebois, Administrative Director for Healthy Communities, Calgary Regional Health Authority

- Muna Muqled, ICU Nurse, Ottawa Civic Hospital

- Ricki Grushcow, Director, Ontario Hospital Association

- Laura Schneider, Manager of Health Programs, Alberta Learning, Government of AB

- Verna Holgate, Executive Director, College of Licensed Practical Nurses of MB

- Lena Nikolsky, Neo-natal ICU Nurse, Winnipeg Health Sciences Centre – Children's Hospital

- Peggy Frederikse, Senior Policy & Project Consultant (also CCLB Board Member), Ontario Ministry of Training, Colleges, and Universities, Access to Professions & Trades Unit

- Rob Boldt, Program Design & Evaluation, Settlement & Multiculturalism Branch, Aboriginal, Multiculturalism, and Immigration Programs Dept., Ministry of Community, Aboriginal and Women's Services (also CCLB Board Member)

- Marianne Kayed, Project Manager, Centre for Canadian Language Benchmarks

- June Rock, Registrar, Alberta Association of Registered Nurses

- Carla Taylor, Advisor – Initial Registration, Registered Nursing Association of BC

- Susan Haglund, Director, Regulatory Services, College of Licensed Practical Nurses of BC

Initial contacts with potential pilot testing sites were made through contacts from the NAC. The requirements for ethics approval at each potential site were explored.

A **confidentiality agreement** was designed for team members who had access to specific information related to the CELBAN.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

For Lessons learned during Stage One, see the How-To Manual, available from the CCLB web site.[4]

The development of test specifications was a lengthy and ongoing process. *It was seen as a circular, not linear process.* At each stage of the project, test specifications were amended, revised, and elaborated upon. Information for writing test specifications was provided by consultants at OISE and also the CanTEST Project Office. A sample of the CanTEST Specifications, as well as examples from the literature review, were helpful. These resources provided a framework for the CELBAN Test Specifications (Confidential Document).

Stage Two: Development of Draft One of the Assessment Tool

Stage Two involved the following components:

- the design of four sub-tests: Listening, Speaking, Reading, Writing

- the development scoring methods, i.e. Speaking & Writing grids (rubrics)

- the development of various forms for candidates

- a review of Draft One of the assessment tool by test reviewers

- finalized plans for pilot-testing Draft One

The information and resources collected during Phase I, *An Analysis of the English Language Demands of the Nursing Profession Across Canada*, were an integral part of the development of the assessment tool. The Phase I data provided authentic texts, tasks, and scenarios for designing the assessment tool, which contributed to the "face validity"[5] of the test. Verbatim data from observations were used to provide a framework for speaking and listening scenarios. The pie chart analysis of the interactions and tasks observed provided helpful information (see **"**Appendix B - Charts" on page 53). Feedback from stakeholders was also considered. Samples of reading and writing during observations provided authentic text models. The nursing consultants provided excellent feedback and support in assuring the authenticity of test items and tasks.

The assessment tool was designed with four separate sub-tests to assess each aspect of language: Speaking, Listening, Reading and Writing. The test development team members provided feedback on the overall framework for the test, individual components, and specific items. Current testing methodology was incorporated into the construct of the assessment tool.

The development of Speaking and Writing grids was a complex process requiring great familiarity with the CLB document. The consultants with expertise in test development were an important resource during this part of the development process. These grids were revised on an ongoing basis throughout the process, as descriptors and performance indicators were clarified and more clearly articulated.

---

[4] The "*How-To" Manual* is a document written by the test developers of the CELBAN, and provides a detailed account of the steps involved in Benchmarking the English Language Demands of a Profession, and Developing an Occupation-Specific English Language Assessment Tool. Information on this resource is available by contacting the CCLB at their website: www.language.ca

[5] "Face validity" is the extent to which a test meets the expectations of those involved in its use, e.g. administrators, teachers, candidates, and test score users; the acceptability of a test to its stakeholders.

Forms were developed for candidates who would be participating in pilot-testing of the first draft of the assessment tool. The following forms were designed:

- a letter of explanation of the project for candidates (see **"Appendix C - Letter of Explanation to Pilot Testing Candidates"** on page 55)

- a questionnaire (see "Appendix D - Questionnaire For Participants In Pilot Testing" on page 54)

- a consent form for candidates (see "Appendix E - Consent Form For Pilot Testing Candidates" on page 57)

- feedback forms for each component of the test (see **"Appendix F - Feedback From Candidates"** on page 58)

- a form to report scores to candidates, including space to list strengths and weaknesses in productive skills - Speaking and Writing (see **"Appendix G – CELBAN Test Results"** on page 59)

Once a complete draft for each component (Listening, Speaking, Reading and Writing) was developed, it was sent to the test reviewers. (The language and test-development experts provide input throughout the development time; however, the reviewers critiqued the completed draft.) Revisions to the test were made based on feedback from the reviewers.

Arrangements for pilot-testing of Draft One were then finalized. The goal was to have a number of provinces and educational institutes involved in the pilot-testing. Students and internationally-educated professionals, who were recently practicing or still undergoing credentialing, were prime test candidates. At each location, establishing contacts who could assist with the organization of the pilot-testing was crucial. An effort was made to do pilot-testing at sites where English for Nursing Purposes programs were being delivered. At these sites, whole classes were available as candidates, and scheduling was most efficient. NAC members provided excellent contact information. At each site, one person agreed to be responsible for the testing arrangements. A small honorarium was provided to those contacts for their assistance.

For Lessons learned during Stage Two, see the How-To Manual.

After the completion of Stage One and Two (12 weeks into the project) an interim report was submitted to the NAC. This report included a progress report and a general summary of Draft One of the assessment tool (types of text, and tasks). A *general* summary was provided because of the confidential nature of the specific test content.

Stage Three: Pilot Testing of Draft One of the CELBAN

Pilot-testing of Draft One (Pilot A) of the assessment tool was conducted in Stage Three. It provided the test developers with qualitative and quantitative data on which to base revisions. It also informed the test developers about the strengths and weaknesses of the assessment tool and the delivery of it, both through statistical analysis and anecdotal reports. Both major changes and fine-tuning resulted from the pilot-testing For the CELBAN, the goal was to pilot-test Draft One with:

- 80 L2[6] internationally-educated professionals in the field (in the process of having credentials recognized).

---

[6] L2 refers to candidates whose native language is **not** English.

- 40 L1[7] students who had completed some field experience in the profession.
- 5 L1 & L2 professionals presently practicing (within 5 years) in Canada.
- Total number of candidates: 125

Pilot-testing was conducted in Toronto (CARE for Nurses Project)and Winnipeg (Red River College) with a total of 126 candidates. The CELBAN was administered, and feedback was collected from the candidates after they had completed all sections of the test. Information was analyzed and utilized in making revisions for the second draft of the test.   Also, various aspects of the administration of the test (e.g. instructions provided to candidates before, during and after; time-frame for completing each section of the test, etc.) were adjusted.

Scoring of pilot tests, recording of results and recording feedback for candidates was accomplished immediately following the administration of the test. In a high-stakes assessment such as this, candidates experience high levels of stress and so the test developers felt it was important to report results as soon as possible. The goal was to mail the results to candidates within one week of the candidate's completion of the test.

Another means of testing language proficiency of the candidates was conducted by the test developers for a comparative analysis. Correlations established informed the test-developers of the validity and reliability of the results on the new assessment. During pilot-testing of the CELBAN, the candidates were assessed using the Canadian Language Benchmarks Placement Test (CLBPT)[8] to establish comparisons.

Once pilot-testing of Draft One was completed at each site, a thorough and immediate statistical analysis of the data was undertaken. Crucial indicators were *Alpha reliabilities* of each component of the test, *indicators of face validity,* and *item discrimination.*

During the time-frame in which pilot-testing occurred, focus groups were also conducted with stakeholders in the same cities. As with Phase I, the focus groups in Phase II were important because they provided feedback from a wide range of stakeholders in each region. It was an opportunity for stakeholders with different perspectives to hear and discuss their concerns. It also provided an opportunity for stakeholders to network with each other in ways that perhaps were not previously experienced.

At the focus groups for the discussion of the CELBAN, items on the agenda included: introductions and greetings, completion of confidentiality agreements, background information on the CLB, a description of Phase I, *An Analysis of the Language Demands of the Nursing Profession across Canada*, and an overview of Phase II, *Developing a Nursing-Specific English Language Proficiency Assessment Tool*. Following the presentations, stakeholders had time for an organized discussion concerning the following 3 questions:

- How does this project address your needs?
- What further needs must be addressed?

---

[7] L1 refers to candidates whose native language is English.

[8] The CLBPT is a task-based assessment tool referenced to the CLB 2000. It is designed to determine the English language proficiency of newcomers to Canada who speak English as a Second Language.   Results are provided in listening, speaking, reading and writing. The proficiency levels are based upon the competencies described in CLB 2000 covering CLB Levels 1-8. For more information contact the CCLB at www.language.ca .

- What are the issues/concerns regarding implementation?

Stage Four: Revision of Draft One & Development of Draft Two

During Stage Four, revisions to Draft One of the assessment tool were made based on analysis of the statistics and feedback from consultants and participants. The importance of statistical analysis and anecdotal data must not be underestimated. Having a team of experts representing a wide range of relevant perspectives was also extremely necessary in revising, re-writing, and adding or deleting test-items or complete sections.

Final arrangements for piloting-testing of Draft Two were also made during this stage. Again, these arrangements were made through contacts from the NAC, and through other contacts established during the course of the project. Networking of ideas and resources was a bonus in having a lot of direct communication and interaction with these contacts.

Plans for the second round of focus groups were also finalized at this time. The project manager and contacts from the NAC assisted the test developers in making these arrangements. It was very beneficial for the test-developers to have the project manager take on the responsibility for arranging focus groups. The location, refreshments, list of participants, and an appropriate time were all arranged from a distance. All these arrangements required a lot of ongoing communication with the contacts at the site. It reduced the stress and workload of the test developers when the project manager handled all the logistical details of the arrangements.

For Lessons learned during Stage Four, see the How-To Manual.

A second interim report on Stage Three and Four (similar to the first interim report on Stage One and Two) was provided to the NAC at the end of this period.

