Proving My Competency One Test at a Time:
Internationally Educated Nurses and the Canadian English Language Benchmark Assessment for Nurses

by

Stefanie K. R. Baldwin-Bojarski

A thesis submitted to the Faculty of Education
in conformity with the requirements of the
Degree of Master of Education

Queen’s University
Kingston, Ontario, Canada
August, 2016

Copyright ©Stefanie Baldwin-Bojarski, 2016
Abstract

Professional language assessment is a new concept that has great potential to benefit Internationally Educated Professionals and the communities they serve. This thesis reports on a qualitative study that examined the responses of 16 Canadian English Language Benchmark Assessment for Nurses (CELBAN) test-takers on the topic of their perceptions of the CELBAN test-taking experience in Ontario in the winter of 2015.

An Ontario organization involved in registering participants distributed an e-mail through their listserv. Thematic analyses of focus group and interview transcripts identified 7 themes from the data. These themes were used to inform conclusions to the following questions: (1) How do IENs characterize their assessment experience? (2) How do IENs describe the testing constructs measured by the CELBAN? (3) What, if any, potential sources of construct irrelevant variance (CIV) do the test-takers describe based on their assessment experience? (4) Do IENs feel that the CELBAN tasks provide a good reflection of the types of communicative tasks required of a nurse?

Overall, participants reported positive experiences with the CELBAN as an assessment of their language skills, and noted some instances in which they felt some factors external to the assessment impacted their demonstration of their knowledge and skill. Lastly, some test-takers noted the challenge of completing the CELBAN where the types of communicative nursing tasks included in the assessment differed from nursing tasks typical of an IENs country or origin. The findings are discussed in relation to literature on high-stakes large-scale assessment and IEPs, and a set of recommendations are offered to future CELBAN administration. These recommendations include (1) the provision of a webpage listing all licensure requirements (2) monitoring of CELBAN
location and dates in relation to the wider certification timeline for applicants (3) The provision of additional CELBAN preparatory materials (4) Minor changes to the CELBAN administrative protocols. Given that the CELBAN is a relatively new assessment format and its widespread use for high-stakes decisions (a component of nursing certification and licensure), research validating IEN-test-taker responses to construct representation and construct irrelevant variance is critical to our understanding of the role of competency testing for IENs.
# Table of Contents

Abstract ......................................................................................................................... ii
Acknowledgements ........................................................................................................ iii
List of Tables ................................................................................................................... vii
List of Figures ................................................................................................................ viii
List of Abbreviations ..................................................................................................... ix
Chapter 1 Introduction ................................................................................................... 1
  Definitions .................................................................................................................... 3
  Rationale ...................................................................................................................... 6
  Context ......................................................................................................................... 8
  Research Questions ..................................................................................................... 13
Chapter 2 Literature Review ......................................................................................... 14
  Advanced English Language Learners and Communicative Language Teaching .... 15
  Advanced English Language Learners and Genre .................................................... 17
  Integration Challenges of IENs .................................................................................. 19
  Research on the CELBAN ......................................................................................... 21
  Validity, Reliability and the Testing Constructs ....................................................... 23
  High-Stakes Large Scale Assessment and Stakeholder Beliefs .............................. 26
  Validity: Consequences and Construct Irrelevant Variance .................................. 27
  Validity & Construct Irrelevant Variance ................................................................. 31
  Context-Sensitive Extrapolation .............................................................................. 33
  Summary .................................................................................................................... 34
Chapter 3 Methodology ............................................................................................... 36
  Methodological Approach ......................................................................................... 36
  Participants ................................................................................................................. 37
  Instrument ................................................................................................................ 41
  Pilot Study .................................................................................................................. 44
  Data Collection .......................................................................................................... 45
  Participant Demographics and CELBAN Readiness .............................................. 46
  Focus Groups & Interviews ..................................................................................... 47
  Data Analysis ............................................................................................................ 49
  Summary .................................................................................................................... 51
Chapter 4 Results ......................................................................................................... 53
  Theme 1 Descriptions of the CELBAN ..................................................................... 55
Theme 2 Consequences of the CELBAN Test Scores ........................................58
Theme 3 Skills Tested by the CELBAN ................................................................61
Theme 4 External Factors Impacting Test Experience ......................................66
Theme 5 Preparing for the CELBAN .................................................................72
Theme 6 Test Conditions and Administration ..................................................74
Theme 7 Authenticity .........................................................................................78
Summary ..........................................................................................................83
Chapter 5 Discussion ......................................................................................84
How do IENs characterize their test experience? ...........................................85
How do IENs describe the constructs measured by the CELBAN? .................89
What, if any, potential sources of Construct Irrelevant Variance do the test-takers describe based on their experience? .........................................................94
Do IENs feel the exam tasks provide an accurate reflection of the types of communicative tasks required of a nurse? ..................................................102
Recommendations for CELBAN administration ...........................................107
Limitations and Directions for Future Research ............................................110
Conclusion ......................................................................................................112
References .....................................................................................................114
Appendix A Ethics Letter of Approval ..............................................................123
Appendix B Semi-Structured Focus Group Guide .........................................124
Appendix C Semi-Structured Interview Guide ..............................................126
Appendix D Letter of Information & Consent Form ........................................127
Appendix E Demographic Information: Sex, Age, Country of Origin & First Language .................130
Appendix F Demographic Information: Language & Nursing Experience ..........132
Appendix G Demographic Information: Language Assessment & Work ...........134
Appendix H Demographic Survey .................................................................136
Acknowledgements

This milestone would not have been possible without the support of Dr. Liying Cheng who inspires young educators, and believes in their potential to become great educators. Thank you for your mentorship, your leadership, and your unwavering support. I must also acknowledge the excellent feedback of Dr. Elizabeth Lee, and Dr. Lynda Colgan whose thoughtful notes have made this a better study.

When I started this degree I wondered who I would meet, and what I would do when my thesis was complete. I am deeply indebted to James McNutt, Esther Heidt Bettney, and Meaghan Low. Your collegiality, warmth, and friendship were so needed and incredibly appreciated. I am so grateful Paul Baker and Nathan Gardi for the opportunity you have given me to apply this theory to learning solution design challenges every day at D2L. There are not words enough to express my gratitude for your ongoing mentorship, and commitment to change the way the world learns. Lihua First, Anna Magolon-Parobek, Mihaela Lucaci, Jenny Hill and the entire Creative Services Department thank you for your advice, thoughts, and support on, and for this project.

Mum, Dad, Dasha and Masha, each of you exemplify a commitment to life-long-learning that has inspired me from my earliest days to finish what I start, and put forth my best effort. Thank you for your example, love and support.

Jesse Baldwin, what more could I ask for that you have not already given? You stood at my side as I chose a harder path, and then you walked it with me. Ani LeDodi VeDodi Li.
List of Tables

Table 1. Minimum Acceptable CELBAN Scores ................................................................. 9
Table 2. Minimum Acceptable IELTS Scores ................................................................. 9
Table 3. Content and format of the CELBAN ................................................................. 11
Table 4. Sensitizing Concepts ......................................................................................... 33
Table 6. Research Questions, Focus Group and Interview Guide Alignment ............... 43
Table 5. Self-Reported CLB Scores and Minimum CELBAN Pass Scores ..................... 47
Table 7. Research Questions, Themes, Codes & Frequencies ........................................ 54
Table 8. Demographic Information: Sex, age, country of origin, & first language ........ 130
Table 9. Demographic Information: Language & nursing experience .......................... 132
Table 10. Demographic Information: Language & Work .............................................. 134
List of Figures

Figure 1. Focus Groups and Interview Participation................................................................. 39
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AELLS</td>
<td>Advanced English Language Learners</td>
</tr>
<tr>
<td>CELBAN</td>
<td>Canadian English Language Benchmark Assessment for Nurses</td>
</tr>
<tr>
<td>CLB</td>
<td>Canadian Language Benchmarks</td>
</tr>
<tr>
<td>CLT</td>
<td>Communicative Language Teaching</td>
</tr>
<tr>
<td>CNO</td>
<td>College of Nurses Ontario</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second Language</td>
</tr>
<tr>
<td>IENs</td>
<td>Internationally Educated Nurses</td>
</tr>
<tr>
<td>IEPs</td>
<td>Internationally Educated Professionals</td>
</tr>
<tr>
<td>L1</td>
<td>First Language</td>
</tr>
<tr>
<td>L2</td>
<td>Second Language</td>
</tr>
<tr>
<td>MELAB</td>
<td>Michigan English Language Assessment Battery</td>
</tr>
<tr>
<td>RN</td>
<td>Registered Nurse</td>
</tr>
<tr>
<td>RPN</td>
<td>Registered Practical Nurse</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

I am a student. I have been one for a long time. When I first saw the Simpson’s episode where the teachers go on strike and Lisa Simpson creates make-work projects to keep herself busy, desperately exclaiming to her mother “I’m oh so good, and oh so smart” finally demanding, “grade me!” I felt a disturbing personal resonance. Messick (1996) puts forth the notion that when we speak about validity, we are not speaking about a property of a test, but rather the meaning or inferences we make from test scores. Where assessment functions as a gatekeeper separating those who are successful from those who are not, personal fulfillment of those assessed may come to be linked with those inferences drawn from test scores (Cheng, Spaling & Song, 2013). Britzman (1997) suggests that the problem with the desire for mastery is that it must pass through the other. In this respect then, personal fulfillment linked to mastery requires an arbiter, or assessment by which an individual may form a part of their personal identity. Messick (1991) outlines a variety of validation approaches with which we may begin to understand “both the existing evidence for and the actual as well as potential consequences of score interpretation and use” (p. 2), or validity of score inference. One validation approach described by Messick involves:

Apprais[ing] the value implications and social consequences of interpreting and using the test scores in the proposed ways, scrutinizing not only the intended outcomes but also unintended side effects—in particular, evaluate the extent to which (or, preferably, discount the possibility that) any adverse consequences of testing derive from sources of score invalidity such as irrelevant test variance (1991, p. 6).
This qualitative study examines the experiences of International Educated Nurses (IENs) who have recently completed the Canadian English Language Benchmarks for Nurses (CELBAN), for licensure and certification to work in the nursing profession in Canada. As the CELBAN is defined by its name as an assessment, this study will use the term “assessment” to refer to it specifically. Due to the fact that the CELBAN functions as a high-stakes test, IENs who must complete it will be referred to as “test-takers,” and its locations as “testing centres.” This study thus refers to high-stakes testing literature to inform its analyses and refers to these studies using the language of their unit of analysis either tests or examinations. This study, in particular, explores the perceptions of the IEN test-takers on: the assessment as a whole, its testing constructs, possible sources of irrelevant test score variance, and the authenticity of the representation of nursing communicative tasks. It is the intention of this study to add to the critical understanding of what is tested by the CELBAN and the role of large-scale language testing in the licensure and certification experience of IENs.

Lather (2009) has suggested that to attempt empathic understanding is to “invite conquest” (p. 20), and suggest a sameness of subjectivities that negates difference and all that can be learned from preparing one’s self “with ears to hear” (p. 18) a story that is not one’s own. For these reasons, my study will treat demographic data with caution and focus collection and analysis on participant descriptions of their experience completing the CELBAN.

By examining test-taker feedback on this Canadian assessment of professional English language competency, this study contributes to a wider understanding of the role of large-scale language testing in the licensure and certification experience of IENs.
Researchers in large-scale, high-stakes assessment have demonstrated the need for such assessments to “obtain additional information from test-takers for validating large-scale assessments” (Cheng & DeLuca, 2011, p. 104). In acknowledgement of the need in this area of research, this study employed a series of focus groups and individual interviews immediately following the completion of their assessment to collect and analyze test-taker feedback on their CELBAN experience.

**Definitions**

**Assessment**

Nagy (2000), defines assessment as an activity which has three roles or functions: gatekeeping, accountability, and instructional diagnosis. For the purposes of this study the term assessment is used as Nagy defines it particularly with respect to his suggestions that assessment is an activity used for these three functions. In particular this term is applied to the CELBAN not only due to its inclusion of the term in its very name, but also due to the functions which are enacted by test-taker completion of the CELBAN. The CELBAN determines who is granted the opportunity to continue in the licensure and certification process, if prior language knowledge is sufficient and what level of language mastery the test-taker possesses. Gatekeeping is described as the function, “in which assessment determines who is granted a privilege such as admission or graduation” (p.262). Accountability is described as the function, “in which assessment is used to decide if schools are working well” (p. 262). Instructional diagnosis is described as the function, “in which assessment is used to find out what students do and do not know, and what to do about it” (p.262).
**Evaluation**

Patton (2012), defines evaluation as “an applied inquiry process for collecting and synthesizing evidence that culminated in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal or plan” (p.3). He further explains that “conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgement about the value of something)” (p.3). This terms is used in this study to describe the conclusions made on the basis of evidence provided through the submission of assessment scores, such as the CELBAN, to determine the readiness of an IEN to be licensed and certified to work as a nurse in Canada.

**High-Stakes Testing**

DeLuca et al. (2013) define high-stakes testing with respect to the decisions made as a result of test measurement which have a significant impact on the test-taker, specifically, “use of this test for high-stakes decisions (e.g., university admissions, scholarship, promotion)” (p. 664). This definition is used to describe the CELBAN as it has direct, and significant impact on the test-takers in terms of their ability to proceed in the licensure and certification process.

**Instrument**

Patton (2002) describes qualitative researchers as a subjective instrument in their collection, analysis, and interpretation of research data. In this respect, I functioned as an instrument for these processes within this study. McMillan & Schumacher (2010), refer to instrumentation as a threat to test validity in “the way changes to the instruments or person used to collect data may affect the results” (p. 112). In this study I adapted a pre-existing interview guide and utilized this guide as another instrument to assist me in the
data collection. In utilizing a pre-existing tool I attempted to mitigate factors which may negatively impact my ability to collect data, specifically those threats identified by McMillan & Schumacher (2010): “the observers may become fatigued or bored or change in some other way so as to affect the recording of data” (p. 112).

**Large-Scale Achievement Testing**

Fox, J. & Cheng, L. (2007) define large-scale achievement testing as “used to measure and ensure student competency or provide system accountability” (p. 11). Patton (2012) defines standardized tests more generally as a tool which “measure[s] the reactions of many respondents in a way that statistical aggregation and analysis are relatively straightforward, following established rules and procedures” (p. 290). These definitions have informed my use of the term Large-Scale Assessment and Large-Scale Testing to describe those opportunities to measure the skills of many respondents in order to report back a measurement that is relatively straightforward.

**Measurement**

Measurement is defined by McMillan & Schumacher (2010) as a process which “involves assigning numbers to things in order to differentiate one thing from another” (p. 150). In the case of the CELBAN, the intention of this assessment is to offer what McMillan & Schumacher (2010) describe as a nominal measurement, or a measurement scale that accomplishes a “naming of mutually exclusive categories or people, events, or other phenomena” (p. 150). In the case of the CELBAN the nominal measurement scale allows for the description, or naming of people, according to CLB levels which ultimately name or describe them as competent in the area of nursing specific English language, or not yet competent.
Rationale

Each year, over fifty percent of all immigrants to Canada choose Ontario as their home (Goldberg, 2001). Although the federal government retains the responsibility for recruitment and selection of immigrants, it remains the charge of the provincial government to oversee the education, training and regulation of professions (Goldberg, 2001). Research has reported only nineteen-to-forty percent of Internationally Educated Professionals (IEPs) obtain suitable employment in their professional field (Ngo & Este, 2006). Cheng, Spaling & Song (2013) have noted that “non-recognition of international credentials and challenges in satisfying the licensure and certification testing requirements have been identified as the largest barriers to successful integration into the workforce, which costs the Canadian economy as much as $5 billion a year” (p. 734).