Stage Five: Pilot Testing of Draft Two of CELBAN

Pilot-testing Draft Two (Pilot B) of the assessment tool was conducted during Stage Five. It was again an informative experience, providing the test-developers with more qualitative and quantitative data on which to base revisions. At this point in the process, more fine-tuning of the assessment tool resulted from the pilot-testing. For the CELBAN, the goal was to pilot-test Draft Two with candidates similar to those tested during Pilot A:

- 80 L2 internationally-educated professionals in the field (in the process of having credentials recognized)
- 40 L1 students who had completed some field experience in the profession
- 5 L1 & L2 professionals presently practicing (within 5 years) in Canada
- Total number of candidates: 125

Pilot-testing of Draft Two was conducted in:

- Ottawa (Algonquin College)

- Edmonton (Grant MacEwan College)
- Calgary (Grant MacEwan College), and
- Vancouver (Vancouver Community College)

A total of 144 candidates were tested.

Note: See "Stage Three: Pilot Testing of Draft One of the CELBAN" on page 12, for explanations of the following components of Stage Five where the same methodology was followed.

Lessons learned during Stage Five, see the How-To Manual.

The following was done at this Stage:

1. **Draft Two was administered and scored.** (During Pilot B, other assessors were trained at Red River College and Algonquin College; also, ESL staff who were familiar with the CLB reviewed the CELBAN and provided helpful feedback at all locations. A draft of an assessment guide for the CELBAN Speaking Test was developed for training purposes.)

2. The L2 internationally-educated nurses were tested using the CLBPT.

3. The data was analyzed.

4. Focus groups were held with stakeholders in the cities in which the pilot-testing was done.

5. An update was provided to the NAC and project manager regarding activities of Stage Five.

Stage Six: Development of the Final Draft of CELBAN and Final Reporting

In this last stage, an official name for the assessment tool was finalized. Suggestions were made by the test developers with input from CCLB Nursing Committee. The Nursing Committee made the final decision. The official name is being registered. Copyright belonging to the CCLB had been clearly established as part of the contract for developing the assessment tool.

Final revisions to the assessment tool were made once the results of the pilot-testing had been analyzed. These revisions included content, time-frames, and instructions for administration, as well as the final test specifications document. This test specifications guide contains a description of the assessment tool, and a general guide to assessors for administering and scoring the speaking and writing assessments. However, f*ormal training of assessors* will be a major component of Phase III*, Implementation of the CELBAN).* The final version of the assessment tool was written, printed, and delivered to the CCLB as per the timeline and in the manner contracted. The Final Version of the CELBAN included the following:

- Listening: Video, Test Booklet, Answer Sheet, Answer Key
- Reading: Test Booklet, Answer Sheet, Answer Key
- Writing: Video, Test Booklet, Answer Sheet, Scoring Grids and Guide
- Speaking: Scoring Grid, Evaluation Sheet, and Guide for Assessors

A general summary of the CELBAN was made available to stakeholders (**see** **"**Appendix H - Canadian Language Benchmarks Assessment For Nurses In Canada (Celban)"** on page 60).

For Lessons learned during Stage Six, see the How-To Manual.

Final reports as per the contract were written and provided to the CCLB and NAC members. (It was necessary to extend the time frame by two weeks because of unanticipated delays.). These reports included the following documents:

- Report on Phase II, Development of the Canadian English Language Assessement for Nurses (CELBAN) (including recommendations for Phase III, Implementation of the CELBAN).

- CELBAN Test Specifications

- "How-To" Manual

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

4 - Project Results

Demographics and Areas Analyzed

Raw data collected during pilot-testing was statistically analyzed. Pilot-testing was conducted on the two drafts of the assessment tool. The combined total of candidates tested during pilot-testing was **270**. The test population was composed of the following:

- 163 internationally-educated nurses
- 98 nursing students (L1 and L2)
- 9 newly-practising nurses (L1 & L2)
- TOTAL: 270 candidates

The following statistical summary represents Pilot B. Data from each of the 4 components of the test (Listening, Speaking, Reading, Writing) was analyzed separately. This analysis included the following:

- demographics
- reliability Alphas
- frequency of distracter selection (CONFIDENTIAL)
- difficulty values
- item discrimination (CONFIDENTIAL)
- comments on validity
- inter-rater reliability (on speaking and writing test components)
- correlation with CLBPT

Feedback from the test candidates was provided to the test developers after the candidates completed each component of the test. This data was also analyzed. The analysis of candidates' feedback included their comments on the following:

- difficulty of test
- length of test
- time allowed for test
- familiarity of content / nursing content
- multiple choice format

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

4 - Project Results ● 17

- chart format
- overall impression of test

Statistical Summary for Pilot B: Data Collection: May – June 2003

Demographics

- Pilot B included 144 participants, who completed at least one section of the test.
- The demographic characteristics of these 144 participants are outlined below.

How to read the tables:

- We'll use the table titled '**Gender**' to describe the meaning of the numbers.
- The title of the table, in this case **Gender**, shows which variable is being described.
- The first column organizes the table. It sorts the data twice, first into valid or missing cases and secondly by the different levels of the variable in question, e.g., male and female.
- The second column is simply a frequency count of how many of the participants fall into the category described in the first column. For example, the '8' under the frequency column means the there were 8 valid, male participants. Similarly there were 126 valid, female cases. The term valid simply means that these participants indicated their gender. Thus, there were 134 valid cases of gender, 8 male and 126 female. There were 10 participants who failed to indicate their gender (missing), but completed at least one part of the test, for a total of 144 participants.
- The percent column calculates the percent based on the total number of participants (valid plus missing). For males and females, the percent column is calculated as follows; percent males = {(8/144) X 100} à 5.6 %; percent females = {(126/144) X 100} à 87.5%.
- The valid percent column uses the only the valid participant total (134) rather than the total number of participants (144). Thus, valid percent males = {(8/134) X 100} à 6%; valid percent females = {(126/134) X 100} à 94%.
- The cumulative percent column gives a running total of the valid percent column (6, 6 + 94 = 100).

GENDER

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	8	5.6	6.0	6.0
	female	126	87.5	94.0	100.0
	Total	134	93.1	100.0	
Missing	System	10	6.9		
Total		144	100.0		

Table 4.1: Gender

Were you educated as a nurse in a country other than Canada

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	76	52.8	53.5	53.5
	no	66	45.8	46.5	100.0
	Total	142	98.6	100.0	
Missing	System	2	1.4		
Total		144	100.0		

Table 4.2: Where Subjects Studied Nursing

Have you passed the Canadian Registered Nurse Exam

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	3	2.1	2.8	2.8
	no	103	71.5	97.2	100.0
	Total	106	73.6	100.0	
Missing	System	38	26.4		
Total		144	100.0		

Table 4.3: Number of Subjects Who Passed the CRNE

Country of Origin

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Canada	44	30.6	30.6	30.6
	USA	1	.7	.7	31.3
	West Africa	1	.7	.7	31.9
	Trinidad	1	.7	.7	32.6
	St. Lucia	1	.7	.7	33.3
	Nigeria	1	.7	.7	34.0
	Romania	4	2.8	2.8	36.8
	Moldova	1	.7	.7	37.5
	Ukraine	1	.7	.7	38.2
	Cambodia	1	.7	.7	38.9
	Philippines	34	23.6	23.6	62.5
	Ethiopia	1	.7	.7	63.2
	Bosnia	2	1.4	1.4	64.6
	Pakistan	2	1.4	1.4	66.0
	Belarus	1	.7	.7	66.7
	Russia	4	2.8	2.8	69.4
	Taiwan	1	.7	.7	70.1
	Poland	3	2.1	2.1	72.2
	Uganda	1	.7	.7	72.9
	Brazil	1	.7	.7	73.6
	Venezuela	2	1.4	1.4	75.0
	Somalia	2	1.4	1.4	76.4
	Haiti	1	.7	.7	77.1
	Japan	1	.7	.7	77.8
	Burundi	1	.7	.7	78.5
	Vietnam	1	.7	.7	79.2
	China	18	12.5	12.5	91.7
	Lebanon	1	.7	.7	92.4
	Iran	3	2.1	2.1	94.4
	Korea	2	1.4	1.4	95.8
	India	2	1.4	1.4	97.2
	Yugoslavia	1	.7	.7	97.9
	Germany	1	.7	.7	98.6
	Peru	1	.7	.7	99.3
	Iraq	1	.7	.7	100.0
	Total	144	100.0	100.0	

Table 4.4: Country of Origin

- Finally, there were some open-ended demographic variables that cannot simply be listed in a frequency table. These variables are outlined in the "descriptive statistics" table below. The first column indicates the question asked of the participant, the 'N' column indicates the number of participants who answered the question, the minimum and maximum columns indicate the lowest and highest value given by the participants, and the mean (the average value of the distribution of scores on that particular question) and std. deviation (the average amount the scores in the distribution vary from the mean) columns provide a measure of central tendency of the responses given for each question.

- The valid N (listwise) indicates that 51 subjects did not answer any of these questions. This group of 51 is made up of mostly Canadian born participants.

	N	Minimum	Maximum	Mean	Std. Deviation
How long have you lived in Canada (months)	82	2.00	359.00	64.0549	58.81599
Education you have had in Canada (months)	63	.50	72.00	14.4802	13.48456
Work experience you have had in Canada (months)	66	1.00	359.00	48.9924	55.79582
Valid N (listwise)	51				

Table 4.5: Descriptive Statistics Summary

Speaking

- A total of 91 participants were measured on the speaking test. Each participant was measured by 2 of a total of 7 raters.

- The speaking test is comprised of 9 categories, each with 6 levels (level 5 to level 10). The level attained by each participant on each category represents the participant's score on that category; the one exception being the 'Intelligibility' category where the participant's score is multiplied by 2. Thus, the total maximum score attainable on the speaking test is 100. Each score is then divided by 10 and the resulting value is the participant's CLB level for speaking.

- The overall agreement or inter-rater reliability for the final CLB level on the speaking test was 0.944.

- The following table indicates 1) the frequency of assignment of each level on the 9 categories of the speaking test by each rater and 2) the inter-rater reliability on each category. For example, the value '3' located under R1 and across from Use of Language indicates that Rater 1 (R1) assigned a level of 5 on the Use of Language category to 3 participants. Rater 2 (R2) assigned a level 5 on the Use of Language to 2 participants. The inter-rater reliability on the Use of Language test was 0.882.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

4 - Project Results ● 21

Score ↓	Level 5		Level 6		Level 7		Level 8		Level 9		Level 10		Inter-Rater Reliability
Test Category 1	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2	
Use of Language	3.	2	10	11	22	29	12	9	27	26	17	14	0.882
Intelligibility	3	1	11	11	24	26	9	8	34	29	10	16	0.894
Organization	1	0	10	9	19	13	10	9	34	45	17	15	0.843
Fluency	5	3	10	17	28	23	12	8	22	28	14	12	0.884
Use of Cohesive Devices	2	2	11	13	16	24	12	9	38	27	12	16	0.886
Vocabulary	5	2	11	11	19	29	10	10	35	23	11	16	0.852
Grammar	4	2	18	13	19	34	11	9	32	22	7	11	0.846
Use of Strategies	0	0	11	13	20	29	11	9	34	27	15	13	0.866
Characteristics of Interaction	2	1	8	8	27	32	10	8	31	25	13	17	0.870

Table 4.6: Inter-Rater Reliability

Listening

- All 144 participants completed the listening test and were included in the analyses.

- The listening component of the test is comprised of 84 questions.

- The listening component is divided into two main tasks, 1) answering 58 multiple choice questions based on five videos, and 2) answering 26 multiple choice questions based on four audio clips.

  - The reliability alphas for the overall test and component tasks are as follows:

    - Reliability for overall listening test: 0.9395

    - Reliability for the video component: 0.9114

    - Reliability for the auditory component: 0.8509

Reading

- Of the total of 144 participants, 121 subjects completed both the reading and skim-scan components of the test. An additional 16 participants completed the skim-scan section only – thus a total of 137 participants completed the skim-scan.

- The reading and skim-scan component consist of 73 questions broken down as follows:

- Parts 1, 2, 3 : Reading Comprehension: 22 questions (Items 1 – 22)

- Part 4 : Cloze: Fill-in-the-blanks   : 35 questions (Items 23 – 57)
- Part 5 : Skim – Scan   : 16 questions (Items S1 – S16)
- TOTAL   73 questions

- The reliability alphas for the overall and component tasks are as follows:

Reliability	Alpha Score
Reliability for overall reading test (all 73 items)	0.9653
Reliability for the reading comprehension (22 items)	0.9209
Reliability for the cloze (35 items)	0.9619
Reliability for skim-scan (16 items)	0.7809
Reliability for cloze and reading comprehension (57 items)	0.9708

Table 4.7: Reliability Alphas

Writing

A total of 143 participants were measured on the writing test. Each participant was measured on 2 tasks, A and B.

1. Task A consisted of an evaluation of 5 writing skills (spelling and legibility, sections completed, main points, supporting details, and point form). Each participant was evaluated on a 4 point scale (level 5 to level 8) and was evaluated by only 1 rater.

2. Task B consisted of an evaluation of 4 writing skills (effectiveness, control of grammar, discourse/fluency, and vocabulary). Each participant was evaluated on a 4 point scale (level 5 to level 8) and was evaluated by 2 raters.

The table below shows the frequency of selection of each level on each writing skill on Task A.

Writing Skill Task A ↓	Level 5	Level 6	Level 7	Level 8	Missing
Spelling and Legibility	9	9	42	83	0
Sections Completed	1	2	21	60	59
Main Points	39	20	28	56	0
Supporting Details	38	38	41	22	4
Point Form	2	6	26	104	5

Table 4.8: Frequency of Selection on Task A

Note: The 59, 4, and 5 missing cases resulted from regrouping the five criteria into two categories.

- The inter-rater reliability for total scores on Task B was 0.905.

- The table below indicates 1) the frequency of assignment of each level on the 4 writing skills of Task B by each rater, and 2) the inter-rater reliability on each writing skill. For example, the value '19' located under R1 and across from Effectiveness indicates that Rater 1 (R1) assigned a level of 5 on the Effectiveness skill to 19 participants. Rater 2 (R2) assigned a level 5 on the Effectiveness skill to 11 participants. The inter-rater reliability on the Effectiveness skill was 0.882.

Score ↓	Level 5		Level 6		Level 7		Level 8		Inter-Rater Reliability
Writing Skills Task B ↓	R1	R2	R1	R2	R1	R2	R1	R2	
Effectiveness	19	11	46	43	53	43	35	36	0.843
Control of Grammar	17	18	55	56	3	35	32	34	0.834
Discourse & Fluency	15	13	35	46	50	43	41	43	0.849
Vocabulary	19	15	33	40	48	48	43	40	0.801

Table 4.9: Frequency of Assignment and Inter-Rater Reliability for Task B

Validity

Validity is evidence that a test is being used appropriately and measures what it sets out to measure. There are 4 major types of validity:

1. **Content validity** (the extent to which the questions on a test are representative of the behaviour that is being measured).

2. **Construct validity** (the extent to which a test measures some theoretical construct)

3. **Criterion-related validity** (the extent to which the scores on a test correlate with scores on some other measure or behaviour).

4. **Face validity** (how test takers perceive the attractiveness and appropriateness of a test).

It is important to understand the different types of validity and when they should be used.

Content Validity

There are no statistical procedures to show that content validity exists. Thus, how do you go about obtaining evidence of content validity?

Content validity is typically established before the test is administered. The generally accepted procedure involves defining the testing universe (the sample of all possible behaviours of the attribute being measured) and developing questions that

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

map onto the particular testing universe. Once developed, the questions are rated by experts who determine the appropriateness of each question to the test universe.

In the case of this test for nurses (CELBAN), both steps (mapping questions to he testing universe, and expert rating of items) were carried out and thus we can be confident that the test is content-valid and representatively samples the nursing language, testing universe.

Construct Validity

The process of establishing construct validity for a test is a somewhat tedious, and requires the gradual accumulation of evidence that illustrates that the test's "test scores" relate to observable behaviours in such a way that they were predicted by the underlying theory.

While there are many different methods that can be used to provide evidence of construct validity, two of the most common methods involve:

1. Correlating the test with other, established tests that measure a similar construct.

2. Showing that different populations of participants, who theoretically should perform differently on the test, do perform differently on the test.

Let's consider these two methods with the CELBAN.

Method 1. Correlation With Other Test: Benchmark Comparisons

A total of 78 participants from Pilot II also had scores for the CLBPT (Canadian Language Benchmarks). The following table shows the correlation between the current CELBAN test and CLBPT levels for each assessment area (i.e., Speaking, Listening, Reading, and Writing). The bolded values along the diagonal represent like assessment areas. The n = represents the number of participants contributing to the correlation.

					CELBAN
		Listening	Reading	Speaking	Writing
CLBPT	Listening	**0.495**	0.444	0.607	0.327
		n = 61	n = 61	n = 61	n = 61
	Reading	0.406	**0.625**	0.281	0.438
		n = 74	n = 74	n = 73	n = 74
	Speaking	0.687	0.484	**0.741**	0.541
		n = 61	n = 61	n = 61	n = 61
	Writing	0.376	0.417	0.342	**0.533**
		n = 74	n = 74	n = 73	n = 74

Table 4.10: Correlation between CLBPT and CELBAN

From a psychometric point of view you would expect to see the highest correlations along the diagonal, between tests that supposedly measure the same thing. The results above do indicate moderate convergent validity since two of the highest correlations, and three of the top six correlations lie along the diagonal.

Method 2. Group Comparisons

Since the test was developed in Canada and based on Canadian English nursing language then it is clear that groups of participants not trained in Canada should be at a disadvantage. Persons then who indicated that they received their nursing training outside Canada should perform below those trained in Canada.

The following table indicates the frequency of nurses either trained in or outside Canada.

Were you educated as a nurse in a country other than Canada

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	76	52.8	53.5	53.5
	no	66	45.8	46.5	100.0
	Total	142	98.6	100.0	
Missing	System	2	1.4		
Total		144	100.0		

Table 4.11: Frequency of Nurses Trained Outside of Canada

Test Component Scores

The following graphs illustrate the differences in scores on each component of the test.

Note: All differences described below are significant at the alpha = 0.05 level, providing evidence of construct validity.

Listening

Figure 4.1: Average Values on Listening Test

Reading

Average Values on Reading Test Components
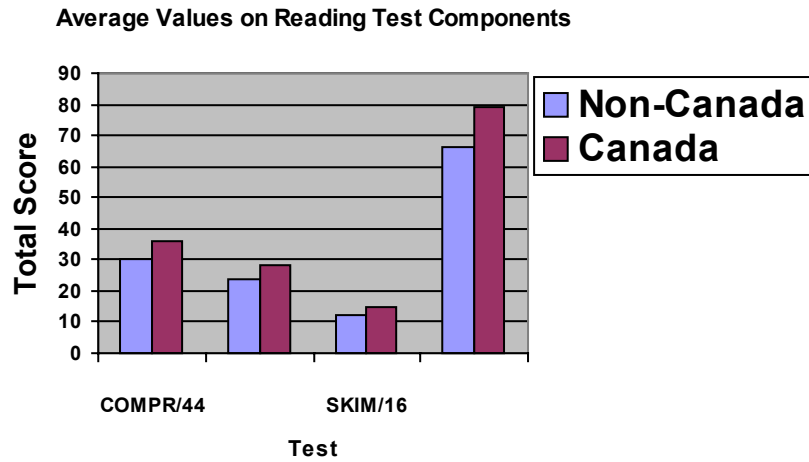
Figure 4.2: Average Values on Reading Test Components

Speaking

Note: All criteria are /10 except Intelligibility, which is /20.

Average Values on Speaking Test
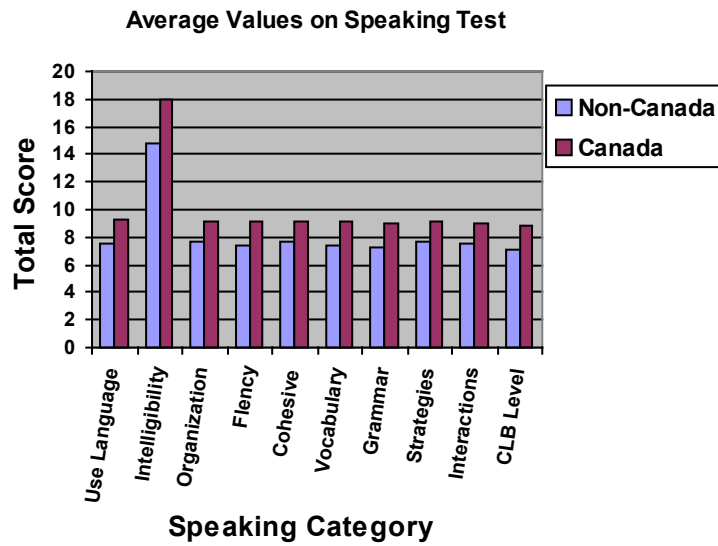
Figure 4.3: Average Values on Speaking Test

Writing

Note: All criteria are /8.