The Canadian Nursing Association has raised concerns that as large numbers of Canadian nurses approach retirement age, a “critical shortage of nursing professionals” could result in Canada (Epp & Lewis, 2009, p. 286).

As of 2006, more than ninety-two percent of Internationally Educated Nurses (IENs) both Registered Nurses (RNs) and Registered Practical Nurses (RPNs) registered with the College of Nurses Ontario (CNO), the provincial regulatory body, were employed (Blythe et al., 2009). In 2007, 10,165 IENs were registered in Ontario, representing approximately eleven percent of the provincial nursing workforce (Blythe et al., 2009). Although one might interpret these statistics optimistically, the CNO has estimated an attrition rate of forty percent among Registered Nurses educated abroad who apply for membership (Blythe et al., 2009). This estimated loss of IENs to the profession suggests a need for further research on how best to support federally selected IENs in
their pursuit of licensure and certification with the CNO, thus facilitating their subsequent economic integration.

Researchers studying those factors which affect IEP integration have raised concerns that:

…if employers dictate that candidates require ‘Canadian’ work experience, and do not recognize the knowledge and experience obtained outside of Canada, this perpetuates racist assumptions and attitudes about the ‘others’’ skills and abilities and maintains the dominant White culture in the Canadian workplace” (Yee et al., 2007, p. 3).

The CELBAN is intended to facilitate integration by offering IENs the opportunity to earn their licensure and certification to practice by demonstrating their English language ability in the area of nursing communicative tasks (Epp & Lewis, 2009). The CELBAN attempts to recognize the professional knowledge of IENs obtained in their home countries by focusing language testing specifically on the language of the profession, an acknowledgement that IENs are already trained in the practice of nursing in another country (Epp & Lewis, 2009). As a national testing instrument of professional language competency, the CELBAN judges what constitutes proficiency based on the receptive and productive language skills demonstrated by the test-takers on their given test date.

Messick (1991) has suggested that the validity of a test may be understood as the “overall evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores or other modes of assessment” (p. 245). He identifies two major threats to test validity: (1) construct underrepresentation and (2) construct irrelevant variance. Construct underrepresentation is defined as an occurrence whereby a test “…is too narrow and fails to include important dimensions or facets of the construct [the
knowledge or skill being assessed)” (p.14). Construct irrelevant variance is defined as an occurrence whereby a test … “is too broad and contains excess reliable variance associated with other distinct constructs as well as method variance making items of tasks easier or harder for some respondents in a manner irrelevant to the interpreted construct” (p.14). In accordance with this view of validity, this study examines the perceptions of test-takers on the skills tested by the CELBAN as well as their experience in completing the assessment. Therefore, this study aims to contribute information to our understanding of the evaluative judgements made from CELBAN scores.

**Context**

Recent changes to the CNO requirements for IENs undergoing the certification and licensure process has meant that applicants must now show language competency by completing either the Canadian English Language Assessment for Nurses (CELBAN) or International English Language Testing System (IELTS) exams (CNO, 2013). Tables 1 and 2 outline the minimum language levels deemed acceptable on each of the aforementioned language assessments to meet the language requirements for the CNO.
Table 1.

Minimum Acceptable CELBAN Scores

<table>
<thead>
<tr>
<th>Test Component</th>
<th>CELBAN Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing</td>
<td>7</td>
</tr>
<tr>
<td>Speaking</td>
<td>8</td>
</tr>
<tr>
<td>Listening</td>
<td>10</td>
</tr>
<tr>
<td>Reading</td>
<td>8</td>
</tr>
</tbody>
</table>

Note 1. As disseminated by the College of Nurses of Ontario's website updated September 10, 2014
Note 2. These are the minimum benchmark scores that must be achieved and includes the standard error of measurement
Note 3. The College will only accept scores from a CELBAN test at a CELBAN administration site. The College does not accept results from CELBAN tests administered by educational institutions for educational purposes.

Table 2.

Minimum Acceptable IELTS Scores

<table>
<thead>
<tr>
<th>Test Component</th>
<th>CELBAN Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing</td>
<td>7</td>
</tr>
<tr>
<td>Speaking</td>
<td>7</td>
</tr>
<tr>
<td>Listening</td>
<td>7.5</td>
</tr>
<tr>
<td>Reading</td>
<td>6.5</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
</tr>
</tbody>
</table>

Note 1. The information above reflects that disseminated by the College of Nurses of Ontario's website updated September 10, 2014
Note 2. These are the minimum benchmark scores that must be achieved and includes the standard error of measurement of 0.5
Note 3. The College will only accept scores from the Academic format of the IELTS test
In the year 2000, a survey of fifty nursing profession stakeholders, conducted by the Centre for Canadian Language Benchmarks, indicated a “strong need for a specialized English language assessment tool to evaluate the English language communication skills of internationally educated nurses seeking registration in Canada” (CELBAN, n.d). This observation is consistent with studies conducted in other areas of medical practice such as Pharmacy Education where it was noted that:

The high levels of speaking, listening, reading, and writing skills required in pharmacy practice are simply not measured in an authentic or ecologically valid way through standardized, generic tests such as TOEFL, which are designed to measure fluency in a conversational, rather than professional, context (Austin & Dean, 2006, p.22).

Research has shown “when a test designed for the purposes of English for Academic Purposes and for university entrance is used for job or professional licensure and certification purposes, construct underrepresentation may occur” (Cheng & DeLuca, 2011, p. 106). In particular, for nurses, the construct underrepresented by general academic English language tests such as IELTS is nursing specific language (Epp, 2006). To address the potential issue of construct underrepresentation (professional language competency), the CELBAN was developed from 2000-2004. The CELBAN was designed specifically to address the language demands or “target language use” of the nursing profession, nation-wide (Epp & Lewis, 2009, p. 285). In particular, IENs receive feedback from CELBAN on productive language skills, highlighting areas of strength and weakness for continued development (Epp & Lewis, 2009). The CELBAN is a paper and pencil test comprised of a series of activities outlined in Table 3 below (CELBANa, n.d):
Table 3.  
*Content and format of the CELBAN*

<table>
<thead>
<tr>
<th>Skill Area</th>
<th>Skills Assessed</th>
<th>Format</th>
<th>Length¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking Assessment</td>
<td>Candidates will be asked questions to demonstrate their ability to: - narrate - describe - summarize - synthesize - state and support an opinion - advise</td>
<td>1 oral interview 2 role plays 2 assessors</td>
<td>30</td>
</tr>
<tr>
<td>Listening Assessment</td>
<td>Candidates are asked to demonstrate their understanding of conversations in the following settings: - hospital - home - clinic - medical office</td>
<td>5 video scenarios 4 audio scenarios multiple choice</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Conversations may be between nurses and: - patients - family members - other professionals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Assessment</td>
<td>Candidates are asked to demonstrate their comprehension of the following text formats: - charts - patient notes - manuals - information texts related to health issues</td>
<td>1 skimming and scanning section (short answer) 1 reading comprehension section (multiple choice and cloze)</td>
<td>50</td>
</tr>
<tr>
<td>Writing Assessment</td>
<td>Candidates are asked to demonstrate their knowledge of the following areas: - the conventions of form filling including the use of correct spelling, legibility of writing, and the use of point form - the inclusion of necessary information including main ideas and supporting details - the conventions of a narrative report including appropriate use of vocabulary, the overall effectiveness of the report, and correct grammar</td>
<td>1 form filling section 1 report writing section</td>
<td>30</td>
</tr>
</tbody>
</table>

Notes: 1. The length is given in minutes.
Introduced in 2004, this test is intended to measure the professional language competencies of IENs. The test takes approximately three and one half hours to complete (CELBAN a, n.d).

The CELBAN tasks are based in both content and context from data collected from the analysis of the English language demands of the nursing profession across Canada (CELBAN a, n.d). The CELBAN is a task based, or genre specific evaluation, which is scored according the Canadian Language Benchmarks (English). The Canadian Language Benchmarks are a collection of levels from 1-12, which describe an individual’s communicative proficiency in English as a second language. A Canadian Language Benchmark (CLB) of 1 describes an individual with little to no communicative ability in English (Centre for Canadian Language Benchmarks, n.d). The Centre for the Canadian Language Benchmarks suggests that a CLB of 8-9 (in all skill areas) indicates that the individual has enough communicative competence in English to begin a college or university program in that language (Centre for Canadian Language Benchmarks, n.d). A CLB of 11-12 indicates a communicative competency akin to a native speaker (Centre for Canadian Language Benchmarks, n.d). The CELBAN can be taken a maximum of three times by the test-taker, but they must wait a minimum of three months between test attempts (CELBAN a, n.d).

As of December 2013, the CNO indicated that the requirement for passing CLBs on the CELBAN for the nursing profession were as follows: Writing 7, Speaking 8, Listening 10, Reading 8 (CNO, 2013). It is important to note that a pass on the listening section represents a level of language close to native-like proficiency (see Table 2). It is
also important to note that if an IEN is unable to meet the minimum level for a pass on the listening section or on any of the other sections of the test taken in a group setting (reading and writing), then the test-taker must re-take this entire section of the exam (CELBAN b, n.d.).

Given that the CELBAN is a relatively new testing format, with widespread use for high-stakes decisions (a component of licensure and certification), further research on test-taker perceptions of the testing constructs (construct representation) and their experience with the assessment (possible sources of construct irrelevant variance) is critical to our understanding of the inferences made from CELBAN test scores and furthers our understanding of the role of professional language competency testing in the licensure and certification of IENs. Therefore, this study addresses the following questions in relation to IENs taking the CELBAN:

**Research Questions**

1) How do IENs characterize their assessment experience?

2) How do IENs describe the constructs (English reading, writing, listening, and speaking proficiency) measured through the CELBAN?

3) What, if any, potential sources of construct irrelevant variance do the test-takers describe based on their assessment experience?

4) Do IENs feel that the CELBAN tasks provide an accurate reflection of the types of communicative tasks required of a nurse?
Chapter 2

Literature Review

In order to situate the issues related to professional language assessment and its impact on Internationally Educated Nurses (IENs), the literature reviewed in this chapter aims to provide an overview of several topics related to adult second language acquisition, literacy in a second language, and issues affecting immigrant integration. First, the literature on Advanced English Language Learners (AELLs) and genre specific approaches to language learning situates the choice of profession specific language assessments. These two sections highlight the academic debates regarding communicative language teaching and the possible limitations of such approaches on explicit vocabulary instruction versus a more holistic approach to reading. The role that language plays in licensure and certification of IENs is then addressed with respect to the integration challenges of IENs in Canada. Specifically, this chapter highlights the importance that research on immigrant integration has placed on including Internationally Educated Professionals (IEPs) in developing the programs and assessments intended to address the needs of this key stakeholder group. The final section of this literature review discusses the approach to test validation taken in the development of the Canadian English Language Benchmark Assessment for Nurses (CELBAN). The efforts taken to integrate IEN feedback in test development in order to establish face validity and the need for additional research in this area is addressed by reviewing CELBAN specific literature.
Advanced English Language Learners and Communicative Language Teaching

Advanced Learners is a term that has been used in the past to refer to college level students whose language proficiency “qualifie[d] them to take courses at upper division college levels” (Swaffer, 2004, p. 19). This term, Advanced English Language Learners (AELLs), is used here to refer to adults who are fluent in one or more prior language, and who continue to refine and develop their English language skills. All participants in the CELBAN may be said according to this definition to be AELLs.

There is a long time split that divides language learning from content learning in language education. However, this arrangement is now being questioned by many professionals in the field of adult second language acquisition (Byrnes & Maxim, 2004; Kern, 2004). Byrnes & Maxim (2004) argue that without moving to wider integration of language supports into content learning, the goal of competent advanced language abilities in a second language will continue to be upheld merely as an ideal, rather than a reality (p. viii). In particular, researchers nowadays question the current prevailing construct for language acquisition: communicative language teaching (CLT) (Byrnes & Maxim, 2004).

CLT may be thought of as an approach to, rather than a method of teaching a second language (Brown, 2007; Coelho, 2004). In CLT classrooms, the focus is not on organizational language forms, and at times fluency is regarded as more important than accuracy (Brown, 2007; Coelho, 2004). While some researchers view CLT as an approach that promotes authenticity in learning, beyond the fulfillment of requirements or tests (Brown, 2007), others are looking at the benefits of adopting explicit, deductive
approaches to teach the writing of specific genres required of advanced second language learners in a given field (Swaffar, 2004). The discussion on CLT in many ways furthers the debate on the question of whether or not AELLs require the development of genre specific language supports (Harris, K. R., & Graham, S., 2006; Hidi & Anderson, 1986; Kern, 2004; Ogle & Correa-Kovtun, 2010; Snow & Uccelli, 2009; Swaffar, 2004).

Research on English for Academic Purposes (EAP) and English as a Second Language (ESL) at the post-secondary level in Canada has demonstrated that English language programs, of either kind, have a significant, direct, positive effect on the academic and social engagement of advanced second language learners (Fox, Cheng, & Zumbo, 2013). This Canadian study which looked at thirty-six English language programs in twenty-six Canadian universities ($n=64$), allowed researchers to focus on a range of L2 learners currently enrolled in English-medium Canadian Universities (Fox et al., 2013). Their findings suggested that AELLs are likely to benefit from additional language scaffolds in EAP and ESL programs regardless of the pedagogically divergent nature of these programs. Working from that same premise, Swaffar (2004) suggests that explicit instruction in the many genres of English discourse is essential because these oral and written rhetorical practices enact culturally embedded communicative situations in a predictable fashion, enabling comprehension.

If AELLs are to comprehend and engage in multiple language discourses and genres, they will require explicit instruction in these kinds of generic discursive practices (Swaffar, 2004). Although AELLs are likely to have encountered a variety of rhetorical genres in their L1, they may have experienced only limited exposure to culturally specific
genres with extended continuous discourse in their L2 (Swaffar, 2004). Meschyan & Hernandez (2004) have suggested that individuals who achieve good command of both oral and written skills in their L2 have an advantage in school and work environments, where they can compete with native speakers and increase their chances of securing a higher standard of living.

**Advanced English Language Learners and Genre**

While there are some comparisons that can be made between the acquisitions of one’s first language to a second language, there are tremendous differences for second language learners in terms of both the cognitive and affective factors impacting learning in adulthood (Brown, 2007). Instructional approaches, which organize language curriculum by having students work through a sequence of phrases, sentences, paragraphs, and finally extended discourse, are criticized by Kern as “eminently logical, but do not mesh well with the psychological needs of language learners who strive to communicate in meaningful, whole acts” (Kern, 2004, p. 10). Similarly, Ogle & Correa-Kovtun (2010) have commented on the negative consequences for motivation that can result when learners’ academic self-identities are threatened by the challenges of content specific vocabulary (p. 532). For AELLS the challenges lie not only in increasing vocabulary or developing greater understanding of sophisticated grammatical forms; but increasingly in the cognitive demands of working with longer texts in terms of recall and synthesis (Swaffar, 2004).