Average Values on Writing Test: Task B and Final CLB Scores
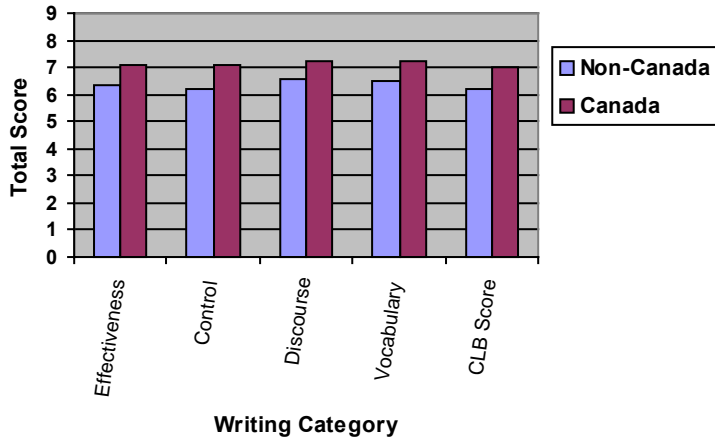
Figure 4.4: Average Values on Writing Test

Face Validity

Face validity involves the test takers perception of the test, specifically if the test measures what it is suppose to measure.

After each section of the test, the participants (test-takers) were asked a series of questions about the test. The results from this assessment are shown below. The question appears at the top of each table and the frequency of responses for each question in the table.

Overall, the data indicate that the majority of participants believed the test was valid.

Listening

The questions were

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	13	9.0	9.1	9.1
	just right	120	83.3	83.9	93.0
	too hard	10	6.9	7.0	100.0
	Total	143	99.3	100.0	
Missing	System	1	.7		
Total		144	100.0		

Table 4.12: Question Difficulty in Listening Test

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

The test length was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too short	31	21.5	22.1	22.1
	just right	88	61.1	62.9	85.0
	too long	21	14.6	15.0	100.0
	Total	140	97.2	100.0	
Missing	System	4	2.8		
Total		144	100.0		

Table 4.13: Length of Listening Test

The time allowed for the test was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too much	5	3.5	3.5	3.5
	enough	78	54.2	54.2	57.6
	not enough	61	42.4	42.4	100.0
	Total	144	100.0	100.0	

Table 4.14: Time Allotted to Do Listening Test

Was content familiar?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very familiar	79	54.9	55.6	55.6
	somwhat familiar	50	34.7	35.2	90.8
	not familiar	13	9.0	9.2	100.0
	Total	142	98.6	100.0	
Missing	System	2	1.4		
Total		144	100.0		

Table 4.15: Familiarity of Content in Listening Test

What did you think of nursing content?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	140	97.2	99.3	99.3
	not effective	1	.7	.7	100.0
	Total	141	97.9	100.0	
Missing	System	3	2.1		
Total		144	100.0		

Table 4.16: Effectiveness of Nursing Content in Listening Test

MC = Multiple Choice

What did you think of MC format?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	133	92.4	94.3	94.3
	not effective	8	5.6	5.7	100.0
	Total	141	97.9	100.0	
Missing	System	3	2.1		
Total		144	100.0		

Table 4.17: Effectiveness of MC Format in Listening Test

What did you think of chart format?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	133	92.4	95.0	95.0
	not effective	6	4.2	4.3	99.3
	3.00	1	.7	.7	100.0
	Total	140	97.2	100.0	
Missing	System	4	2.8		
Total		144	100.0		

Table 4.18: Effectiveness of Chart Format in Listening Test

What did you think of video section?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	126	87.5	91.3	91.3
	not effective	12	8.3	8.7	100.0
	Total	138	95.8	100.0	
Missing	System	6	4.2		
Total		144	100.0		

Table 4.19: Effectiveness of Video Section in Listening Test

What did you think of audio section?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	126	87.5	90.0	90.0
	not effective	14	9.7	10.0	100.0
	Total	140	97.2	100.0	
Missing	System	4	2.8		
Total		144	100.0		

Table 4.20: Effectiveness of Audio Section in Listening Test

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Overall impression of listening?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very good	39	27.1	27.5	27.5
	good	71	49.3	50.0	77.5
	fair	32	22.2	22.5	100.0
	Total	142	98.6	100.0	
Missing	System	2	1.4		
Total		144	100.0		

Table 4.21: Overall Impressions of Listening Section in Listening Test

Reading Test

The questions were

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	6	5.3	5.3	5.3
	just right	98	86.7	86.7	92.0
	too hard	9	8.0	8.0	100.0
	Total	113	100.0	100.0	

Table 4.22: Difficulty of Reading Questions

The length of the test

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too short	6	5.3	5.3	5.3
	just right	99	87.6	87.6	92.9
	too long	8	7.1	7.1	100.0
	Total	113	100.0	100.0	

Table 4.23: Length of Reading Test

Time allowed for reading comprehension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too much	5	4.4	4.4	4.4
	enough	90	79.6	79.6	84.1
	not enough	18	15.9	15.9	100.0
	Total	113	100.0	100.0	

Table 4.24: Time Allowance for Reading Comprehension

Content was familiar

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very familiar	35	31.0	31.0	31.0
	somewhat familiar	67	59.3	59.3	90.3
	not familiar	11	9.7	9.7	100.0
	Total	113	100.0	100.0	

Table 4.25: Familiarity of Reading Content

Think of nursing content

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	110	97.3	100.0	100.0
Missing	System	3	2.7		
Total		113	100.0		

Table 4.26: Effectiveness of Nursing Content in Reading Test

Think of MC format

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	105	92.9	95.5	95.5
	not effective	5	4.4	4.5	100.0
	Total	110	97.3	100.0	
Missing	System	3	2.7		
Total		113	100.0		

Table 4.27: Effectiveness of MC in Reading Test

Think of Fill-In-Blanks

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	100	88.5	92.6	92.6
	not effective	8	7.1	7.4	100.0
	Total	108	95.6	100.0	
Missing	System	5	4.4		
Total		113	100.0		

Table 4.28: Effectiveness of Fill-In Blanks in Reading Test

Overall impression of read comp

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very good	35	31.0	31.8	31.8
	good	60	53.1	54.5	86.4
	fair	15	13.3	13.6	100.0
	Total	110	97.3	100.0	
Missing	System	3	2.7		
Total		113	100.0		

Table 4. 29: Overall Impression of Reading Comprehension

Skim and Scan refers to the reading method where a person quickly looks for specific information in a text, for example,. in a phone book.

Skim and Scan Section

The questions were

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	15	11.0	11.0	11.0
	just right	112	82.4	82.4	93.4
	too hard	9	6.6	6.6	100.0
	Total	136	100.0	100.0	

Table 4.30: Difficulty of Skim & Scan Questions

The length of test was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too short	25	18.4	18.5	18.5
	just right	102	75.0	75.6	94.1
	too long	8	5.9	5.9	100.0
	Total	135	99.3	100.0	
Missing	System	1	.7		
Total		136	100.0		

Table 4.31: Length of Skim & Scan Test

Time allowed skim-scan

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too much	5	3.7	3.7	3.7
	enough	82	60.3	60.7	64.4
	not enough	48	35.3	35.6	100.0
	Total	135	99.3	100.0	
Missing	System	1	.7		
Total		136	100.0		

Table 4.32: Time Allowance for Skim & Scan Test

Was the content familiar?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very familiar	62	45.6	46.3	46.3
	somewhat familiar	62	45.6	46.3	92.5
	not familiar	10	7.4	7.5	100.0
	Total	134	98.5	100.0	
Missing	System	2	1.5		
Total		136	100.0		

Table 4.33: Familiarity of Skim & Scan Test Content

Think of nursing content

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	128	94.1	100.0	100.0
Missing	System	8	5.9		
Total		136	100.0		

Table 4.34: Effectiveness of Nursing Content in Skim & Scan Test

Think of quest and answer format

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	124	91.2	96.9	96.9
	not effective	4	2.9	3.1	100.0
	Total	128	94.1	100.0	
Missing	System	8	5.9		
Total		136	100.0		

Table 4.35: Effectiveness of Q & A format in Skim & Scan Test

Think of Part 1: patient infor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	126	92.6	99.2	99.2
	not effective	1	.7	.8	100.0
	Total	127	93.4	100.0	
Missing	System	9	6.6		
Total		136	100.0		

Table 4.36: Effectiveness of Part 1of Skim & Scan Test

Think of Part 2: Progress notes

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	effective	122	89.7	96.8	96.8
	not effective	4	2.9	3.2	100.0
	Total	126	92.6	100.0	
Missing	System	10	7.4		
Total		136	100.0		

Table 4.37: Effectiveness of Part 2 of Skim & Scan Test

Overall impression Skim-Scan

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very good	46	33.8	34.6	34.6
	good	65	47.8	48.9	83.5
	fair	22	16.2	16.5	100.0
	Total	133	97.8	100.0	
Missing	System	3	2.2		
Total		136	100.0		

Table 4.38: Overall Impression of Skim & Scan Test

Speaking Test

Overall impression of speaking

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very good	47	57.3	58.0	58.0
	good	30	36.6	37.0	95.1
	fair	4	4.9	4.9	100.0
	Total	81	98.8	100.0	
Missing	System	1	1.2		
Total		82	100.0		

Table 4.39: Overall Impression of Speaking Test

The length of the test was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	just right	82	100.0	100.0	100.0

Table 4.40: Length of Speaking Test

Was content familiar?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very familiar	40	48.8	48.8	48.8
	familiar	39	47.6	47.6	96.3
	not familiar	3	3.7	3.7	100.0
	Total	82	100.0	100.0	

Table 4.41: Familiarity of Speaking Test Content

Do speaking task reflect nursing?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	yes	74	90.2	96.1	96.1
	no	3	3.7	3.9	100.0
	Total	77	93.9	100.0	
Missing	System	5	6.1		
Total		82	100.0		