Comprehension of complex continuous text is made more challenging at advanced levels by the need to identify the implications of a particular genre to convey information,
as well as an understanding of the multiple layers of interpretation within a larger, socio-historical context (Swaffar, 2004, p. 21). At the advanced post-secondary level, students may be challenged by the identification of culturally embedded practices within a given genre. For example, Swaffar notes encyclopedia articles are typically guided by particular features, such as the likelihood that this type of article will not include multiple voices that may challenge master narratives of the culture it serves (Swaffer, 2004, p. 23). Perhaps for just this reason, Kern suggests that advanced students may be able to develop their critical framing skills by simplifying complex narratives, thus forcing them to identify what they consider to be the essential elements of the story (Kern, 2004).

Interestingly, the CELBAN’s writing section requires IENs to do precisely this, i.e., summarizing the most important points of briefing, and writing a narrative report that effectively conveys the essential components of a scenario in nursing (CELBAN a, n.d.).

Comprehension of complex continuous text in a second or additional language is further compounded by the constraints of engaging with an unfamiliar text genre, especially one whose intended audience is necessarily limited (Swaffar, 2004). For example, gerontological palliative care articles do not anticipate a wide audience, thus they are typically written according to a particular style atypical of a more reader-friendly text. The challenges of coming to terms with academic language can be exacerbated by the anxiety and stress that accompany high-stakes testing in post-secondary academic settings (Snow & Uccelli, 2009), such as in high-stakes language testing. These forms of large-scale, high-stakes assessments tend to require writing in a persuasive or analytic genre, and are often graded according to criteria that refer to academic language forms (Snow & Uccelli, 2009). In these settings, AELLs are likely to benefit from the
background, or schematic, knowledge they have gained from membership within a specialized reader community, such as the principles of gerontological care, and nursing care—albeit in their L1 (Swaffer, 2004). However the conventions of the writing genre may provide new challenges where these conventions may not be shared by the AELLs L1. Some (Snow & Uccelli, 2009; Swaffer, 2004) suggest additional instruction in the generic forms of a particular context may be useful for AELLs. It may logically follow then, that assessments of proficiency in a genre specific area may be best directed to those conventions and style of that context, as they are with the CELBAN.

Integration Challenges of IENs

Most immigrants enter a period of low income earning soon after migration, and some will never reach the earning potential commensurate with their skill level (Blythe et al., 2009). For nurses, it has been reported that the longer they remain unemployed in their field, the more upgrading they will require, resulting in the greater likelihood that they will choose not to pursue nursing as a viable professional option (Blythe et al., 2009). One approach the Ontario provincial government has taken to address barriers to the economic integration of Internationally Educated Professionals (IEPs) has been to adopt bridge training programs, defined as those that:

Help skilled newcomers…find employment commensurate with their skills and experience in Ontario…in general they are designed to give internationally trained individuals: an assessment of their education and skill[s],…preparation for license or certification examination, occupation-related language training, and individual learning plans to identify any added training needed (Ministry of Citizenship and Immigration, 2013).

In 2008, the Ontario Regulators for Access Consortium (ORAC) published a manual with the purpose of “summarizing experiences with bridging programs in Ontario and to assist
bridging educators in continuing to develop and enhance the quality of their programs” (Austin, 2008, p. 3). In summarizing the key components of successful bridging programs in Ontario, ORAC has effectively established a key framework for assessing success which includes: partnerships between key stakeholders, integration of language support with training and professional skill competencies, providing context for the cultural conception of the professional role in Canada, and engaging employers in order to maximize post-program connections (Austin, 2008, p. 12). An Australian researcher, a medical doctor himself, has proposed XII necessary components for training programs for internationally educated medical graduates including: an in-depth understanding of the demographics of the International Medical Graduate (IMG) populations accessing these programs, an approach to education that involves building upon prior IMG learning and experience, an understanding of cross-cultural learning styles, and a need to look after the families of IMGs (Couser, 2012). One third of Couser’s recommendations focus on increasing awareness of cultural differences, and the necessity of thinking critically about the needs of IEPs who relocate their lives, and often families, in pursuit of international work. Such attention to the familial and financial stressors which can impact economic integration may help to shed light on some of the factors which may affect access to licensure and certification requirements such as high-stakes certification assessments which may require a proportionally large investment of time, and money for a family that has recently entered a “period of low income earning soon after migration” (Blythe et al., 2009). Perhaps out of a desire to address these factors which are irrelevant to testing or assessment constructs, both Couser’s views and those expressed in the Bridging to Success manual emphasize the central importance of IEPs in their role as key
stakeholders and partners, in programs and services intended to serve them (Couser, 2012; Austin, 2006).

**Research on the CELBAN**

Systematic searches of academic journals in Canadian nursing reveal very limited research on the topic of the CELBAN. As of July 2016, searches of the journal, *The Canadian Nurse* yielded two articles, one of which is not attributed to an author, and provides a link to the CELBAN website for further information (The Canadian Nurse, 2005). A French version of the same journal and article, again without attribution to an author and with a link to the CELBAN website is provided in *Infermiere Canadienne* (2005). The second article, appearing in 2006, mentions the CELBAN twice in two sentences describing the Canadian Nursing Association’s participation in the national advisory committee for CELBAN (Highlights from the 2005 Canadian Nurses Association Annual Report, 2006). A 2007 study addresses the CELBAN within the broader context of licensure of IENs in the province of Alberta, but the CELBAN itself does not constitute the research focus (Ogilvie et al., 2007). A search of TESL Canada, in July of 2016, yielded one result which cites a CELBAN document describing the Centre for the Canadian Language Benchmarks of internationally trained nurses (MacPherson, S., Turner, D., Khan, R., Hingley, W., Tigchelaar, A., Dustan Lafond, L, 2004).

One of the most comprehensive resources on the CELBAN may be a chapter written by two developers of the CELBAN (Epp & Lewis, 2009). This chapter appears in a wider text on creating a culturally inclusive nursing environment (Epp & Lewis, 2009). The limited number of research articles which addresses the CELBAN specifically is
indicative of the novelty of professional proficiency assessments which judge language competency specific to the performance of job tasks.

A presentation offered by one of the CELBAN developers, Lucy Epp, in 2006, at the Metropolis Conference in Lisbon, Portugal opens the discussion of the CELBAN as “A Model for Language Assessment in Canada” (Epp, 2006). Her presentation outlines the limitations of traditional ESL testing services for specific professions: “content of tests does not represent the workplace context, reliability and validity is not based on the target population,” and suggests these limitations can be mitigated through the use of a professional language competency test, of which the CELBAN is the foundational model. This point is emphasized by the statement made in the same presentation that “The Centre for Canadian Language Benchmarks has developed an occupation-specific English language assessment tool for nurses as a model for other professions and occupations” [emphasis added] (Epp, 2006). Epp’s comments at the 2006 conference emphasized that the use of tests like IELTS, TOEFL and MELAB for the purposes of assessing professional language competency would likely lead to construct underrepresentation: “Some who can pass the test, cannot function in the workplace—and others who take these language tests and fail, may have the potential to function successfully on the job” (Epp, 2006). The CELBAN has thus been designed with the intention of ensuring that those with professional competence are not unfairly disadvantaged by language testing which assesses language proficiency in subject areas unlikely to affect their work performance.

Ogilvie et al. (2007) emphasize that the very ability to “practice in one’s chosen profession can be a major factor in integration of immigrants into Canadian society and
thus a powerful factor in attraction and retention of highly skilled immigrants” (p. 223). While the CELBAN remains an under-researched area affecting the integration of IENs, several researchers have commented on the barriers to licensure faced by IEPs hoping to re-establish their careers in Canada (Blythe et al., 2009; Cheng et al., 2013; Myles et al., 2006; Ogilvie et al., 2007; Ngo & Este, 2006; Walton-Roberts & Hennebry, 2012; Yee et al., 2007). The development of the CELBAN was intended to ease some of the barriers to licensure faced by IENs in particular, as the assessment was intended to address the language demands of the nursing profession specifically (Epp & Lewis, 2009).

Validity, Reliability and the Testing Constructs

This paper relies heavily on the work of Messick (1991, 1996) to define test validity, and threats to test validity. This is due to the seminal nature of these papers as cited by (Fox et al., 2013; Haladyna & Dowling, 2004; DeLuca et al., 2013; Cheng & DeLuca, 2011). Validity may be understood as “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions based on test scores of other modes of assessment” (Messick, 1991, p. 1). Tests may be used for a variety of evaluative purposes. In the case of the CELBAN assessment, test results guide the decisions to progress or hold back an application of IENs for their licensure and certification.

Fundamental to test validation is the gathering of evidence either, or in some cases both, convergent and discriminant evidence (Messick, 1991). It makes sense then that the earliest steps concerning the construction of the assessment focused on the gathering of empirical evidence concerning the key language demands of on the job for nursing professionals. One of the phases of the development of the CELBAN entitled, “Analysis
of the Language Demands of the Nursing Profession” identified the following key language challenges in the nursing profession for IENs:

1. Speaking and listening identified as most challenging, with comprehensibility a major concern
2. Fast pace of the workplace
3. Phone interactions
4. Terminology and jargon
5. Need to ask for clarification
6. Use of appropriate tone and register
7. Need for assertiveness in certain situations
8. Ability to respond to negativity
9. Writing: charting and legibility
10. Non-Verbal Communication
11. Culture of the Workplace (Epp, 2006).

The construction of the CELBAN included “analysis of target language use, pilot testing drafts of the new assessment tool with the target population, and the use of rigorous measures of reliability and validity through statistical analysis of results” (Epp & Lewis, 2009, p. 287). Coefficient alpha reliability scores for the second draft of the CELBAN pilot were reported as: Speaking .944, Listening .939, Reading .965, and Writing .905. On this basis, CELBAN developers conclude that the CELBAN “clearly exceeds reliability criteria” (Epp & Lewis, 2009, p. 293). That several assessment aspects were measured is consistent with a form of evidence gathering described by Messick (1991) as convergent evidence: whereby several research findings are found to be in agreement supporting a particular claim or interpretation. Additionally, some efforts were made in the piloting phase to determine the “face validity” of the CELBAN as assessed by the test-takers (Epp & Lewis, 2009, p. 296), a methodological approach that is more consistent with qualitative research. Face validity is defined by the CELBAN developers as whether or not “the stakeholders believe that the test measures what it is supposed to
measure” (Epp & Lewis, 2009, p. 296). The survey questions included in the publication by Epp & Lewis (2009) assessed the test-takers perceptions of the construct validity of the assessment, but do not appear to address construct-irrelevant variance factors. It is important to note that validity is not a characteristic which is either present or absent, but rather a matter of degree to which validity evidence “becomes enhanced or contravened by new findings” (Messick, 1991, p. 2). For instance, if the face validity study had yielded results which contradicted the statistical reliability testing, then the evidence might be described as discriminant. What this means is that in a case where some evidence suggests one finding and other evidence suggests yet another contradictory finding, the inference that either measure has by itself captured reality may be discounted. While qualitative and quantitative studies draw from differing methodological approaches, the evidence which is yielded from these various studies may contribute to our collective evaluative judgments on the degree to which that evidence supports the actions taken based on the result of the assessment.

Qualitative researchers such as Denzin and Lincoln (2008) maintain that “objective reality can never be captured. We know a thing only through its representations” (Denzin & Lincoln, 2008, p. 5). While validity researchers such as Messick maintain that validity is neither entirely present or absent, the key issues of test validity are: “meaning, relevance, utility of scores, the import or value implications of scores as a basis for action, and the functional worth of scores in terms of the social consequences of their use” (Messick, 1991, p. 4). Therefore, the use of mixed methods described in the early construction of the CELBAN suggests its developers were concerned with how best to represent, and thus know their evolving creation, the CELBAN, and what the outcomes
of this assessment would mean for its stakeholders. However, as indicated above, one representation of the CELBAN which was absent from its piloting was the notion of construct-irrelevant variance factors, i.e., factors which represent a facet of testing validity. Test validation, not only in the case of the CELBAN but in all assessments, is an ongoing process of gathering and evaluating data. As such, validity evidence may be considered to be always incomplete, and therefore:

Validation is essentially a matter of making the most reasonable case, on the basis of the balance of evidence available both to justify current use of the test and to guide current research needed to advance understanding of what the test scores mean and how they function in the applied context (Messick, 1991, p. 2).

In the case of the CELBAN then, the approach to collecting evidence must be guided by the current balance of evidence available, which to-date remains to be primarily offered by the test developers themselves. Therefore additional research on the CELBAN would help to contribute to the evidence with which we might evaluate the current use of the test and its wider role within the certification and licensure process for IENs.

**High-Stakes Large Scale Assessment and Stakeholder Beliefs**

A 2014 study sought to contribute to the validity argument for the Test of English as a Foreign Language internet based test (TOEFL iBT) utilizing both focus groups, survey data, and stimulated recall to sample stakeholder beliefs concerning task authenticity to construct representation (Malone & Montee, 2014). The TOEFL iBT has experienced a shift in test item type from its earlier iterations, e.g., the paper-based test to the current format, e.g., internet-based which requires test-takers to demonstrate their knowledge through a series of integrated language skill tasks (Malone & Montee, 2014)—an approach which is shared to some extent by the CELBAN. In integrated
language tasks, academic language ability is assessed by blending skill areas to promote authenticity. For example, a test-taker’s reading proficiency may be assessed in part by their demonstration of comprehension through a writing task, as might be expected in a post-secondary educational environment. Malone and Montee explain that, “stakeholder beliefs about the correspondence between test tasks and the skills they purport to test provide important support for test validity” (p. 2). This is especially true in regard to the CELBAN where assessment tasks are intended to model target language usage of the profession. Where it is not clear to test-takers what target language usage they are meant to demonstrate in the completion of an assessment item, the faith we may have in the score reported by that assessment question—as it relates to their skill level may be called into question.