Table 4.42: If Speaking Test Tasks Reflect Actual Nursing Speaking Tasks

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Role Play 1 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	5	6.1	6.1	6.1
	just right	77	93.9	93.9	100.0
	Total	82	100.0	100.0	

Table 4.43: Difficulty of Role Play #1

Role Play 2 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	2	2.4	2.4	2.4
	just right	78	95.1	95.1	97.6
	too hard	2	2.4	2.4	100.0
	Total	82	100.0	100.0	

Table 4.44: Difficulty of Role Play #2

The questions were

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	4	4.9	5.0	5.0
	just right	74	90.2	92.5	97.5
	too hard	2	2.4	2.5	100.0
	Total	80	97.6	100.0	
Missing	System	2	2.4		
Total		82	100.0		

Table 4.45: Difficulty of Questions in Speaking Test

The assessors made me feel

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	comfortable	79	96.3	97.5	97.5
	uncomfortable	2	2.4	2.5	100.0
	Total	81	98.8	100.0	
Missing	System	1	1.2		
Total		82	100.0		

Table 4.46: Comfort Level During the Speaking Test

Writing Test

Was the content familiar?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very familiar	74	52.1	52.5	52.5
	somewhat familiar	56	39.4	39.7	92.2
	not familiar	11	7.7	7.8	100.0
	Total	141	99.3	100.0	
Missing	System	1	.7		
Total		142	100.0		

Table 4.47: Familiarity of Content in the Writing Test

My overall impression of writing test is

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very good	37	26.1	27.0	27.0
	good	75	52.8	54.7	81.8
	fair	25	17.6	18.2	100.0
	Total	137	96.5	100.0	
Missing	System	5	3.5		
Total		142	100.0		

Table 4.48: Overall Impressions of Writing Test

The form was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	7	4.9	5.0	5.0
	just right	128	90.1	90.8	95.7
	too hard	6	4.2	4.3	100.0
	Total	141	99.3	100.0	
Missing	System	1	.7		
Total		142	100.0		

Table 4.49: Difficulty of the Writing Test Form

The length of Task 1 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too short	15	10.6	10.8	10.8
	just right	122	85.9	87.8	98.6
	too long	2	1.4	1.4	100.0
	Total	139	97.9	100.0	
Missing	System	3	2.1		
Total		142	100.0		

Table 4.50: Length of Writing Task #1

The time allowed for Task 1 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	enough	111	78.2	78.2	78.2
	not enough	31	21.8	21.8	100.0
	Total	142	100.0	100.0	

Table 4.51: Time Allowance for Writing Task #1

The report was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	5	3.5	3.5	3.5
	just right	130	91.5	91.5	95.1
	too hard	7	4.9	4.9	100.0
	Total	142	100.0	100.0	

Table 4.52: Difficulty of the Report in the Writing Test

The length of task 2 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too easy	12	8.5	8.5	8.5
	just right	125	88.0	88.7	97.2
	too hard	4	2.8	2.8	100.0
	Total	141	99.3	100.0	
Missing	System	1	.7		
Total		142	100.0		

Table 4.53: Length of Writing Task #2

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

4 - Project Results ● 39

The time allowed for Task 2 was

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	too much	9	6.3	6.5	6.5
	enough	94	66.2	67.6	74.1
	not enough	36	25.4	25.9	100.0
	Total	139	97.9	100.0	
Missing	System	3	2.1		
Total		142	100.0		

Table 4.54: Time Allowance for Writing Task #2

Note: After pilot-testing was completed, it was discovered that Part One of the Reading Comprehension Test was not secure. This passage was originally purchased from the consultants at the CanTEST Project Office. They replaced this passage (Passage A) with a new passage (Passage B). Passage B was pilot-tested with 618 L2 students at the University of Ottawa. Data on this passages was analyzed by the CanTEST Project Office.

The two passages were compared in the following ways:

1. Mean comparison

   - Passage A (total testing population, both L1 and L2):
        Mean .68

   - Passage A (international only)   Mean .64

   - Passage B (total population; all L2) Mean: .64

2. The two passages were completed by 23 internationally-educated nurses at Red River College (one additional nurse completed only Passage A), to confirm that the level of difficulty was comparable. The results were:

   - The average percentage correct, 75%, was identical for the two passages.

   - The variances were similar with standard deviation for Passage A=1.16 and standard deviation for Passage B=1.31.

Based on recommendation from the CanTEST Project Office and the statistician, these comparisons indicate that the two passages are interchangeable as subsections of the Reading Comprehension Test.

NOTE: Ongoing statistical analysis is necessary to track the success rate (post-test) of internationally-educated nurses through performance evaluations conducted several months after they have obtained and retained nursing positions. Performance evaluations conducted on these nurses would also lend support for the validity of the assessment tool, if the 'performance indicators' used in the evaluation are congruent with the test construct and the performance indicators used in the tool. To be most accurate, the evaluation would need to be designed by the test developers (using a set of pre-determined criteria to ensure this congruency), and conducted by an unbiased individual. It is desirable to obtain longitudinal data of this type as concrete evidence of the validity of the test.

40 ● 4 - Project Results

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

In addition to the above formal statistical analysis, feedback received during focus groups was recorded (see Recommendations for Implementation, of this document).

5 - Phase II Conclusions

Conclusions

The end result of the project has been the development of an English language proficiency assessment tool for internationally-educated nurses. The statistics indicate the reliability and validity of the CELBAN. In addition, the feedback received from candidates and colleagues during pilot testing was extremely positive. Candidates were very satisfied with the CELBAN and anxious to hear *when* the CELBAN would be approved and available. They said that CELBAN more appropriately measured their English proficiency skills in a nursing context than any other test they had previously taken. They believed that CELBAN provided an excellent alternative as an assessment tool to determine their threshold English language proficiency for entering the nursing profession in Canada.

The importance of Phase I report, *An Analysis of the English Language Demands of the Nursing Profession Across Canada***, cannot be overstated in developing an occupation-specific language proficiency assessment tool.** The test-developers continually referred back to this valuable resource throughout the development process.

The involvement and commitment of stakeholders, including regulatory bodies, is essential to the process from the outset. Also, the e**stablishment of a team of experts** (linguists, second language test & measurement consultants, statisticians, test reviewers, nursing consultants) to assist on the test development team is crucial. In addition, a **National Advisory Committee** is important as part of the process of development of an occupation-specific language assessment tool.

Based on the experience with this project, it is suggested that a minimum of 12 months be provided for the development of an occupation-specific language assessment tool. This provides adequate time to secure funding, obtain ethics approval, and create and execute an appropriate, well-designed plan, *including* pilot testing. It is important to identify pilot testing sites which are representative of the target population across Canada. The development time would *follow* a 6-month time period for the analysis of the English language demands of the profession, (which has clearly been established as necessary groundwork for the development of the tool).

Other professions face the same issues regarding assessing the Engllish language proficiency of internationally-educated candidates. **The development of the CELBAN provides a model for other professions as they seek to address these concerns.**

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

6 - Recommendations for CELBAN

Recommendations for Implementation

Phase III, *Implementation of the CELBAN* will be the next step in the process. It will be an extensive and crucial phase. It must be well-organized and methodically executed, with buy-in from all stakeholders before, during and after the process. It should include the set-up of an administrative center, (central location for data collection and storage), set-up of assessment centre(s), training of assessors, writing subsequent versions of the assessment tool, and publicizing/marketing the CELBAN. Funding must be in place for start-up costs, with incremental funding provided during the initial implementation time (the first few years) until the "business" of conducting CELBAN assessments is well-established.

The following issues/concerns regarding implementation were expressed by participants at focus groups held in six cities in response to the following question: "What are the issues / concerns regarding implementation?" Comments and recommendations by the test developers follow each concern outlined.

Purpose of CELBAN

The mandate of CELBAN needs to be clearly articulated. This is NOT a placement test for entry/exit in programs, but a proficiency assessment for internationally-educated nurses to demonstrate threshold competency necessary for licensure as part of requirements for accessing the profession).

Recommendation #1

All educational institutes, licensing bodies, employers, immigrant referral agencies, and other stakeholders need to have a clear understanding of what the CELBAN is, and is NOT, to be used for. This information needs to be clearly explained to internationally-educated nurses seeking direction. The brochure designed by the CCLB, which introduces the CELBAN, is an excellent starting point. Updating the newly-established CELBAN web-site, and creating links to nursing websites, immigration websites, etc. would also be useful in providing clarification and information to interested parties about the mandate of the CELBAN.

Administration & Quality Control of CELBAN

An area of concern is the establishment of assessment centres – selection, and time-frame for set-up of centres for testing (regional, provincial, etc.). Related to this, is the concern regarding assessor training (i.e. where, when, by whom?). Also, ongoing monitoring and follow-up of assessors is important. Ongoing recording-keeping and statistical analysis, as well as the development of new versions of the CELBAN are security-related issues needing attention. There is also some concern about possible liability issues (i.e. claims of reliability and face validity need to be defensible). An appeal process needs to be in place along with policies regarding re-tests, i.e. re-test of entire test vs. single component failed (e.g. speaking only); time frame before re-testing can occur; validity period for results; record-keeping of success and failure rates of test-takers; time-frame for providing test results and feedback (re: strengths and weaknesses in productive skills – Speaking & Writing) to candidates.

Recommendation #2

The administration of the CELBAN is a very important aspect of implementation. First, establishing an administrative centre for CELBAN is a crucial part of the process. It needs to be centrally-located and accessible to representatives of all relevant stakeholders. Systems for keeping statistical records, storage and retrieval of data, and registering candidates, etc. need to be established. Policies for testing and re-testing and filing appeals need to be developed and clearly articulated. It is advisable to start with one (or two) testing centre(s) so that implementation is done in a controlled and progressive fashion.

Recommendation #3

Training of assessors is the second important step in standardizing the delivery of the CELBAN. Regular monitoring and follow-up of assessors is also recommended.

Recommendation #4

The third step in the administration of the CELBAN is the development of other versions of the assessment tool. To maintain the integrity of the test, it is necessary that several versions be available as soon as possible. The security of the test is a crucial element in developing a high-stakes assessment. Having several versions available assists in maintaining this security.

Cost and Accessibility

The CELBAN must be cost-effective, affordable (perhaps a loan program could be established for test-takers), comparable in price to other similar tests, and accessible.