Validity: Consequences and Construct Irrelevant Variance

Researchers have noted that in large scale assessments there are intended as well as unintended consequences of test-takers’ test performance: “…test performance has consequences on educational promotion, employment policy, and programming decisions, and test-takers’ social-emotional health” (Cheng & DeLuca 2011, p. 104). Nagy (2000) has proposed that assessment serves three functions: gatekeeping, accountability, and instructional diagnosis. The CELBAN can be seen as serving each of Nagy’s assessment functions. In its capacity to determine “who is granted a privilege such as admission or graduation” (Nagy, 2000, p. 262), the CELBAN functions as a gatekeeper to licensure and certification ensuring only those that are deemed to have sufficient knowledge of Canadian nursing language are permitted to advance in their pursuit of provincial licensure and certification. As a secondary role, the CELBAN is
used as an accountability measure, “to judge the quality of education” (Nagy, 2000, p. 262), in particular to assess whether or not the international education of the prospective nurse in question is deemed adequate by Canadian standards. Lastly, the CELBAN may be used by bridging education providers with the IENs in pursuit of further language instruction to improve upon areas of professional language weakness as revealed through its role as an “instructional diagnosis tool” (Nagy, 2000, p. 262). However if bridging education programs were to use test results as a diagnostic tool to direct future learning, these actions may be seen as undermining the validity of the test itself by narrowing the curriculum. Nagy (2000) cites the work of several researchers who have investigated the influences of high-stakes testing toward “inappropriate teaching to the test…and the effects of teaching of test-taking skills” (p. 270). Nagy’s suggestion is that if a test is meant to measure the knowledge of a particular construct, then learning opportunities may be lost when the focus shifts from learning about the construct to learning about how to respond to questions about the construct in a given testing format.

Although research in assessment has shown that all tests vary in their ability to represent the testing construct to some extent, validity is also affected by the way in which the test-takers approach common test formats. For example, Rupp, Ferne, & Choi (2006) have offered evidence in support of the theory that test-takers approach multiple-choice questions as problem-solving tasks as opposed to comprehension tasks. When test-takers respond to comprehension tasks (the testing construct) in a manner consistent with process-of-elimination problem solving type strategies, the validity of that testing format may be called into question. Rupp et al. (2006) argue that where test-takers respond to a reading comprehension question in a manner inconsistent with their thoughts about the
passage at the time of their reading, their response is dictated not necessarily by their understanding, but by their understanding of the purpose of the question itself. Rupp et al. explain “Reading is first and foremost a purposeful activity…while the purpose for reading influences a reader’s type of engagement with a text” (p.444). Furthermore, Rupp et al. (2006) explain”

One of the main implications of a purpose for reading is that it guides readers in the selection of their strategies, the range of skills they draw on, and the intensity with which they draw on each skill…The main purpose of responding to MC questions about reading passages is, undoubtedly, to answer them correctly, and so test-takers select their strategies accordingly to optimize their chances for success (p. 447). This may have further implications for aspects of the CELBAN which ultimately intends to assess language skills through an occupation specific context and the familiar genre of nursing vocabulary. As Sasaki’s (2000) study has shown, test-takers who read a “culturally familiar cloze text tried to solve more items and generally understood the text better…” (p. 85). In particular what this piece of information may call into question is the potential for test-takers to attempt to problem solve beyond the construct of reading comprehension in order to select answers that they feel are more representative of their professional knowledge, what is in this case the “culturally familiar” (Sasaki, 2000, p. 85). Where test-takers may answer questions with the intention of demonstrating deep subject specific knowledge, as opposed to completing a reading comprehension assessment based on a given text, scoring ambiguity may result. Cheng & DeLuca’s (2011) observation then, that “…It is critical that the link between test validity and consequences of test use from multiple stakeholder perspectives within language assessment be established so that we can better justify the use of test scores in
pedagogical practice” (p. 105) is especially relevant in the CELBAN assessment completion context.

Researchers have observed that “anxiety is another well-documented cognitive phenomenon closely related to performance and achievement. Tests can cause anxiety, which may contribute to construct irrelevant variance” (Cheng & DeLuca, 2011, p. 107). Cheng and DeLuca’s (2011) study of the relationship between aspects of the testing experience and test-takers’ perceptions of test validity and use, revealed a number of construct irrelevant factors that can negatively impact the perception of validity within a high-stakes, large-scale testing environment. In total eight themes emerged from the study including: test administration and testing conditions, timing, test structure and content, scoring effects, preparation and test taking strategies, test purpose, psychological factors, external factors and consequences (Cheng & DeLuca, 2011). Test administration and testing conditions included the concepts of systematic and random biases. Systematic biases were defined as those that disadvantage the group of test-takers “by virtue of test administration protocols” (Cheng & DeLuca, 2011, p. 110), e.g., a limited ability to return to previous test questions for a computer based test. Random biases were cited more frequently than systematic biases by test-takers’ (Cheng & DeLuca, 2011). Random biases included “distracting environmental conditions, poor quality of test formats (e.g., low volume on tape recorders), and inconsistent invigilation protocols” (Cheng & DeLuca, 2011, p. 110). This study revealed that environmental factors were large contributors to test-takers’ perceptions of test validity (Cheng & DeLuca, 2011).

A 2013 study on the factors affecting student performance on the Test of English as a Foreign Language internet-based-test (TOEFL iBT) sought to identify potential issues
with the TOEFL iBT, including construct-irrelevant variance impacting the score variance of test-takers (DeLuca et al., 2013). For the purposes of the TOEFL iBT study, the researchers defined score variance as “the difference between the student’s true score and his/her observed score on the test as attributed to a variety of factors” (DeLuca et al., 2013, p. 663). Factors affecting score variance impact our understanding of validity, where it is defined as: “the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests” (Haladyna & Dowling, 2004, p. 18). In other words, our choice of what is measured as a factor of score variance, and what is not, can affect the claims made about the degree to which theory and evidence concur, or validate the testing instrument. High-stakes testing situations demand increased evidence of validity as the outcomes which follow the interpretation of a given test score can be life altering. Reducing or eliminating threats to validity, establishes greater confidence in the interpretation of test scores (Haladyna & Dowling, 2004). In particular, researchers have identified Construct Irrelevant Variance as a “serious problem in high-stakes testing” (Haladyna & Dowling, 2004, p. 17).

**Validity & Construct Irrelevant Variance**

Haladyna & Dowling (2004) cite the following discussion of Construct Irrelevant Variance by Messick: “…educational achievement tests, at best, reflect not only the psychological constructs of knowledge and skills that are intended to be measured, but invariably a number of contaminants” (p. 18). It is to be expected that a certain amount of anxiety may result for test-takers in high-stakes assessment environments such as the CELBAN, however researchers have noted that while it may be closely related to
performance and achievement, it may also contribute to construct irrelevant variance. (Cheng & DeLuca, 2011).

DeLuca et al. (2013) identified three themes which emerged from the focus groups conducted with key informants following the completion of the TOEFL iBT under the pattern Construct Irrelevant Variance: testing environment, test design, and score reports. Testing environment factors which were found to have contributed to construct irrelevant variance included: the amount of travel time required to access a testing location, the time of day at which the test began as well as its duration, ambient noise in the test taking facility, administration procedures which interrupted test-takers, as well as security protocols which led to personal discomfort e.g., requirement to remove shoes despite the cool temperature in the test taking room (DeLuca et al. 2013). Test design factors included: the effect of speededness (length of the test, or timing of complex tasks) on test-takers’ perceived performance, comprehension of instructions, consistency of test design patterns to signal important information and the ambiguity of distracters in multiple choice questions (DeLuca et al., 2013). Score report factors predominantly consisted of the ambiguity of the score report received by test-takers following the completion of the TOEFL iBT, leaving test-takers unclear on how they may have improved particular areas of their test taking performance in re-test, e.g., did the use of subject specific knowledge in an integrated reading and writing task that went beyond the bounds of the reading passage hinder a test-taker’s performance (DeLuca et al., 2013)?

The findings of Cheng & DeLuca (2011) and DeLuca et al. (2013) have emphasized the importance of: the testing environment, test design, and score reporting on the effects of construct irrelevant variance. These in combination with the additional eight themes that
emerged from the Cheng & DeLuca (2011) study constitute sensitizing concepts that will be applied within a wider deductive approach to the coding of focus group and interview data following the completion of the CELBAN see Table 2.

Table 4.
*Sensitizing Concepts*

<table>
<thead>
<tr>
<th>Cheng &amp; DeLuca 2011</th>
<th>DeLuca et. al., 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test administration and testing conditions</td>
<td>Testing environment</td>
</tr>
<tr>
<td>Timing</td>
<td>Test design</td>
</tr>
<tr>
<td>Scoring effects</td>
<td>Score reports</td>
</tr>
<tr>
<td>Test structure and content</td>
<td></td>
</tr>
<tr>
<td>Preparation and test taking strategies</td>
<td></td>
</tr>
<tr>
<td>Test purpose</td>
<td></td>
</tr>
<tr>
<td>Psychological factors</td>
<td></td>
</tr>
<tr>
<td>External factors and consequences</td>
<td></td>
</tr>
</tbody>
</table>

**Context-Sensitive Extrapolation**

Patton (2012) describes extrapolation as involving “logically and creatively thinking about what specific findings mean for other situations, rather than the statistical process of generalizing from a sample to a larger population” (p. 302). Of this approach the suggestion is made that Cronbach’s original intention for the concept of extrapolation was to interpret the findings “in light of stakeholders’ and evaluators’ experiences and knowledge, and then [further] applied or extrapolated using all available insights, including understanding about quite different situations” (Patton, 2012, p. 302). The intention of extrapolation then is to re-focus interpretation from the search for an absolute truth, to greater concern for conclusions that are “reasonable, justifiable, plausible, warranted, and useful” (Patton, 2012, p. 302).
Summary

CELBAN test-takers may be described as Advanced English Language Learners (AELLs) in that their language learning needs are guided by the depth and range of their understanding of their L1 as well as the support they require for L2 or greater learning motivation as unique from L1 learning experiences (Brown, 2007; Kern 4004; Ogle & Correa-Kovtun, 2010). Researchers of AELLs are re-evaluating approaches to teaching which prioritize content learning over language learning or vice-versa (Byrnes & Maxim, 2004; Kern, 2004). Promotion of either communicative language teaching or explicit, deductive methods of teaching receptive and productive genre specific English language skills focus on increasing the authenticity of the language learning experience for AELLs (Brown, 2007; Swaffar, 2004). Studies on English for Academic Purposes (EAP) and English as a Second Language (ESL) at the post-secondary level in Canada has demonstrated that English language programs, of either kind, have a significant, direct, positive effect on the academic and social engagement of advanced second language learners (Fox, Cheng, & Zumbo, 2013). Genre specific methods of teaching and assessment are particularly important for AELLs where comprehension of a complex continuous text requires not only understanding of the words themselves, but the implications of the selection of a particular genre to convey information (Swaffar, 2004).

As an assessment of professional language competency the CELBAN offers AELLs the opportunity to demonstrate their language skills through the completion of nursing genre specific tasks.

This genre specific language assessment opportunity is critical for evaluating the language competency of IENs, for whom it has been reported that the more upgrading
they require, the greater the likelihood that they will not choose to pursue nursing as a viable professional option (Blythe et al., 2009). Although to date, the literature on the CELBAN is somewhat sparse (N.A., 2005; N.A., 2006; MacPherson, 2004), its test developers have outlined the limitations of traditional ESL testing services in assessing the genre specific language skills of IENs (Epp, 2006). As the CELBAN offers IENs and the public a score by which to interpret nursing language competence, ongoing efforts to contribute validity evidence to the assessment and the scores offered by it, are required.

Other high-stakes, large-scale language assessments have been studied in efforts to contribute to the validity argument of their use, by analyzing test-taker data on their beliefs concerning task authenticity of construct representation (Malone & Montee, 2014; Cheng & DeLuca, 2011). Because factors effecting score variance, or perceived score variance impact our understanding of validity, it is essential to gather and analyze test-taker perceptions of their assessment experience to further our understanding of the “degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests” (Haladyna & Dowling, 2004, p. 18). In light of Patton’s (2012) description of extrapolation as “logically and creatively thinking about what specific findings mean for other situation…” (p.302), this study drew from the research reviewed here to inform the approach to collection and analysis of test-taker perceptions of the CELBAN.
In this chapter I describe my approach to the research, and the method I applied to address my research questions. This chapter provides a description of: the participants who volunteered to take part in this study, my recruitment process, the research instrument and piloting phase of my study, data collection, and data analysis. In the end, I present an overview of the themes and codes derived from the data analysis. The study described here received the approval from both the Education Research Ethics Board and General Research Ethics Board at Queen’s University to ensure minimal risk to both the researcher and the participants (see Appendix A).

Methodological Approach

A qualitative approach was used to gather and interpret the data of the study. This study focused on the Canadian English Language Benchmark Assessment for Nurses (CELBAN) and attempted to learn more about what CELBAN scores tell us from listening to test-takers who described how they “perceive it, describe it, feel about it, judge it, remember it, make sense of it, and talk about it with others” (Patton, 2002, p. 104). As a relatively new assessment of English language ability, noted for its unique approach to assessing English language skills required in a professional nursing context in Canada, I felt one of the best ways to understand what is communicated by CELBAN test scores was to learn from test-takers how they felt about the assessment, and how they felt it assessed their areas of skill. In this way I attempted to offer validity evidence for the CELBAN by “…making the most reasonable case, on the basis of the balance of evidence available both to justify current use of the assessment and to guide current
research needed to advance understanding of what the test scores mean and how they function in the applied context” (Messick, 1991, p. 2). Denzin & Lincoln (2008) tell us that “objective reality can never be captured. We know a thing only through its representations” (p. 5). This is as true of assessing English language skills as it is of understanding the CELBAN experience itself. In an effort to know, and to convey to others, the degree of mastery of English language skill of an IEN, we depend, in part, on a CELBAN score to represent that level of skill. In order to understand more of this representation, a CELBAN score, I have undertaken to review and analyze, qualitatively, the experiences of XVI CELBAN test-takers in order to advance our collective understanding of the CELBAN test taking experience.

This study employed a qualitative approach using focus group discussions and one-on-one interviews. Three, one-and-one-half hour long focus groups were conducted with XVI Internationally Educated Nurse (IEN) participants who were required to complete the CELBAN as a component of the licensure process in Canada. Additionally, nine one-half hour one-on-one interviews were conducted following the completion of test-taker speaking assessments. Focus groups and interviews used the interview guide adapted from a pre-existing tool utilized in the TOEFL iBT investigation of construct representation and construct irrelevant variance (DeLuca et al., 2013).

**Participants**

XVI IENs who were required to take the CELBAN as part of their licensure and certification process took part in this study. Six participants of a possible ten participated in the focus groups and interviews in Hamilton, and ten out of a possible forty-five participated in the focus groups and interviews in Toronto. All participants shared their
CELBAN experience with me in a focus group, interview setting or both. [Figure 1 below shows the breakdown of participants who shared their CELBAN experiences in either a focus group or an individual interview format.]

Due to the format of the group and individual assessments of the CELBAN, focus groups were intended to follow the group assessment (reading, writing, and listening). Individual interviews were intended to follow individual speaking assessments. Figure 1 demonstrates the number of participants who contributed to this study, their participation in a focus group, an interview, or both, as well as the location of their participation. Whole Report Interview refers to instances in which participants were unable to attend a focus group following their group assessment but shared their CELBAN experience with me following their group and individual assessments in an interview on the day of their CELBAN. Whole report focus groups refers to instances in which participants were unable to attend an one-on-one interview following their individual speaking assessment but shared their CELBAN experience with me following their group and individual assessments in a focus group on the day of their CELBAN. Individuals who offered a whole report (responded to questions on both the group and individual assessments) are not included in the tally of participants who completed both a focus group and interview, but are reported separately under whole reports. The unit of analysis in this study is the individual IEN, as the CELBAN does not operate in isolation, rather it is designed to assess the skills of individuals. Therefore information was gathered from individual responses on the research questions indicated by the Semi-Structured Focus Group and Interview Guides in Appendices B and C. In both Focus Group and Interview settings
these guides offered an overview of the questions and possible flow of conversation I anticipated prior to meeting with the participants.