Recommendation #5

In terms of cost, it must be kept in mind that two trained assessors are required for the Speaking component, which is conducted with each individual candidate. Also, two trained assessors are required to evaluate the Writing component. One trained invigilator can administer and score the Listening and Reading components, as they are multiple choice format with an answer key provided. The complete test takes approximately 2 1/2 hours. The cost should be comparable to tests with similar demands for assessors and candidates. In addition to cost, the CELBAN needs to be accessible for internationally-educated nurses living anywhere in Canada. Regional assessment centres, and/or portable assessors may eventually facilitate this accessibility.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Follow-Up

In addition to providing feedback, perhaps assessors could provide referrals/suggestions for remediation if candidates are unsuccessful.

Recommendation #6

It is advisable to provide information and assistance to candidates who attempt the CELBAN by referring them to appropriate programs if they are unsuccessful in meeting the standards set by the CELBAN. This would best be facilitated by a "circular" not linear process flow chart (see **Process** issues below).

Recommendation #7

It is also important that ongoing research be done in tracking the candidates who successfully access the workforce. Ongoing statistical analysis and tracking of successful test-takers provides a means for assessing "predictive validity" (criterion-related validity) for the assessment tool.

Recommendation #8

Performance evaluations of internationally-educated professionals, in theory, lend support to the criterion-related validity of the assessment tool. However, the 'performance indicators' used in the evaluation need to be congruent with the indicators used in the tool. The evaluation needs to be given by an unbiased individual using a set of pre-determined criteria which relate to the construct. Ideally the evaluation should be designed by the test developers.

Note: There was a suggestion that cross-assessment comparisons be made with other tests (e.g., TOEFL, MELAB). Caution is advised in carrying out this type of comparison, as different tests may not be testing the same language tasks, and a comparison of scores could be questioned in terms of appropriateness.

Information and Publicity

When and *how* will information be available to the public?

Recommendation #9

As each step of implementation is carried out, the CCLB will continue to provide information and promote the CELBAN. Information about the CELBAN has already initially been provided in various formats: a CELBAN brochure, information on the CCLB website, and links to nursing websites. Publicity in nursing journals and professional publications may be an additional source for circulating information.

Development of Other Tests

There is a need for preparatory materials and practice-tests for CELBAN. There is also a need for a "predictor" test for potential candidates prior to arrival in Canada or prior to attempting CELBAN (e.g. CLBA, CLBPT, generic "health practitioner" language proficiency test, etc.) An entry test (separate from CELBAN) for L2 students applying for nursing programs would be very useful for educational institutions assessing entry-level proficiency for L2 nursing students.

Recommendation #10

In Phase I of the project, Benchmarking the Nursing Profession, a wealth of information was obtained from which to design CELBAN preparatory materials and nursing-specific tests. The CLBA and the CLBPT are useful, though generic, proficiency tests which could be used as predictor tests. Some institutions have designed in-house health-related English Proficiency tests for entry to special programs. The need for a standardized general "health practitioner" language proficiency test is an issue needing further investigation.

Provision of Programs

There is a need for provision of programs (i.e. English for Nursing Purposes , or ENP-type programs) to assist candidates in preparing to take CELBAN, or remedial programs for "gap-filling" after candidates have been unsuccessful with CELBAN and had weaknesses identified. It would also be very helpful if financial assistance or loan programs could be made available for candidates taking preparatory program.

Recommendation #11

Again, in Phase I of the project, Benchmarking the Nursing Profession, there was a wealth of information obtained while analyzing the English language demands of the Nursing Profession from which to design an ENP curriculum. Specific modules could be designed to address the gaps identified to candidates after unsuccessfully attempting the CELBAN.

Buy-In From Stakeholders

There is a need for acceptance by institutions and licensing bodies of the legitimacy of the CELBAN. A single national standard of a group of acceptable language proficiency assessment(s) recognized across all jurisdictions is needed. It is also time for a shift to occur from the responsibility being mostly that of the internationally-educated nurse, to a shared responsibility by the "system"(all stakeholders) to assist these nurses to obtain and retain employment. Perhaps advocacy is needed for a process for socialization into the Canadian healthcare "culture".

Recommendation #12

The feasibility study conducted prior to Phase I confirmed the sense that stakeholders agreed there was indeed a need for developing a nursing-specific language assessment tool. That initial support must continue throughout the process of implementation. Un-voiced concerns, unfair biases, or firm reluctance to change the status quo will lead to confusion and chaos during the implementation phase. A smooth process of implementation is preferable for stakeholders, and especially for internationally-educated nurses.

Establishment of "Process"

There is a need to establish a clearly defined process which is transparent to all stakeholders, and especially to internationally-educated nurses. **This process must be a single national process**. There needs to be a connection to recruiters and employers in the process, as well as a built-in support system for nurse-to-nurse

46 • 6 - Recommendations for CELBAN

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

connection. The exact placement of CELBAN (and/or other language tests) in the process needs to be clearly established.

Recommendation #13

A process chart (which has circular, not linear, aspects) needs to be refined, approved, and followed by all stakeholders. This will more readily facilitate the process of assisting internationally-educated nurses to obtain and retain employment as expeditiously as possible.

Recommendation #14

Finally, the following reality was clearly articulated by a member of one of the focus groups:

Not only do internationally-educated nurses represent a multicultural and diverse population, but they also serve clients who represent the same. As such, they need to be regarded as an asset, not a liability.

The bottom line is that everyone, all the stakeholders as well as the general public, benefit when internationally-educated nurses can be integrated into the Canadian workforce as fairly and efficiently as possible.

7 - CELBAN Test Development and Administrative Reference Documents

- Phase I, Benchmarking the Nursing Profession - An Analysis of the English Language Demands of the Nursing Profession Across Canada

- How-To Manual: Developing an Occupation-Specific Language Assessment Tool Using the Canadian Language Benchmarks

- CELBAN General Test Specifications(Confidential)

48 • **7 - CELBAN Test Development and Administrative Reference Documents**

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

8 - Appendices

Appendix A – Bibliography (References)

BOOKS:

Alderson, J. Charles, Clapham, Caroline and Wall, Dianne. Language Test Construction and Evaluation. Cambridge University Press. 1995.

Bachman, Lyle F. Fundamental Considerations in Language Teaching. Oxford University Press. 1990.

Bachman, Lyle F. & Palmer, Adrian S. Language Testing in Practice. Oxford University Press. 1996.

Bailey, Kathleen M. Learning About Language Assessment. Dilemmas, Decisions, and Directions. Newbury House. Heinle & Heinle Publishers, 1998.

Buck, Gary. (Editors - Charles Alderson & Lyle F. Bachman) Assessing Listening. (Cambridge Language Assessment Series.) Cambridge University Press. 2001.

Cushing Weigle, Sara. (Editors - Charles Alderson & Lyle F. Bachman) Assessing Writing. (Cambridge Language Assessment Series.) Cambridge University Press. 2002.

Douglas, Dan. Assessing Languages for Specific Purposes. (Editors - Charles Alderson & Lyle F. Bachman) Assessing Listening. (Cambridge Language Assessment Series.) Cambridge University Press. 2000.

Gronlund, Norman E. How to Make Achievement Tests and Assessments. Allyn & Bacon. 1993.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

8 - Appendices ● 49

Kirk, Jerome & Miller, Marc. <u>Reliability and Validity in Qualitative Research</u>. Sage Publications, 1986.

Long, Michael H. & Richards, Jack C,. Editors. <u>Methodology in TESOL</u>. A Book of Readings. Heinle & Heinle Publishers. 1987 & 2001.

McNamara, Tim. <u>Language Testing</u>. Oxford University Press. 2000.

Millman, Jason. & Green, Jennifer. "The Specification and Development of Tests of Achievement and Ability". (Editor Robert L. Linn) <u>Educational Measurement</u>. Collier Macmillan Publishers.

Patton, Michael Q. <u>Creative Evaluation</u>. Sage Publications, 1981.

Potter, Patricia A. & Perry, Anne Griffin. <u>Basic Nursing</u>. A Critical Thinking Approach. (Four Edition). Mosby. A Harcourt Health Sciences Company. 1999.

Sides, Marian B. Korchek, Nancy. <u>Successful test-taking – Learning Strategies for Nurses</u>. Lippincott. 1998.

Sommers-Flanagan, Rita & John. <u>Clinical Interviewing</u>. John Wiley & Sons, Inc. 1999.

Weir, Cyril. <u>Understanding & Developing Language Tests</u>. Prentice Hall. 1993.

Williams, Diana. <u>Writing in Practice</u>. A Practical Guide for Health Professionals. Jessica Kingsley Publishers Ltd. 2002.

ARTICLES:

Brown, James D. and Hudson, Thom. "*Alternatives in Language Assessment*" TESOL Quarterly. Vol. 32. No. 4 Winter 1998.

Brown, James D. and Hudson, Thom. "*A Focus on Language Development: Expanding the language Proficiency Construct Across a Variety of Tests (Technical Report #21)*" U. of Hawaii, Second Language Teaching & Curriculum Centre. p. xi + 230

Celce-Murcia, Marianne. "*Communicative Competence: A Pedagogically Motivated Model with Content Specifications*". <u>Issues in Applied Linguistics.</u> Vol. 6 No. 2 1995.

50 ● 8 - Appendices

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Kitao, S. Kathleen & Kitao, Kenji. "*Testing Communicative Competence.*" The Internet TESL Journal, Vol. II, No. 5, May 1996.

Kitao, S. Kathleen & Kitao, Kenji. " *The Internet TESL Journal*", Vol. II, No. 7, July 1996.

Kitao, S. Kathleen & Kitao, Kenji. "*Testing Listening*", "*Testing Reading*", "*Test Design I & II*", The Internet TESL Journal.

Lumley, Tom. "*Perceptions of Language-trained Raters and Occupational Experts in a Test of Occupational English Language Proficiency*" ESP. Vol. 17, No. 4. pp.347-367. 1998.

Mendelsohn, D. & Stewart, G. "*Developing the Midwives' language Proficiency Test: A field-specific performance test.*" 1999. ESP Malaysia 6, 11-27.

Norris, John M. "*Identifying Rating Criteria for Task-based EAP Assessment" A Focus on language test development: Expanding the language proficiency construct across a variety of tests*" (Technical Report #21, pp. 163-204). Honolulu: University of Hawaii, Second language Teaching and Curriculum Centre.