**Figure 1**

*Focus Group and Interview Participation*

![Focus Group and Interview Participation](image)

Note 1 Whole Reports are not included in # participants who completed both

As the primary focus of this research was to examine test-taker perceptions of the testing-constructs and their experience with the assessment, to inform our understanding of inference made from CELBAN test scores, purposeful sampling was used to identify potential IENs preparing to write the CELBAN for the purposes of licensure and certification. Patton (2002) suggests that the “purpose of purposeful sampling is to select information-rich cases whose study will illuminate the questions under study” (p. 46). A purposive sampling technique was used to contact potential participants who met the criteria of having registered to write the CELBAN at an assessment centre in Ontario for the purpose of fulfilling a licensure and certification requirement. An Ontario organization which began administering the CELBAN in the Fall of 2014 assisted in the
recruitment process. Touchstone Institute Competency Evaluation Experts, formerly the Centre for the Evaluation of Health Professionals Educated Abroad (CEHPEA), distributed a copy of an e-mail inviting participation in the study as well as Appendix D Letter of Information and Consent Form to IENs registered to take part in the CELBAN as part of their licensure and certification requirements to work as a nurse in Canada, at the Hamilton and Toronto test centres in the Winter of 2015. The recruitment e-mail sent out to possible participants explained that interested individuals would be provided with a free pizza dinner as well as a ten-dollar Shoppers Drug Mart gift card in appreciation and recognition of their time and efforts spent contributing to this study. The total sample size was limited by the two data collection dates chosen to coincide with the CELBAN assessment date availability, and the availability of test-takers on their assessment date. Where individuals indicated interest in participating but were unable to attend a focus group or commit their time for the total duration of the focus group, they were invited to participate in an individual interview at a later time. Those XVI participants represented individual cases selected purposely to facilitate in-depth inquiry into the phenomenon of study: the CELBAN test-taking experience. Participation in the study was voluntary. A letter of information and a signed and returned consent form was used as consent to participate in this study (see Appendix D).
**Instrument**

**Focus Group and Interview Guide.**

The instrument used in this study (Appendices B & C) has been adapted with permission from DeLuca, Cheng, Fox, Doe and Li (2013). This instrument was used to establish a semi-structured interview guide for both the focus groups and individual interviews (see Appendices B & C).

Table 6 provides an overview of the Focus Group and Interview Guides and how they were intended to align to the research questions of this study. Section 1 of the focus group instrument, and the beginning of the Interview Instrument were designed to focus on participant experiences with the CELBAN overall. General perceptions were considered to be useful in answering the first research question regarding how IENs characterize their assessment experience. The second section of the instrument asked more specific questions regarding the assessment itself, encouraging participants to describe those sections of the assessment which tested their English language skills. This section was intended to help me understand how these IENs describe the constructs of the CELBAN. In the focus group interview guide the question on technical difficulties and timing, as well as the question in the interview guide on test-taker awareness of CELBAN administrators or assessors in the room were intended to identify what, if any, possible sources of construct irrelevant variance may have impacted the test-taking experience. Lastly, both the focus group and interview guides asked questions about whether or not the test-takers felt the assessment tasks provided an accurate representation of the types of communicative tasks they would have to complete in a
Canadian nursing setting. These questions were intended to address the last research question focused on task authenticity.
Table 5.
Research Questions, Focus Group and Interview Guide Alignment

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Sample Focus Group and Interview Guide Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) How do IENs characterize their CELBAN experience?</td>
<td>Section 1 of both Focus Group and Interview Guides ex. &quot;Tell me in general about your experience with the CELBAN&quot; and &quot;Describe your experience with the oral interview&quot;</td>
</tr>
<tr>
<td>2) How do IENs describe the constructs (reading, writing, listening, and speaking proficiency) measured by the CELBAN?</td>
<td>&quot;How do you feel the assessment tested your English language skills?&quot; and &quot;How do you feel the oral assessment tested your English language skills?&quot;</td>
</tr>
<tr>
<td>3) What, if any, potential sources of construct irrelevant variance do the test-takers describe based on their CELBAN experience?</td>
<td>&quot;Were there any technical difficulties during the test?&quot; and &quot;Were you aware of any observers in the room while you completed the assessment?&quot;</td>
</tr>
<tr>
<td>4) Do IENs feel that the assessment tasks provide an accurate reflection of the types of communicative tasks required of a nurse in Canada?</td>
<td>&quot;Do you feel the exam tasks were similar to what you would have to do as a nurse while working in Canada?&quot; and &quot;Do you feel these types of tasks are realistic in terms of the interactions you have in a Canadian hospital?&quot;</td>
</tr>
</tbody>
</table>
Pilot Study

To prepare for the data collection phase, a semi-structured interview guide was piloted with an Advanced English Language Learner who is also an Internationally Educated Professional with previous experience with large-scale, high-stakes English language testing for the purposes of admission into a post-graduate educational program in Canada. McMillan & Schumacher (2010) have advised that “a pilot test is necessary as a check for bias in the procedure, the interviewer, and the questions” (p. 206). The pilot encouraged revisions to the order, structure, and clarity of language in the interview guide, particularly for lengthy questions which the pilot participant noted as being challenging to respond to. For example, the second question in the pilot interview guide “How did each language skill section of the test provide an accurate reflection of the types of communication demands faced by nurses in Ontario?”; was revised to be stated more simply as “Do you feel these types of tasks are realistic in terms of the interactions you have in a Canadian hospital?” Reflections on possible sources of bias in the piloted questions lead to revision of the piloted question “Can you describe the security procedures of the test?” This was changed to a more neutral “How would you describe the security procedures of the test?” Transcriptions of the interviews and focus groups revealed that in the data collection phase this question regarding security procedures was rephrased to a far more neutral “Tell me, what are the steps you take to write the test when you arrive? What do you do first?”

Piloting the semi-structured interview guide, and following the evolution of question types is an effort to achieve greater rigour in my approach to qualitative research as it is defined by Anafra, Brown & Mangione (2002), “the attempt to make data and
explanatory schemes as public and replicable as possible” (p. 28). I believe that piloting the questions prior to collecting the data from participants in the Hamilton and Toronto test locations better prepared me to think about asking open ended questions, that allowed the test-takers to speak about their experience, as opposed to merely answering questions that I was interested in for the purposes of this study.

**Data Collection**

Data collection was comprised of three, one-and-one-half hour long focus groups as well as nine one-half hour individual interviews in the winter of 2015. Data collection took place at the St. Charles Adult Learning Centre in Hamilton, and the Touchstone Institute in Toronto, Ontario on the day of test-takers’ CELBAN assessment. Due to the format of the assessment’s two-part structure (a group assessment for the reading, writing, and listening sections and an individual assessment for the speaking section of the assessment) focus groups were intended to follow the group section of the assessment with individual interviews following the speaking section. In Hamilton, one participant offered their views on the entire CELBAN experience in an interview setting, as opposed to describing only their oral assessment. This was due to the scheduling of their speaking assessment which occurred at the time of the Hamilton focus group. Similarly, in Toronto three participants described their entire CELBAN test taking experience in the format of a Focus Group. This occurred as result of their speaking assessments being scheduled in the morning while the two prior focus groups, and their group assessment was scheduled for the afternoon, at which time I was conducting more one-on-one interviews.
Participant Demographics and CELBAN Readiness

Appendices E, F, and G provide detailed demographic data on the participants in this study. Appendix E describes participants by sex, age, country of origin and first language learned. Appendix F describes participants by the language they speak at home, the country the participant received their nursing credentials in, the language of their nursing instruction, and the number of years the participant has been studying English for. Appendix G describes participants by the prior language assessment they completed by name, the language levels reported by that prior language assessment, whether they are currently working in a health care related field, and whether they are working part or full time.

Eleven females and five males took part in this study all of whom self-identified their age range as being between twenty-to-thirty and thirty-to-forty years old. The majority of the participants in my study were from India originally (11) and the Philippines (2). Most had studied nursing in India (9) and the Philippines (2) (See Appendix A). All except one participant identified the language of their nursing instruction as being English, the one who did not, left that response blank. Eleven of the XVI participants identified a language other than English as the language they speak when at home. The number of years of nursing experience ranged from one-to-twelve years prior to coming to Canada (see Appendix F). Participants, except for one who left the response blank, identified the following as prior English language assessments they had completed: The Canadian Language Benchmark (CLB), the Canadian English Language Proficiency Index Program (CELPIP), and the International English Language Testing System (IELTS) (See Appendix G). Participants were asked to indicate their
prior assessment levels in reading, writing, listening and speaking. Participant responses were reviewed and summarized below. Although not all participants completed the CLB assessment, these represented the reported highest and lowest participant scores. The CLBs define the English language levels assessed by the CELBAN, as such they provide a good overview of how prepared participants may have been to complete the CELBAN at the time of their assessment. It is interesting to note that the highest self-reported prior assessed listening score was not high enough to constitute a pass on the CELBAN. See Table 5 for an overview of reported and minimum pass levels on the CELBAN.

Table 6.

<table>
<thead>
<tr>
<th>Assessed Skill</th>
<th>Minimum Self-Reported CLB Score</th>
<th>Maximum Self-Reported CLB Score</th>
<th>CELBAN Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing</td>
<td>4</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Speaking</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Listening</td>
<td>7</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Reading</td>
<td>6.5</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

Note 1. As disseminated by the College of Nurses of Ontario's website updated September 10, 2014
Note 2. These are the minimum CELBAN benchmark scores that must be achieved and includes the standard error of measurement
Note 3. The College will only accept scores from a CELBAN assessment at a CELBAN administration site. The College does not accept results from CELBAN assessments administered by educational institutions for educational purposes.

Focus Groups & Interviews

The combination of both focus groups and individual interviews benefited the data collection process in a variety of ways. Focus groups are fundamentally a way of listening to participants and learning from them as a group (Krueger & Casey, 2000) as they share a common experience. In this study, the common experience shared by
participants is the CELBAN test-taking experience. The inclusion of the focus group method offered the opportunity for participants to interact with a small group of their peers to share their thoughts, feelings, attitudes, and ideas about their experiences with the CELBAN. This is particularly relevant for a research focus which asked participants to consider aspects of the assessment completion process which were not their primary focus during the assessment. The ability to speak in a group setting allowed participants the opportunity to offer a more detailed account of the experience as other participants shared experiences which were similar or different than their own. The individual interviews which took place following the speaking assessment provided participants an additional opportunity to contribute thoughts or feelings they may not have shared in a group setting.

Focus groups and interviews took place in unoccupied classrooms or assessment rooms at the St. Charles Adult Learning Centre in Hamilton, and the Touchstone Institute in Toronto, Ontario on the day of test-takers’ CELBAN assessment. Although a more neutral location is often selected for this type of data collection, in an effort to ensure that participants were afforded the greatest flexibility in their individual assessment times I felt it would be most comfortable for participants to remain at the assessment site. This offered a degree of convenience for participants in that they were not required to exit the assessment centre to eat dinner and rush back, as such, part of my recruitment process included the offer of a free pizza dinner as well as a ten dollar Shoppers Drug Mart gift card which was given to every participant. In each session, for which I had not previously met the participant or participants, I began with an explanation of the purpose of the study, information about the recording of the focus groups and interviews, and a request
for the participants to complete a brief demographic survey (Appendix H). I recorded each focus group and interview on two different recording devices so that in the subsequent analysis phase, I was able to review a secondary recording of the sound was muffled or distorted on one. I began my questions in both the focus groups and interviews with general questions focusing on the CELBAN assessment experience overall example: “How would you describe your test taking experience today?” Both focus groups and interviews followed the established interview guide adapted from DeLuca et al. (2013), and were audio-recorded and transcribed verbatim to facilitate analysis. In total six hours of focus groups and interviews were recorded and CXVII pages of single spaced, Times New Roman, XII point font, transcript pages were collected.

Data Analysis

Focus group and interview transcripts were coded deductively via the sensitizing concepts derived from Cheng & DeLuca (2011) and DeLuca et al. (2013) on examining of test-takers’ experience and their contribution to validity. The deductive coding method “starts with an idea or theoretical framework and uses the data to verify or disprove the idea” (Schadewitz & Jachna, 2007, p. 5). Using a theoretical framework has been noted to be especially useful in allowing researchers to be guided by the sensitizing concepts which further “sensitizes the analyst[s] to the possibility of a category or behaviour that either has been overlooked in the data or is logically a possibility in the setting but has not been manifested” (Patton, 2005, p. 470).

In total, eleven sensitizing concepts derived from Cheng & DeLuca (2013) and DeLuca et al. (2011) were used to guide the deductive coding of focus group and interview transcript data: testing environment, test design, score reports, test structure and
content, preparation and test taking strategies, test purpose, psychological factors, external factors and consequences. Data analysis was conducted first by myself, the researcher, alone. I reviewed the transcripts and applied codes guided by my sensitizing framework. Where a concept or description central to one of my questions was not addressed by a code based upon my sensitizing framework, I created a new code to address this concept. I then reviewed previously coded transcripts to see if this concept was described there as well.

In total, XXI codes emerged to describe the data. I reviewed these codes and grouped them under broader themes which more broadly described those codes. I reviewed my codes and thematic framework with my supervisor, and aligned my themes to the guiding research questions of this study. In total, 7 themes describe the data. After having deductively applied codes with the guidance of my sensitizing framework, and grouped these under themes, I reviewed my coding and thematic structure with another Masters of Education candidate to explore the opportunity to more specifically define my themes and the appropriate grouping of codes under those themes. At this time all identifying information was removed from the transcript sections we reviewed and pseudonyms were applied to the data to ensure anonymity of data. Following this review, the structure outlined in Table 6 was determined to best represent the data according to the themes outlined there. This approach to collaboration on the review and analysis of codes allowed for additional refinement of the way in which I defined the themes of the data and thus was intended to offer evidence of efforts toward establishing inter-rater reliability. In light of Weitzman’s (2000) caution regarding the use of auto coding programs that, “We may fail to check to see what passages were actually coded in the
autocoding process, and to use our own intelligence to analyze whether they fit” (p. 808), I used NVivo 10 to facilitate the recursive process of applying research driven themes to fit deductive codes.

Summary

As I prepared to collect and analyse my data I took several steps to improve the quality and credibility of my findings. In relying upon a previously utilized instrument for data collection, I drew upon best practices laid down in prior research. I adapted this data collection instrument for the purposes of my own research, and piloted this tool to better prepare me to ask relevant, and appropriate questions of my participants. This pilot revealed a need to revise questions for simplicity, order, and bias. This revision to the data collection tool also sensitized me to the need to ask open ended questions in the data collection phase.

After the data collection phase, reviewing my codes and themes with my committee helped me to consider the alignment between the critical themes that emerged from my research and the guiding research questions of my study. Furthermore, taking the opportunity to review my codes and thematic structure with another Master of Education candidate allowed me the opportunity to consider if I was applying codes in a manner unique to my own perceptions and personal bias, or if my codes and themes aligned to the thoughts of other researchers. Reviewing this structure with a colleague was another important step in my research for which I engaged with a process of inter-rater reliability to add to the trustworthiness and credibility of my findings.

During a review of my research analysis with my committee it was noted that a group of participants were frequently quoted. Although XVI participants offered their
CELBAN experiences in total, the results and discussion draw upon input from XV of the XVI participants with some such as Jake, Sandy, and Penelope appearing more frequently. This increased frequency is a result of the in-depth, detailed descriptions these participants engaged in during the interviews and focus groups which touched upon those themes critical to my study in particular. My intention in selecting quotations for inclusion in this study was to provide evidence of the types of comments and thoughts shared by participants which most significantly comment upon a critical theme shared by other participants.

After review and reflection with my committee on those comments included in this study, I returned to my transcripts to review the thoughts shared by all of my participants to ensure that I had not failed to include a divergent position or viewpoint that might impact our understanding of the CELBAN experience from this group of participants. Although a direct quotation appears from all participants with the exception of Sam, that is not to suggest that Sam’s comments were not reviewed in detail or that his thoughts are not included in this study. Rather, I found that the views of this participant were represented by the data included here and that the quotations included below.

In the following chapter, I will discuss the major findings from my focus groups and interviews and how the quotations included below contribute to our understanding of the CELBAN, its testing constructs, possible sources of construct irrelevant variance and the authenticity of the assessment. In chapter five, I will interpret these findings and discuss them in relation to prior research on high-stakes large-scale AELL testing.
Chapter 4

Results

In this chapter, I will present the results of all data collected during the study. I have organized the results of this data collection by the 7 themes which I feel describe the findings of this research. As noted in Chapter 3, twenty-one codes emerged from the data. From these codes, seven broad themes emerged. Themes and codes are grouped in Table 7 below according to their correspondence with the four research questions of this study.

Frequency of the occurrence of a code is included in Table 7 to denote the commonality of these shared experiences, however frequency was not used as a means by which to arrive at the themes. For instance, comments made by participants were coded under psychological factors 108 times. However, the most commonly occurring word that might lead participant comments to be coded under psychological factors was ‘stress.’ For the purposes of my study I was most interested in learning more about the specific sources of stress, especially in regard to sources which may have impacted the participants’ ability to demonstrate their knowledge and skills in the assessment. As my results and discussion will elaborate on, a more significant aspect of the psychological factors, or stresses which may have impacted the test taking experience, was the challenge or stress related to locating resources which clearly described the expectations and format of the CELBAN.
Table 7.
Research Questions, Themes, Codes & Frequencies

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Themes and Codes</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How do IENS characterize their assessment experience?</strong></td>
<td><strong>Theme 1 Descriptions of the CELBAN</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 1 CELBAN assessment as preference</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>• Code 2 Ease</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>• Code 3 Challenges</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>• Code 4 Discomfort</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>• Code 5 Assessment aspects described positively</td>
<td>96</td>
</tr>
<tr>
<td><strong>How do IENS describe the constructs measured by the CELBAN?</strong></td>
<td><strong>Theme 2 Consequences of the CELBAN Test Scores</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 6 Next steps for self</td>
<td>12</td>
</tr>
<tr>
<td><strong>What, if any, potential sources of Construct Irrelevant Variance do the test-takers describe based on their experience</strong></td>
<td><strong>Theme 3 Skills Tested by the CELBAN</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 7 Language skills assessed</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>• Code 8 Nursing skills assessed</td>
<td>61</td>
</tr>
<tr>
<td><strong>Do IENS feel the assessment tasks provide an accurate reflection of the types of communicative tasks required of a nurse?</strong></td>
<td><strong>Theme 4 External Factors Impacting Test Experience</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 9 Immigration and CELBAN</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>• Code 10 Psychological factors</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>• Code 11 Licensure process</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>• Code 12 Family pressure</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>• Code 13 Financial stressors</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td><strong>Theme 5 Preparing for CELBAN</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 14 Knowledge of CELBAN</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>• Code 15 CELBAN resources</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td><strong>Theme 6 Assessment Administration and Conditions</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 16 Timing</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>• Code 17 Fairness protocols</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>• Code 18 Assessment structure and design</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>• Code 19 Administrative challenges to completion</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>• Code 20 Exceptional administrative organization</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td><strong>Theme 7 Authenticity</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Code 21 Cultural Differences</td>
<td>5</td>
</tr>
</tbody>
</table>
In some instances less commonly shared views may be represented by a low frequency of coding in the data, for example the comments made on cultural differences (see code 21 in Table 7), but the significance of these comments merit their prioritization as a significant contributor to a theme as it sheds light on a given research question. Drawing on previous literature which has used qualitative data analysis to derive key themes indicative of individual and systemic level barriers to integration as perceived by Internationally Educated Professionals (IEPs) (Yee et al. 2007), this study attempts to “allow the primary data to speak as much as possible on their own [in order to enable] the reader to see how the research team [in this case myself, my committee and my Master of Education colleague] saw the data” (p. 14). To protect the identity of participants, each individual was assigned a pseudonym as indicated in Appendices E, F and G.

The results described here are a reflection of the thoughts and feelings shared with me by the participants over two days of data collection at two test centres, one in Hamilton and one in Toronto. The participants shared their personal experiences and thoughts on: the assessment generally, the impact their results would have on their lives, the skills they felt were tested by the assessment, factors external to the test which may have impacted their demonstration of their language skills, their approach to preparation for the assessment, their thoughts on the administration and conditions of the test, and lastly, the authenticity of the test.

**Theme 1 Descriptions of the CELBAN**

Key to my study was a desire to know more about how test-takers perceive the CELBAN and their test taking experience. I wanted to know about the aspects of the test which they found to be self-affirming in terms of their professional language skills, and I
also wanted to know about the aspects of the assessment which they find to be challenging. I began every conversation with a participant by asking questions about how they would describe the CELBAN. This section contains a general overview of the kinds of comments made about the CELBAN test experience generally.

Mary’s response was positive overall, she explained that in her opinion the CELBAN:

…asked us to write something we are used to writing. Because we write nurse’s report[s]. Always we document things-so that’s something nice- because they’re not asking you to explain a graph…which is what exactly happens in IELTS where they give you some population graph or things which- it’s ok it’s good- but it’s not as related to nursing whereas here everything was nursing.

Aspects of the CELBAN which were described positively, were often also described as being comfortable, or putting the test-taker at ease because the context of the assessment was one which was familiar to IENs, i.e., communicative tasks related to patient care. During the focus groups and interviews, many participants described the CELBAN in positive terms and on many instances also contrasted their positive experience with the CELBAN with less positive prior language proficiency tests that are did not focused on nursing specifically. Mary’s positive comments on the CELBAN function as an acknowledgement of her sense of fairness of the assessment as she explains her opportunity to demonstrate her English writing skills were contextualized within her role as a nurse. While the CNO accepts specific language scores on either the CELBAN or IELTS as evidence of language proficiency (CNO, 2013), test-takers like Sandy cited either concerns that the IELTS was too general a language exam, or appreciation for the fact that the CELBAN is specific to the communicative tasks of a nurse. Sandy described the CELBAN positively in contrast to her experience with IELTS saying: “The scenarios are much…easier, and from your life experience, patient care area. IELTS is different.”
On a different CELBAN assessment day in Toronto, Emily’s comments agreed with those of Mary in Hamilton saying: “In general I think it was very good because it’s all from medical from nursing content. It’s everyday hospital scenario. So, I think it’s a good way to assess nursing language, English language for nurses.”

There were test-takers who had less positive experiences. For example Penelope, who was unsure of where the CELBAN really fit in terms of her overall certification process, she explained:

I think in Ontario they d[o]n’t require [you] to pass the English exam first but—in other province[s] before you start studying your nursing you have to clear your English exam…if you clear English that will also help with your nursing exam as well… I already clear[ed] all [my] exam[s] for nursing—I didn’t clear this exam. If I didn’t clear this exam there is [no] meaning of [the] nursing exam and that is all a waste of time and money. So I think CNO should [make it a requirement to] take this exam first and then the nursing exam.

Test-takers are under a great deal of pressure when they come to write the CELBAN assessment. During my interviews and focus groups, test-takers described a variety of challenges and sources of discomfort related to the assessment or aspects of the assessment. Some challenges are discussed in greater detail as under the section on External Factors Impacting Test Experience.

While the point made by Penelope above is not one that was repeated by many other IENs with whom I spoke, it is significant in the context of our wider understanding of how the CELBAN is perceived by test-takers, particularly as Penelope points out, in the province of Ontario. Internationally Educated Nurses are already nurses in the countries in which they were trained and employed previously. When they commit to the certification process in Canada, their intention is to resume their work in their field of
practice. Penelope’s comments are incredibly important because they highlight how the wider certification process can affirm or undermine one’s professional expertise.

For Penelope, a “Pass” on her professional nursing assessment, administered and written in English, was not enough to signify her eligibility to work in her field, as she had not yet passed her English language exam. Her comments suggest that, in her view, the certification experience as a whole could be improved by establishing the English language assessment as a pre-requisite to the professional competency exam to maintain the face-validity of the overall certification experience of IENs. Penelope’s prior certification exam results at the time of our meeting may indicate that while she has the professional nursing knowledge to practice, her English communication skills at the time of her first CELBAN attempt were not at a level required for a certified nurse in Canada. However, in the eyes of this test-taker the experience feels like one in which she must question the meaning of her result on the professional nursing exam itself.

To Penelope, a “Pass” on her certification exam may have “no meaning” because without a pass on her language exam she cannot be certified to work as a nurse in Ontario. It may also have “no meaning” because although the professional certification exam results were a pass, these results conflict with her prior language exam scores, suggesting she would not be able to successfully practice in her profession in Ontario as the professional certification exam is intended to indicate.

**Theme 2 Consequences of the CELBAN Test Scores**

It is through the work of Samuel Messick (1996) that that we are guided to think of validity not as a property which is either present or absent from an assessment but rather as the meaning or inferences we make from test scores. For this reason, I feel it is
important to know more about the consequences the reported CELBAN scores will have on the lives of the test-takers I spoke with. I asked test-takers questions about what they would do if they passed their CELBAN. This section includes the responses that I feel best illustrate the impact of the inferences that are drawn from the CELBAN scores of the test-takers I spoke with.

Jake came to Hamilton to complete his second attempt at the CELBAN, he explained:

There’s nothing you can do it is—it is always on the CELBAN…Today I was supposed to write at the end the exams [RPN] in Winnipeg—but I was not allowed to because I missed the listening and then the writing part—the last one. My listening was 9 but I need a 10, my writing was 6 but I need a 7. So they tell me I can’t write today so that’s why I’m here—I’m doing it and if all goes well the next…RPN exam will be in May…I’ve already done the bridging program. Just waiting for the CELBAN-If it goes through…then I will start living in Canada. Because I’m not surviving, living here in these kinds of jobs—like minimum wage jobs—you’re not living you’re just surviving here in Canada.

Listening to my participants describe and characterize their test experience a common theme which emerged was the consequence of the CELBAN test score on test-takers’ life. At no time, was the consequence of the test score more powerfully described than when Jake described his survival in Canada to date. These comments contextualize the CELBAN testing experience as one of many opportunities that IENs will have to demonstrate their competency to gain certification in their profession. As Jake explains the minor consequence of a CELBAN score may mean, for some, the need to re-take the assessment, as he did himself. On a deeper level though, the consequence of the CELBAN score is the difference for Jake, between survival and life. To survive, without the opportunity to work in his field is not a life for Jake. He characterizes the CELBAN as one of a series of exams required for certification, with his next one already lined up at
the time of our focus group. These exams are intended to summarize and convey the skills of an individual who has entered Canada with expertise in nursing from their country of origin. It is important to note that in the case of Jake, he described having already completed a bridging program. This tells us that as a test-taker Jake has already invested significantly in both time and money to show equivalence where possible in his education and to close gaps in his educational transcripts. The CELBAN as one of the professional exams remaining for his certification, may then be understood to represent to a significant marker in his progress and pursuit of the rights and responsibilities of working as a nurse in Canada. While the assessment itself is intended to assess the professional language skills of test-takers, the consequence for test-takers who take the test in various parts of the country may be temporal, financial, and professional.

Jake explains that as a result of a past failed attempt at the CELBAN he decided to complete the assessment again. However, as the professional certification process is governed by a tight timeline for which all certifying exams and processes must be completed within a pre-determined duration of time, test-takers must strategize test dates and locations which will allow them to demonstrate their competency within the allotted time. As the timeline for certification comes closer to an end for each IEN, temporal and financial consequences may be increased as they seek out additional opportunities to show competency. In the case of Jake, the CELBAN would not be offered in his province at a time that would allow him the flexibility to make his deadlines to complete his next assessment; and thus, it is the CELBAN, for this test-taker, that is the gatekeeper to live a life of professional work in Canada. Jake explains further:
“Unfortunately, the College of Nurses, Practical Nurses told me in October they’re going to close my file in February 15\textsuperscript{th}—that means that if today the CELBAN is not successful that will be the end of- (laughter) ya”

“You will pass.” (Sandy)

“Ya the nursing career. Then I will look at other things. But I would encourage everybody to keep trying so long as you still have the time, ya.” (Jake)

To understand how test-takers characterize their CELBAN test experience it is important to understand the most extreme consequences of the test score. In the case of Jake the consequences of the test were substantial enough to place him, a new Canadian, working a minimum wage job on an airplane from Manitoba to Toronto, and later ground transit from Toronto to Hamilton to complete his second attempt at a professional language assessment. So extreme are the consequences of this assessment in the lives of the test-takers, as to be characterized by Jake as an experience which could end his nursing career in Canada. It is within this context of time, testing and opportunity that test-takers consider the impact their CELBAN score have on their next steps, and in the case of Jake it is described as the difference between survival and life.

**Theme 3 Skills Tested by the CELBAN**

In my research on the CELBAN, I asked questions about what kinds of skills the participants felt each section of the assessment was testing. The responses below capture participant responses which described the CELBAN as accurately and fairly assessing the language skills required of a nurse, as well as instances where test-takers described uncertainty in how to respond to assessment questions.

When asked about which skills he was tested on in the reading section Jake explained:
…there was a part of the skimming and scanning where…I was able to look for specific information from the information given and was also…able to extract information from the passage, which is…good because it will test my knowledge in skimming and scanning and comprehension also.

Participants frequently identified specific language skills tested by the assessment and associated these skills with particular questions, question types or sections of the assessment. That test-takers saw a relationship between assessment sections and language skills suggests congruence between the types of tasks assessed by the CELBAN and what Messick (1996) deems “the content aspect of construct validity” (p. 248). This piece of information is critical to our understanding of authenticity and fairness in the assessment of language skills because the intended purpose of the CELBAN was to specifically address the language demands or “target language use” of the nursing profession, nationwide (Epp & Lewis, 2009, p. 285).