Norton, B. & Stewart, G. "*Accountability in Language Assessment: Challenges of Authenticity and Cultural Diversity.*" The Canadian Modern Language Review, 1996: 56, 2, 223-224.

Peirce, B. & Stewart, G. "*The Development of the Canadian Language Benchmarks Assessment.*" TESL Canada Journal. 1995: 14, 2, 17-31.

OTHER:

IELTS (International English Language Testing System), *The IELTS Handbook* 1999.

Carleton University, *CAEL (Canadian Academic English Language Assessment) website*: www.carleton.ca

Nagy, P. –"*An Examination of the Validity of The Canadian Language Benchmarks Placement Test*" (Final report), 2002

Twenty-Five Years of Second Language Teaching at the University of Ottawa. 1995.

Occupational English Test. T.F. McNamara. (descriptive info. - http://ericae.net/tc3/)

Policy Brief - Organization for Economic Co-operation and Development. "*International Mobility of the Highly Skilled.*" July 2002.

Canadian Nurses Association, *The Canadian RN Exam Prep Guide, 3rd edition,* Canadian Nurses Association, 2000.

VIDEOS

Introduction to Physical Assessment Techniques, Medical Training Associates. P.O. Box 833 Moncton, N.B. Canada. EIC 8N6. 1993.

Physical Assessment for Nurses - History Taking., Medical Training Associates. P.O. Box 833 Moncton, N.B. Canada. EIC 8N6. 1993.

Rapid Physical Assessment of the Critically Ill Patient, AIMS Media, 9710 DeSoto Avenue, Chatsworth, CA 91311-4409

Orthopedic Assessment., Canadian Learning Company Inc.   Head Office; 63 Mack Avenue, Scarboro (Toronto), Ontario. M1L 1M5

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Appendix B - Charts

Charts illustrating analysis of observations data.

Chart 1 illustrates the situational use of language by nurses during the observations.

Situational Uses of Language

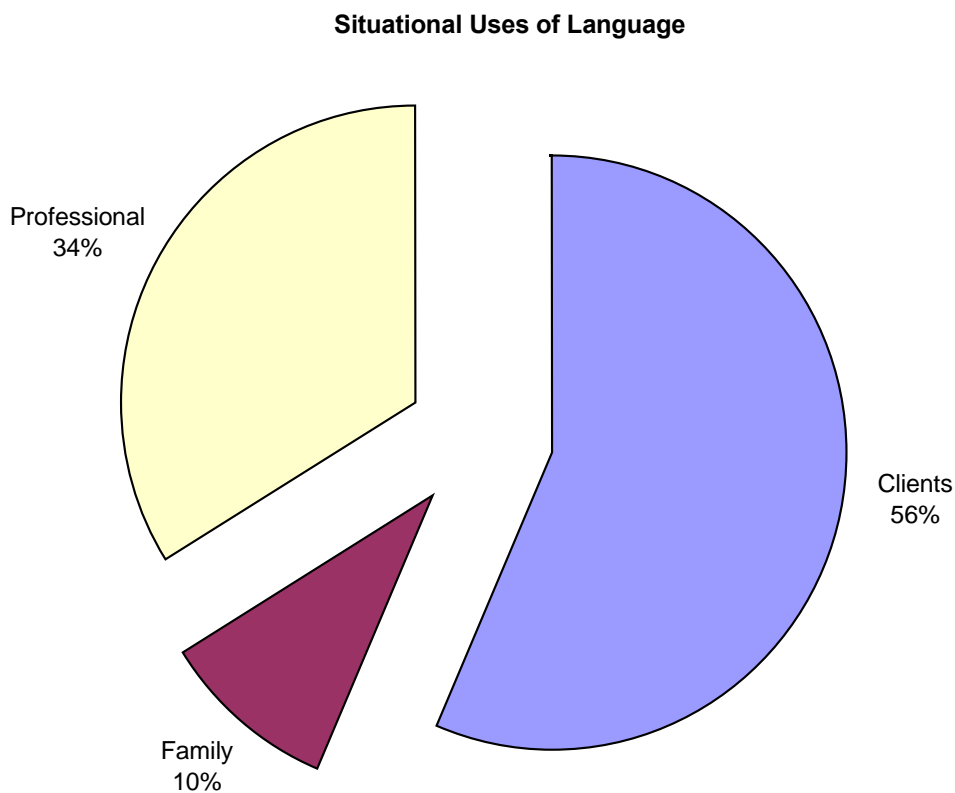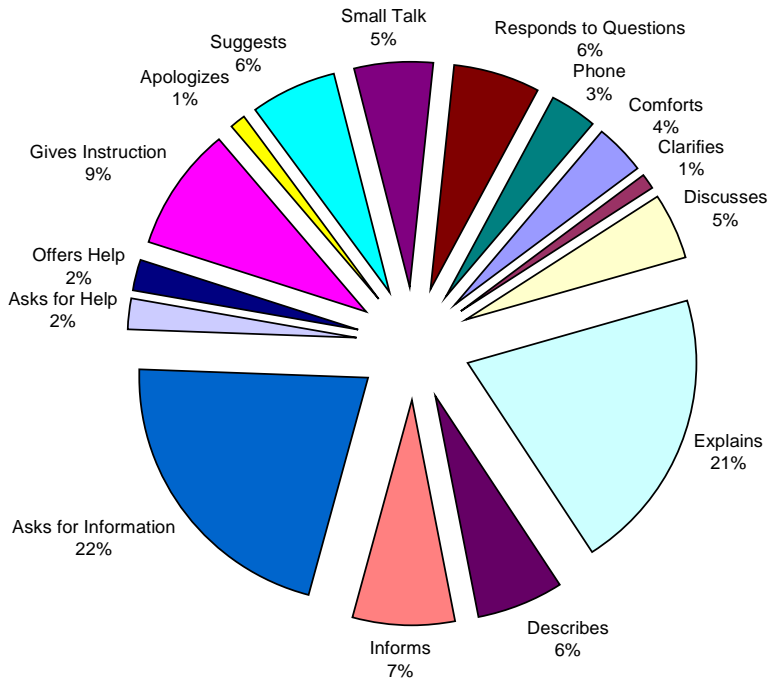Professional
34%

Clients
56%

Family
10%

Chart 2 illustrates the types of tasks observed, and the percentage of time that was spent on each task, based on all the observations.

Language Tasks

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Appendix C - Letter of Explanation to Pilot Testing Candidates

To Whom It May Concern:

We would like to invite you to participate in a project in which we are doing research regarding the English language assessment of internationally educated nurses. Presently the language tests that are being used for internationally educated nurses do not reflect the language the nurses actually use on the job. We are exploring the possibility of developing a test that will use content and test language more specifically related to the nursing profession. Using this information, it is hoped that a more appropriate and relevant process will be in place to allow internationally-educated nurses to demonstrate their English language ability in the profession.

You need to know that participating in this project does not pose any risk to you. All the information that we gather will be strictly confidential, and your name will not be used in any reports. Data will be stored in a locked cabinet, and will be destroyed two years after the completion of the project.

If you agree to participate in the study, you will be tested using two English language tests. One test will be based on the language used in the nursing profession. It will test your speaking (a separate 45 minute interview), listening, reading and writing. The complete test will take about 3 hours. You will also be asked to respond to some questions about the test.

The other test, the Canadian Language Benchmarks Placement Test (CLBPT) will take about one hour. Again, you will be tested on speaking, listening, reading and writing. The results of this test will determine your language level in terms of Canadian Language Benchmarks. If you have taken the CLBA or CLBPT in the past six months and have official proof of your score, you do not need to take the CLBPT again. Just have proof of your score available for us.

Only the researchers will know what your test results are. If you want to know your test results you can contact us using the contact information at the top of the consent form that you signed. We will only give you the results if you request them.

It is important for you to know that you do not have to participate in this project. Participation is voluntary. Also, you can change your mind about participating at any time. Just contact someone at your school, or one of us if you do change your mind. This will not affect your job or school status in any way.

Thank you for considering this project.

Yours truly,

Lucy Epp   Catherine Lewis

Appendix D - Questionnaire For Participants In Pilot Testing

APPENDIX D: **QUESTIONNAIRE FOR PARTICIPANTS IN PILOT TESTING**

ID #

1. Name:_____
   FAMILY   GIVEN

2. Gender: male   female

3. Age: _____ 20-25   _____26-30   _____ 31-35

   _____36-40   _____41-45   _____ 46 +

4. Country Of Origin _____

5. First Language _____

6. Were you educated as a nurse in a country other than Canada?
   YES   **NO**

If **NO**, go to question # 11.

If **YES**, answer the following questions:

How long have you lived in Canada? _____
                                    years   months

7. Describe any education or work experience you have had in
   Canada.

Education In Canada	How Long?
1.	
2.	
3.	

Work Experience In Canada (Job Title) (Including Volunteer Work)	How Long?
1.	
2.	
3.	

8. In which country (or countries) did you attend school (primary and high school)?

☐

_____

9. In which country (or countries) did you receive your nursing education?

☐

_____

10. In which language(or languages) did you receive your nursing education?

☐

_____

11. Have you met the English language requirement for nursing in your province? (e.g., TOEFL or IELTS)

YES   **NO**

☐

If yes, which test/s did you take? _____

Which was your most recent test and what was your score?

Test: _____   Score: _____

FOR ALL PARTICIPANTS:

12. Have you passed the Canadian Registered Nurse Exam?

☐

YES   **NO**

If YES, when? _____

If you wish to receive your results on these tests, please write you complete mailing address and e-mail address here:

Street Address: _____

City, Province, Postal Code: _____

e-mail address: _____

Appendix E - Consent Form For Pilot Testing Candidates

APPENDIX E: Consent Form For Pilot Testing Candidates

Pilot Testing Consent Form

Part 1: Researcher Information
Name of Principal Investigator: Lucy Epp, Instructor/Researcher Affiliation: Red River College, Language Training Centre Contact Information: Suite 300-123 Main Street, Winnipeg, MB; R3C 1A3; phone: (204) 945-6151; e-mail: lepp@rrc.mb.ca
Name of Co-Investigator: Catherine Lewis, Instructor/Researcher Affiliation: Red River College, Language Training Centre Contact Information: Suite 300-123 Main Street, Winnipeg, MB; R3C 1A3; phone: (204) 945-6151; e-mail: clewis@rrc.mb.ca

Part 2: Consent of Subject	Yes	No
Do you understand that you have been asked participate in a project?		
Have you read and received a copy of the attached information sheet?		
Do you understand the benefits involved in taking part in this research study?		
Have you had an opportunity to ask questions and discuss the study?		
Are you willing to allow researchers to use samples of your unidentified writing and speaking to train assessors?		
Do you understand that you are free to refuse to participate or withdraw from the study at any time? You do not have to give a reason and it will not affect your work or study situation.		
Has the issue of confidentiality been explained to you? Do you understand who will have access to the data gathered?		