Jake made an important distinction between the two types of reading tasks in the quote above. Firstly, he specified that one skill tested by the CELBAN was to identify key pieces of information from a reading selection, the skimming and scanning section. He then differentiated this reading skill, identification, from general comprehension, in which he specified the test-taker must not just find information from the text but, “extract information from the passage” (Jake). The importance of this skill is confirmed also by Sandy who explains:

In reading I would say that that in our career we have to go through different articles and journals…so this was really a good method…because we had to comprehend, we had to skim for the important information…so if we had to go to [a] patient’s history we can be able [to] understand [the] patient’s history exactly. So this was really a good way to…identify one’s skills for understanding comprehension in terms of reading.
This comment highlights how closely language skills are intertwined with the performance of the tasks as Sandy identifies not only the skills tested, scanning for significance and comprehension, but also how integral these skills are to the performance of daily activities, reading patient history and ongoing professional development through professional reading and research.

At times, discussions with participants revealed the complexity of assessing language skills which are so deeply embedded in the performance of empathy and patient satisfaction (akin to customer service). This complexity was often highlighted in discussions with participants on the oral assessment, which included a nurse-patient role-play. Test-takers often spoke positively of the oral assessment. Some of their comments highlighted an appreciation for its approximation of authenticity, particularly in how well the oral assessment captured the difficulties that arise when communicating health care instructions. However, test-takers with whom I met did not mention an appreciation for this component’s assessment of their technical oral language skills such as: “forming open ended questions, using intonation for clarification, selecting the correct tense” (The CELBAN Centre, 2015). For example:

When examiners say, “my patient died,” “my father died because of stroke when…I was like one year old” so I say, “Oh I’m very sorry to hear that, it must be very hard for you.” Then she was making expression like she was very sad. And so I was feeling like, at the time I was feeling like, she is examiner why is she feeling like that because—this is exam so I have to empathize right. “I’m sorry”—like that. But she showed the real feeling so I liked that…. Then after she say[s] “oh thank you for acknowledging my feelings, I appreciate this” and so she respond[ed] to me like this one. So it’s like a little bit reality—not much (Penelope).

Penelope described her confusion when the assessment required a demonstration of language skills that simultaneously identified knowledge of cultural competency, in terms
of the expected performance of the nursing role in a Canadian context. Although official
documents from the CELBAN Centre indicate the intention of the oral assessment is to
test the ability to, “ask questions to gather information…Use grammar accurately to form
questions…as well as asking direct and precise questions to elicit clear answers” (The
CELBAN Centre a), the direction of the role play indicated by Penelope suggests that the
assessor was prompting the test-taker, not for a question—but for statements of empathy
and understanding—a point which Penelope admits gave her a moment of pause
regarding which test criteria she was expected to meet through this interpersonal
challenge. It is useful to note that in reaction to this experience Penelope indicates her
suspicion that the assessor was looking for an expression of empathy, and later indicates
that perhaps this was also intended to provide an authentic effect. Demonstrating an
awareness for linguistic expressions of cultural competency which are required to ensure
patient satisfaction, represents an entire set of professional skills in and of themselves.
Although the CELBAN is intended to assess language skills specifically, the nursing
specific context of the test at times offered moments of confusion to this participant.

Some of the test-takers with whom I spoke described sections of the CELBAN
which they felt tested their nursing skills. For example Sandy explained:

…for [a patient that has] rheumatoid arthritis…there was a blank, what should we
put under the knee? Like we have to select either knee, wound, or what. So if I
was clear about rheumatoid arthritis, [it] is [a] joint problem…there’s no
wound…if I know that this is joint problem that’s why I selected knee…If I was
not familiar with rheumatoid arthritis as a disease…I might have thought it’s a
cardiac disease or something like related to any wound so I might have selected
the wound. I might have selected the wrong thing. Or pillow should be under the
knee or under the head. Like because it’s a joint disease so I selected the ah one
pillow should- shouldn’t be under the knee because it’s a joint problem- knee’s a
joint- so because of my nursing knowledge I feel I selected the correct thing, the
correct choice.
In the quotation above, the participant describes her decision making process in responding to a specific assessment question. In this instance she explained that her selection of a possible answer was dictated by her professional knowledge rather than focussing specifically on her language skills. In this case the participant believed that utilizing her professional knowledge would be an advantage and a good test taking strategy for this assessment question, and likely other questions in the assessment as well. This type of test taking strategy is similar to that described by Rob, another test-taker, however his responses to the question differed as a result of his interpretation of the patient need.

In our focus group, Rob, agreed with Sandy’s line of reasoning for answering this particular exam question explaining:

So what she said is entirely correct because if you don’t have proper knowledge about that topic, you will not be able to fill the form. Because like what she said in the question is like ‘… pillow should be placed under dash’ and the options are head, joints and knee. So we know that rheumatoid arthritis is a joint problem so we’ll of course select the answer “knee.” But if you don’t have that knowledge about rheumatoid arthritis—because we [are] usually placing pillows under the head—you might choose the option like head.

The responses from Sandy and Rob indicate the challenges some participants encountered in the assessment by attempting to demonstrate both their professional knowledge and their English skills. Although Sandy and Rob both employed the same test taking strategy (which relied on their assessment of the patient need) their answer responses differed.

The type of question Sandy and Rob described is a multiple choice Cloze exercise which appears in the reading section of the CELBAN. The Centre for Canadian Language Benchmarks (2014) has issued a Test Taking Strategies document which outlines the
following instructions to complete Cloze exercises: “Read the entire passage once through quickly, ignoring the blanks to get the gist (the general content of the text). Then read again for meaning, line by line to select the best option from the multiple choices provided to fill in the blanks.” This instruction indicates that the best test taking strategy in this case would be to focus primarily on the context of the reading passage. This suggests Sandy and Rob may have benefitted from adjusting their test taking strategy from one which prioritized their sense of patient need, to one which identified the context clues of the paragraph.

**Theme 4 External Factors Impacting Test Experience**

The CELBAN should not be studied in isolation from the wider context in which it operates, specifically a context of immigration and professional licensure. Every test-taker to whom I spoke can be described as a New Canadian looking to work as a nurse. Having completed their training as a nurse in another country, each test-taker enters into the Canadian licensing system on a timeline for which they must complete all necessary examinations to show equivalency of their training to the Canadian model. As I spoke to these test-takers, I noted that frequently IENs discussed aspects external to the test which had impacted their test–taking experience. These factors often directly related back to their experience of the licensure process as a new immigrant to Canada. While these factors do not appear to form the subject matter being tested, or testing criteria of the test, I feel it is important to give voice to the experiences shared with me which described how many test-takers had come to take the CELBAN—an assessment which is designed to address the nursing language proficiency of nurses trained outside of Canada.
Amanda expressed feelings of frustration which focused on the discrepancy between her belief that she would be able to work as a nurse in Canada upon her arrival, and her current situation:

It’s not our fault—we are spending money, we are spending time...We are immigrants here. So we have other things to do too. We have to survive for ourselves too. We are doing the much heavier...like the other kind of jobs which are very frustrating for us too. Because we are not illiterate people here. They called us alright. There’s a skilled worker online—you can apply—you can come to Canada work as a Doctor, work as [an] engineer, work as a nurse—and now they’re saying alright you have to start at zero. Why? If you don’t feel us competent enough then you should stop doing this to us.

While the process of immigrating to a new country can bring about its own challenges as policies and documents change, the process is made more complex for individuals who apply to enter Canada as a skilled worker in a particular area—in this case nursing. Although it may be expected that the individual(s) show documentation of their certification for the purposes of immigrating as a skilled worker, this is not to suggest that the license or credentials of another country while admissible for the purposes of applying for immigration under this category of worker, may be admissible for the purposes of working in that field once the individual has entered the country. As immigration and certification are their own separate processes, they come with their own financial obligations as applicants must cover the administrative fees of scheduling meetings or examinations, and assessing or marking documents. All of this may certainly be made all the more challenging for IENs as they learn the certification process to work in their field will be different depending on where they choose to re-certify within their new country. At the time that I met with CELBAN test-takers in both Hamilton and Toronto, the process for writing the CELBAN was undergoing significant changes which
would impact when and where the assessment might be written as well as the success criteria for the group section vs. individual (speaking) section of the assessment. The experience of continual process change is disorienting for anyone pursing a complex and multifaceted goal. Several test-takers described a level of distrust which can arise when high stakes decisions may be attributed to a reported score which may seem arbitrary in the face of significant life change. Jake provided insight into the experience when he offered the following:

What I was always thinking about is—why should they make the process so tough for Internationally Educated Nurses and even other profession—why should they make it tough?...[M]ost of the people coming to Canada…they may be using English like as the official language—but not the first language. But here again the requirement is like really high for me...So it has been made very cumbersome…And like some have quit and they don’t want to keep on trying because they try—sometimes it just means like 25 you say like no sorry you can’t go ahead.

At several points test-takers like Jake expressed frustration with a process which he described as “so tough for IENs and even other profession[s]”, especially where those individuals have made it known at the outset of their immigration application that their intention is to work in their field within the country. While all assessments must establish baseline success criteria, Jake’s comments highlight the challenge for some IENs to accept that assessment results which just fall short of the success criteria may prevent or delay an IEN from working in their field in their new country.

Immigration brings about considerations on settlement, or where to settle in one’s new country. These choices can be influenced by a variety of factors including proximity to a community where a person, or family believes they may be successful in finding employment, proximity to other family members, and cost of housing. In Hamilton I met
with Jake who had flown to Ontario from Manitoba to complete his CELBAN in a tight
time frame that would meet his safe practice requirements. In Toronto, I met with Janet
who had also flown to Toronto from Winnipeg earlier that week, and Emily who had
arrived in the city from Bellville that morning. I was surprised to hear Emily’s account of
her day as she explained the logistics behind her test date:

“4:30 in the morning? So you left at 4:30 in the morning you were here by 8:15
and your test started here at 9 and then when is your speaking exam?”
(Interviewer)

“At 3” (Emily)

“At 3 O’clock so by 3:30” (Interviewer)

“3:30 or 4 probably if I go” (Emily)

“You’re gonna get back in the car?” (Interviewer)

“Think so, I guess so” (Emily)

The limited number of testing locations available in the province for writing the
CELBAN means that test takers located outside of the GTA or Golden Horseshoe area
are required to travel long distances either on the same day as their assessment, or take on
the additional financial burden of taking more than one day off for the assessment and
paying for accommodations while they are away from their home. For those test-takers
who have settled in communities outside of these densely populated areas, their commute
on the day of their exam, could mean a day of nearly eight hours of driving sandwiching
an assessment experience spanning seven hours, as in the case of Emily. It should also be
noted that from April to November 2016, there are no test dates listed for any testing
centres in the province of Ontario outside of the Touchstone Institute in Toronto
(CELBAN Centre).
The timing and scheduling of assessment components was described both positively and negatively by some test-takers. In the case of Anna, in Toronto, timing was described as a stressor as it pertained to the scheduled time of her CELBAN assessment components. In the case of her oral English assessment, Anna was scheduled and to begin her assessment at 3 in the afternoon. However she explained: “They said my time is 3 o’clock. But they became free before that, and so, I just went for the test around 1:50 I guess, and it was quite in a hurry. Oh I was so nervous, because it is all of a sudden.” Anna’s comments highlight aspects of administration external to the test which have the ability to impact a part of test experience. Variability in test times is expected by test-takers, as evidenced by the fact that Anna was in the waiting room for the oral English assessment more than an hour prior to her start time. Her comments “it was quite in a hurry” coupled with the observation that she was asked to begin her oral English assessment just over an hour before her scheduled start time, indicate that prior test-takers were completing their oral assessments faster than expected, or perhaps that some were unable to make their test day and time. Regardless of the reasons, Anna’s Oral English assessment was rescheduled an hour before her anticipated start time, that she felt a sense of hurry which increased feelings of nervousness entering the assessment may be described as a factor which impacted her test taking experience on the day of her CELBAN assessment.

Timing and scheduling of assessments was not described by all test-takers as a negative external factor that impacted their test taking experience. Nancy, like Anna, completed her assessment at the Toronto location and similarly noted that her assessment did not start exactly when it was scheduled. Her oral English assessment was scheduled
for 1:30 in the afternoon, but explained that her assessment was delayed by about 10 minutes. Nancy described her feelings on the oral English assessment this way: “I thought the speaking was really well organized, and I felt really…relaxed.” She didn’t mind the shift in timing for the assessment and said “But I didn’t have any problem with the time because they did really need to manage and organize everything that’s why.” Furthermore, she noted that the administrative side of the oral English assessment provided a very positive experience “the interview was so like, pleasant. All the assessors were nice.”

During the Toronto focus group, several participants discussed the stresses that impacted and accompanied them on their test day explaining:

“And they said you have to be psychologically confident—you know—but psychologically it’s giving us stress. It’s very stressful” (Amanda)
“Getting through this difficult situation and squeezing them and getting all of their” (Syd)
“Money even time” (Amanda)
“And at last just giving them back to the full service no” (Syd)
“You know how many things I’m lose right now while I’m here? My job, my study, my exams, and my—you know I’m living alone. So that’s very stressful. No family here and you are struggling like it’s too much” (Jessie)

Their comments highlight the multiple and complex challenges that impact the CELBAN test-takers including the challenge to remain psychologically strong while completing a series of examinations in a new country, within an allotted timeframe. These stressors may be exaggerated for individuals who must take on lower paying work while endeavoring to show competency or equivalency in their professional training, and cover the costs of the associated assessments and examinations which will afford them the opportunity to work in their field and earn a wage commensurate with their level of skill
and educational investment. For some test-takers their goal to begin working in their field in Canada is just part of a wider plan to provide for and eventually support their other family members in joining them in this new country.

**Theme 5 Preparing for the CELBAN**

Overwhelmingly, test-takers described their appreciation for an examination like the CELBAN which recognizes the specific language skills and knowledge IENs bring to the field of nursing. Just as frequently they spoke of their need for additional educational resources to outline the exam components and provide formative assessment opportunities to evaluate gaps in their current language skills to offer a focused and measured approach to their assessment preparation. The most frequently cited concern in regard to CELBAN preparation was the challenge in locating CELBAN specific study materials:

> I went into the internet and the libraries to look for resources and there were none-so that really scared me. But then…my friend told me that because we don’t have any idea, it’s better that we take a prep course. Which I did take, paying a lot of money, but I think it was no- It was no help. Because I think no one really knows what CELBAN is. Even the people claiming to be teaching CELBAN have no idea what the CELBAN is at all. Because what I studied in the class and what I did today was totally, entirely different (Mary).

Whether ultimately to the test-taker’s advantage or disadvantage, to date there are no CELBAN books published by The CELBAN Centre that are akin to the types of preparation books available for the IELTS or TOEFL test. While IELTS and TOEFL cater to a global population with a wide variety of professional and educational goals, the CELBAN’s audience is far more limited: IENs in Canada. There are increasing numbers of test preparation or FAQs sheets which are being created and posted by The CELBAN
Centre, and other sites which claim to have expertise in this area as well, for example: www.celban.biz, or http://dear-kim.com. Mary indicated that as a result of the challenges she experienced in locating CELBAN resources, she chose to pay for a preparation course not endorsed by The CELBAN Centre. Mary explained that the course she took was not reflective of the exam structure or questions.