Part 3: Signatures
This study was explained to me by: _____ Date: _____
I agree to take part in this study. Signature of Research Participant: _____ Printed Name: _____
Witness (if available): _____ Printed Name: _____
I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate. Presenter: _____ Printed Name: _____
* A copy of this consent form must be given to the subject.

Appendix F - Feedback From Candidates

APPENDIX F: Feedback From Candidates

SPEAKING

1) My overall impression of the speaking test is …

          (a) very good   (b) good   (c) fair

2) The length of the speaking test was… (a) too short   (b) just right (c) too long

3) Was the content of the test familiar to you?

        (a) very familiar   (b) somewhat familiar   (c) not familiar

4) Do the speaking tasks on the test reflect the nursing profession? YES   NO

5) Role Play One (asking questions) was… (a) too easy   (b) just right   (c) too hard

6) Role Play Two (giving instructions) was…(a) too easy   (b) just right   (c) too hard

7) The questions were …   (a) too easy (b) just right   (c) too hard

8) The assessors made me feel…   (a) comfortable   (b) uncomfortable

OTHER COMMENTS _____

_____

LISTENING COMPREHENSION

1) The questions were…   (a) too easy   (b) just right   (c) too hard

2) The length of the test was… (a) too short   (b) just right   (c) too long

3) The amount of time allowed for the listening test was…

          (a) too much   (b) enough   (c) not enough

4) Was the content of the test familiar to you?

          (a) very familiar   (b) somewhat familiar   (c) not familiar

5) What did you think about the following aspects of the test (check your choice):

	effective	not effective
a) nursing content		
b) multiple choice format		
c) chart format		
d) video section		
e) audio section		

6) My overall impression of the listening comprehension test is …

          (a) very good   (b) good   (c) fair

OTHER COMMENTS:

_____

_____

SKIMMING AND SCANNING

1) The questions were…   (a) too easy   (b) just right   (c) too hard

2) The length of the test was… (a) too short   (b) just right   (c) too long

3) The amount of time allowed for the skimming and scanning test was…

                              (a) too much   (b) enough   (c) not enough

4) Was the content of the test familiar to you?

                  (a) very familiar   (b) somewhat familiar   (c) not familiar

5) What did you think about the following aspects of the test (check your choice):

	effective	not effective
a) nursing content		
b) question and answer format		
c) Part One (patient information)		
d) Part Two (progress notes)		

6) My overall impression of the skimming and scanning test is …

                  (a) very good   (b) good   (c) fair

OTHER COMMENTS _____

_____

READING COMPREHENSION

1) The questions were…   (a) too easy   (b) just right   (c) too hard

2) The length of the test was… (a) too short   (b) just right   (c) too long

3) The amount of time allowed for the reading comprehension test was…

                         (a) too much   (b) enough   (c) not enough

4) Was the content of the test familiar to you?

               (a) very familiar   (b) somewhat familiar   (c) not familiar

5) What did you think about the following aspects of the test (check your choice):

	effective	not effective
a) nursing content		
b) multiple choice format		
c) fill-in-the-blanks format		

6) My overall impression of the reading comprehension test is …

            (a) very good   (b) good   (c) fair

OTHER COMMENTS _____

_____

WRITING

1) Was the content of the test familiar to you?

        (a) very familiar   (b) somewhat familiar   (c) not familiar

2) My overall impression of the writing test is …(a) very good   (b) good   (c) fair

TASK ONE: FORM

3) The form was …   (a) too easy   (b) just right   (c) too hard

4) The length of Task One was… (a) too short   (b) just right   (c) too long

5) The amount of time allowed for Task One was…

        (a) too much   (b) enough   (c) not enough

TASK TWO: REPORT

6) The report was … (a) too easy   (b) just right   (c) too hard

7) The length of Task Two was… (a) too short   (b) just right   (c) too long

8) The amount of time allowed for the writing test was…

        (a) too much   (b) enough   (c) not enough

OTHER COMMENTS _____

_____

Appendix G – CELBAN Test Results

CELBAN Test Results

NAME _____   ID #_____

Canadian Language Benchmarks Placement Test (CLBPT) Scores	
Speaking	
Listening	
Reading	
Writing	

Canadian English Language Benchmarks Assessment for Nurses (CELBAN) Scores	
Speaking	
Listening	
Reading	
Writing	

ANALYSIS OF STRENGTHS AND WEAKNESSES OF PRODUCTIVE SKILLS

SPEAKING	Strengths	
	Weaknesses	
WRITING	Strengths	
	Weaknesses	

Benchmarks recommended for Nursing in the Canadian Nursing context	Speaking	8
	Listening	9
	Reading	8
	Writing	7

Appendix H - Canadian Language Benchmarks Assessment For Nurses In Canada (Celban)

Summary: This test consists of four sections which assess four English language skills: speaking, listening, reading and writing. The contents of the test are based on the results of Phase I of the project (Benchmarking the English Language Demands of the Nursing Profession across Canada*), in which an occupation specific language analysis of the nursing profession was carried out across Canada. In that project, language demands of the nursing profession were established using the Canadian Language Benchmarks[9]. Results are reported as Canadian Language Benchmark Levels

Speaking

- The speaking assessment includes an oral interview and role plays.

- Two assessors will carry out the assessment.

- Speaking assessment will be tape recorded for future reference if needed by assessors to verify score.

- Context and content of speaking tasks are based on data collected from Phase I.

- Tasks were created with input from nursing instructors/ consultants to ensure authenticity.

- The speaking assessment includes two role plays in which the candidate is asked to interact with the "patient" (one of the assessors) by asking questions to obtain information, and to give instructions and offer explanations. In addition, the candidate will be asked to answer questions to demonstrate ability to narrate, describe, summarize, synthesize, state and support opinion, and advise.

- Total time: 30 minutes

- Criteria for scoring speaking tasks is based on CLB descriptors

  - General use of language

  - Intelligibility

  - Organization

  - Fluency

  - Use of cohesive devices

  - Adequacy of vocabulary for purpose

  - Grammar

  - Use of strategies

  - Characteristics of Interaction

- Scores are assigned as CLB levels

---

[9] available from the Centre for Canadian Language Benchmarks @ www.language.ca

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

- Feedback on strengths and weaknesses is provided

Listening

- Context and content of scenarios are based on data collected from Phase I.
- All scenarios were created with input from nursing instructors/consultants to ensure authenticity.
- Five video scenarios (in various settings including hospital, home, clinic, and medical office).
- Four audio scenarios (phone calls and shift-to-shift reports).
- Scenarios include interactions between nurses and patients, family members, and other professionals.
- Question format: multiple choice (some in chart format)
- Total time: 50 minutes

Reading

- Context and content of reading texts are based on data collected from Phase I.
- All texts were created with input from nursing instructors/ consultants to ensure authenticity.
- The reading assessment includes two sections:
  - skimming and scanning (10 minutes)
  - reading comprehension (40 minutes)
- Text includes various formats such as charts, patient notes, manuals, and information texts related to health issues.
- Question format:
  - short answer questions (skimming and scanning)
  - multiple choice questions (reading comprehension), including a cloze exercise

Writing

- Context and content of writing tasks are based on data collected from Phase I.
- Tasks were created with input from nursing instructors/ consultants to ensure authenticity.
- The writing assessment includes two sections:
  - Form-filling (10 minutes)
  - Report writing (20 minutes)
- Criteria for scoring writing tasks is based on CLB descriptors:
  - Criteria for form-filling:

- conventions of form filling (spelling, legibility, point form)
- necessary information included
- Criteria for report writing:
    - effectiveness
    - grammar
    - discourse/fluency
    - vocabulary for purpose/content
- Scores are assigned as CLB levels
- Feedback on strengths and weaknesses is provided

Glossary of Terms

CanTEST

CanTEST is a language assessment tool developed by the University of Ottawa which also provides results as CLB scores. It offers testing of higher CLB levels.

CCLB

Centre for Canadian Language Benchmarks

CELBAN

Canadian English Language Benchmarks Assessment for Nurses, a product created by a test development team led by Red River College in Manitoba, for the Centre for Canadian Language Benchmarks in 2002 - 2003.

CLB

Canadian Language Benchmarks describe what adult second language learners can do, using English, at twelve levels of proficiency, or benchmarks. The twelve Benchmarks are separated into three progressive levels of proficiency and cover 4 skill areas, Listening, Speaking, Reading, and Writing.

Face validity

Face Validity is the extent to which a test meets the expectations of those involved in its use, e.g. adminstrators, teachers, candidates, and test score users; the acceptability of a test to its stakeholders.

L1

A native speaker of a language.

L2

Second language speaker. For example, L2 English speaker is someone who does not have English as their first (or native) language.

NAC

National Advisory Committee, was a group of nursing stakeholders who provided advice and information to the test developers and researchers in Phase I and Phase II of the project. In addition, some NAC members helped coordinate arrangements for focus groups and pilot test site locations.

OISE

Ontario Institute for Studies in Education at the University of Toronto.

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool

Glossary of Terms ● 63

Regulatory Bodies

The professional organization responsible for registering professionals in that profession to practice within a specific jurisdiction. For example, each province has regulatory bodies responsible for registering Registered Nurses and for Practical Nurses.

Stakeholders

In the case of the Benchmarking the Nursing Profession project, stakeholders usually refers to representatives from nursing regulatory bodies, nursing schools, healthcare representatives and employers, and internationally-educated nurses.

TOEFL

Test of English as a Foreign Language, a general language proficiency test, used in Canada and abroad for English language assessments.

TSE

Test of Spoken English, is a tool used to evaluate spoken English communication skills in Canada and abroad. It is owned by the same organization that runs TOEFL.

64 ● Glossary of Terms

The Development of CELBAN (Canadian English Language Benchmarks Assessment for Nurses A Nursing Specific English Language Assessment Tool