In 2015, an update was made to the website hosting official CELBAN resources, and as of March 2016, seven CELBAN tip sheets have been posted to the CELBAN Centre website (CELBAN Centre), and a CELBAN Readiness Self-Assessment is also posted. There are also documents describing “Test-Taking Strategies,” and “What to Expect on Test Day” (The CELBAN Centre). These documents provide additional support and answers to frequently asked questions by IENs who are interested in taking the assessment. Although test-takers often want to support one another by sharing their own personal experiences of the CELBAN, official documents are important for setting down in clear language the structure, and expectations of the assessment. Comments like those made by Janet indicate that although time and effort was spent in registering for the assessment, she was not fully aware at the time of her assessment of how her responses would be scored or how high her scores would have to be to successfully complete the assessment. For example, during a focus group in Toronto she stated:

And for the score of CELBAN it’s like you have to get 10 out of 10 for the listening. I don’t know how- like how many mistakes you can afford to get in that one. And for the rest it’s like 8- so I don’t know how many mistakes (Janet)

Although Janet’s demographic survey revealed that she had previously had her CLB levels assessed, at that time, only one of her language area scores would have been high enough to meet the minimum requirement for the CELBAN. In the focus group that Janet
attended she also revealed that she was unaware of the policy regarding how soon after
the completion of a CELBAN attempt a test-taker may attempt the assessment again. In
fact, of all of the attendees in the first Toronto focus group only one, Rose, was familiar
with the policy. I asked the participants what the next steps would be for somebody who
completed the CELBAN and was not happy with their score. Rose explained “They have
to wait for three months to get another test date and they have to pay again.” All other
participants in the focus group asked why “three months?” Rose further explained “Yes,
three months it says so in…celban.org that if you fail—you can also apply for re-
marking.” These pieces of information may suggest that in the case of Janet and some of
the other participants in the first Toronto focus group, additional preparatory materials
may have been useful in determining if they were ready to complete the CELBAN, and
how they might plan out their next steps and certification timelines if unsuccessful on
their assessment attempt.

Theme 6 Test Conditions and Administration

Comments from test-takers at both the Hamilton and Toronto assessment centres
highlight the challenges that can occur when administering long, complex paper and
pencil assessments, especially in regard to the provision of test taking supplies. For
example, as a paper and pencil test, the CELBAN features a question booklet and an
answer booklet. Test-takers frequently described frustration at not being allowed to take
notes on their question booklet to guide their responses in their answer booklets:

We had…answer booklet and question booklet…the thing that I didn’t like about
this listening exam was they…didn’t allow us to write anything on the question
paper. Because there is a video going on, two people are having a conversation—
and you are doing a multitasking skill in this listening…we’re looking at the
screen…you’re listening what they’re talking, and you’re also reading the
questions. And the question had all the options and the answer booklet had where we need to mark the answers right. So, they didn’t let us and mark anything and the way we are trained to do IEL-listening and IELTS is we have, we use the key words- you know because the conversation doesn’t stop (Mary)

Mary highlighted an important point in regard to the structure of the exam. Several tasks, such as the listening task described above are integrated tasks—requiring the test-takers to listen, observe, read questions and indicate a response. These tasks require the test-takers to perform several tasks at once, or multitask, as Mary described it. In addition to these operations, test-takers must also ensure that at all times they are reading the written questions from one booklet and copying their answers and any pertinent details from the question booklet into their answer booklet.

I asked whether or not test-takers were allowed to record some of their thoughts or practice part of their exam response on a separate piece of paper, scrap paper for their own reference. Mary explained that for her test day this was not an option, and furthermore she was concerned that any notes she took on her answer sheet might invalidate her responses, although she did feel the need to make some notes:

There’s no scratch paper nothing…that’s one thing which I found to be very difficult you’re not able to write anything I don’t know if it’s good because your answer[s are on] scantr[on]. I guess if you write anything apart from marking [your answer selections], the computer may not take it—that’s what my concern was…They said you could write on your answer sheet. But I think these are all Scantr[on] which…goes into the computer…I’m not sure how it’s going to- because I did write down some things on my answer paper. So that’s one concern (Mary)

Particularly for a paper and pencil test, it might be assumed that test-takers would have the access to scrap paper for note taking. Certainly, the absence of any note taking space outside of a scantron sheet opens up the possibility—as in Mary’s case—of test-takers
including notes on their scantron sheet which could affect the automated grading of test responses leading to score variance not reflective of the test-takers skill level.

As with many other types of high stakes assessment, test-takers are given very strict instructions for what they can or cannot bring into the exam room with them. In the case of the CELBAN, some test-takers spoke about the challenges they experienced with writing down their responses, not only due to the limited space for note taking but also due to the restriction on the number of pencils permitted to each test-taker. Penelope explained

But they didn’t give the pencils—they only give one pencil…and whenever we need[ed a] pencil…we have to raise the hand so that takes time and that distract[s] other people as well right. So at least they [should] have to give two pencils and one sharpener (Penelope).

Penelope’s comments seem to offer practical advice for a simple improvement that could be made to the test taking experience. Offering even one more pre-sharpened pencil to test-takers would surely help to reduce the downtime experienced by a test-taker whose lead has broken in their pencil as well as the possible disruption to other test-takers as they await the receipt of another pencil with which to complete the test. With even one more pencil, test-takers may not require a sharpener which could disrupt others as a result of the noise or distraction of its use during the test. Offering two pencils to test-takers seems like a practical and reasonable solution that would also help to ensure that a skipped question is more representative of a test-taker’s decision to leave a question blank as opposed to the possibility that they did not have access to a pencil at the time in which they may have otherwise recorded an answer.

On the spoken assessment, test-takers noted other types of challenges related to the administration of the assessment. Some test-takers spoke about their concern that there
was only one assessor in the room while they completed their speaking assessment. In particular, they speculated that it might be difficult for one assessor to complete all of the tasks required on their own, such as: taking notes, grading, listening and responding. Some expressed a concern that the assessor may miss a critical piece of information shared by the test-taker, such as Penelope who explained: “It’s very difficult to speak if somebody's not responding [to] you…. [The] examiner is very busy in writing the comment—so he will not give you [an] answer to the question.” I asked Penelope if she was ever asked to repeat a question because the assessor was taking notes and needed to hear the question again before responding, she explained:

Ya, sometimes they [the assessor] ask—because they’re busy writing. So what I [was] thinking they should do… is [have] two examiners in the room. So one examiner should take the comment and another should be the patient and maintain the eye contact (Penelope).

While some test-takers noted their appreciation for the level of authenticity afforded by the role play part of the assessment, certainly as Mary noted the inclusion of a second assessor for the purposes of continuity and cohesion in the oral assessment would be beneficial both to the assessors grading the quality of the spoken responses as well as the test-takers offering them.

While there were common themes which emerged from my discussions with test-takers on the administrative challenges which impacted their test experience, there were many test-takers who also indicated numerous ways in which invigilators, and the overall administration of the exam was seen as advantageous or beneficial to the overall test experience. For instance some test-takers noted the very clear and consistent instructions given for the oral assessment. Sandy noted that:
When I entered into the room they gave me time to get relaxed, they introduced themselves…they explained [to] me how the speaking test will go, they asked me question[s], then they explained [to] me how the role play will be taking place. So, things were pretty organized…

Test-takers like Sandy were appreciative of opportunities to have a positive interaction with test invigilators noting their smiling faces (Anna), and their attention to possible stressors like the level of volume for the listening section (Penelope). While these test-takers indicated their concerns about how they would perform on their test day, they were happy to note instances in which they felt test administrators or invigilators offered support or acknowledgement of their feelings.

**Theme 7 Authenticity**

When meeting with participants, I was interested to learn more whether or not they felt the assessment met its goal to accurately reflect the types of communicative tasks required of a nurse. In this way, I felt I could begin to understand how test-takers perceive the authenticity of the CELBAN. To learn about test-takers’ thoughts on this topic, I asked questions about whether or not they believed the assessment tasks, or questions, provided an accurate reflection of the types of communicative tasks required of a nurse. What follows are, what I believe to be, the most significant responses to my questions on the topic of authenticity.

On the topic of the listening section of the assessment, Joe explained: “I think they’re recording from the life situations.” Some participants observed that listening activities included recordings which also captured background noises commonly heard in a nursing setting. This was described as similar to what often occurs in a clinical setting when conversing with patients. For example, Mary explained “…you need to be good
with your communication, [with] your listening because this is exactly what happens in a nursing setup. You’re listening, there’s a lot of noise going on…you cannot assume things.” The inclusion of background sounds in the listening assessment was noted by some participants as a point of contrast to other types of English language tests they had completed in the past, for example the IELTS. As Joe explained “…the clarity of the recordings [of the CELBAN] compared to IELTS…For IELTS…maybe the questions are difficult but the clarity of the questions or the speaking qualities—come more than compared to CELBAN.” On this point I felt it was important to probe for additional information on the type of clarity that was noted by this participant. At this point in the focus group I (“I:”) asked the following:

  I- Ummhmm so can you tell more about that, “the clarity”—do you mean the kinds of voices you heard?
  Joe- No it’s, like there are lots of other noises are there in the listening
  I- In the recording
  Joe- ya, in the recordings. I think they’re recording from the life situations.
  I-Ya
  Joe- Not like for the other exams
  I- Why do you think they did that?
  Joe- We can hear…some background noises that happens around that situation, in the general atmosphere. Or same like if you are recording something out from here, there are some noises of [a] teacher, or the movements of the agendas, same [as] we can hear in the CELBAN exam too.

The observation made by Joe that background noises seem more likely to be included in the CELBAN listening section as opposed to other general assessments of English language skill, in conjunction with Mary’s remarks that in a nursing setting “there’s a lot of noise going on” suggests that the listening section of the CELBAN exam more
authentically captures the demands placed on the listening skills of a nurse while working on the job. Although background noise may be considered a source of score variance in a general test of English listening skills, it is interesting to note that within a nursing specific context it is described by some participants as an assumed reality of the challenges of job performance.

Participants also described the way in which listening activities captured not only the sounds commonly heard in a clinical care setting, but the process or approach to conversations with patients that may be employed. For example, one participant noted that if you are meeting a patient for the first time and need to discover more information about their medical history, it is important to ask open ended questions in order to encourage the patient to willingly disclose information related to their lifestyle choices. This participant noted this type of behavior to be demonstrated in the listening activity which is presented in the form of a video, “she’s [the nurse] indirectly asking the questions to the patient to get information for what she has in the paper” (Rob).

Although I was unable to follow up on this point, a question arose in my mind regarding the complexity of assessing listening skills in the contextualized format of the CELBAN. As to Mary’s point regarding the importance of not making an assumption about what a person has heard and the importance of eliciting information from a patient in an open ended manner, I wonder about the challenges that may be faced by a test-taker who is asked a question about what they have heard in conversation. For example, if a patient tells the nurse that they are a social smoker, or casual smoker; is the nurse expected to categorize this responses in the form a ‘yes’ or ‘no’ or are they expected to provide additional background? Although the test-taker may hear the word smoker, as a
nurse do they feel the need to clarify what they have heard for the sake of accuracy in their reporting, to ensure they have not communicated an assumption by way of generality? According to Grabowski & Dakin (2014) authenticity of task is defined by the test tasks being a close approximation of the Target Language Use (TLU) tasks. In this regard then, a demonstration of nursing listening skills may not involve the recording of words used by a patient, but more specifically the recording of the precise information relayed or withheld during a nursing assessment.

When asked about how realistic the test was as it compared to working as a nurse, many test-takers responded with comments on the oral component of language assessment. This part of the exam asks test-takers to complete a role play in which they gather patient history and offer discharge instructions to a patient. Not surprisingly then, many participants connected positively to this component of the exam in regard to how closely it felt like the work they do most often. Although many test-takers recognized this exam component’s attempt to incorporate on the job types of tasks into the exam, many also noted its limitations. For example Jake acknowledged appreciation for the acted responses of the assessors but also indicated the tension that was created between the test environment and the approximation of the work environment:

When I was taking- from the guideline…the [patient] history taking, it was a good guideline, however it was a little bit hard to end it. You know like in a real situation there’s a procedure on how you can end a conversation of the nurse patient interaction-but in that form [the role-play] I think it was nothing like that. So, I had to …figure out, what, how can I conclude here? So ya, I think that’s one part that was a little bit challenging…how can I conclude here? I was not sure how to- where do I send the patient? What’s the next level here? (Jake).

In this case, Jake indicates an appreciation for the structured nature of the role play and a willingness to act along with it, but this is challenged when the structure of the role play
misses a procedural step which would follow from the conversation. In this case there was no location to send the patient to. It was Jake, in fact, who would shortly be required to leave the room but in the typical work scenario, Jake is tasked with sending a patient either to another area of the hospital, a pharmacy or an alternate location.

Some participants noted moments in the test that stretched their perception of authentic representations of the practice of nursing. For example, one participant explained that in part of her speaking assessment she was asked to demonstrate how she might convince a patient to change their behavior for the sake of their health. Her response highlights how authentically representing a nursing conversation is challenged by the various epistemologies of practice held by test-takers from all over the world:

So it's not generally done in our place [country]. So it's a bit hard for us to deal in that manner…That we are dealing with someone to convince them. It’s generally the duty of the doctors…So thinking that way I think it was a bit hard for me to have to like convince someone to do exercise when the person is not actually willing to do that. So, I had not undergone such a circumstance…so it a bit hard for me to relate to it (Stephanie).

This response identifies an awareness of the differences in the way in which nursing is practiced in Canada versus in other countries. Although a discussion in which a nurse would attempt to convince a patient of anything would be seen as inauthentic and unprofessional in Stephanie’s country of origin, her description reveals a sense of belief in the authenticity of this practice as authentic in a Canadian context. This observation is consistent with research in the area of Internationally Educated Pharmacists (IEP) conducted by Austin & Dean (2006) in which they defined four broad areas unique to the learning needs of internationally trained pharmacists. The fourth area noted by Austin & Dean (2006) was the challenge of redefining the way in which a profession is practiced in a different country, they explain: “reconceptualizing one’s professional paradigm of
practice was a difficult shift for many foreign-trained pharmacists, particularly when coupled with Western democratic notions regarding patient autonomy, decision-making and confidentiality” (26). Austin & Dean (2006) referred to a simulated patient case where IEPs administered birth control to a fourteen year old girl without a parent present, and were later asked about whether or not the girl had been given the medication by a parent. While a pharmacist may encounter this scenario in many different countries, how they respond to it depends on what is expected legally and professionally by that pharmacist in the country in which they are practicing. Similarly, while Stephanie may have come in contact with several patients in the past who required counselling on healthy lifestyle choices, in the past it has been expected that she would refer this type of patient to a doctor, not offer the counselling directly.

**Summary**

The findings from the focus groups and interviews contributed to the analysis of test-taker’s perceptions of the CELBAN, its constructs, possible sources of construct irrelevant variance, and authenticity of language and tasks. The thematic analysis presented here provides a contribution to CELBAN validation otherwise absent from the current literature. A more detailed discussion of these results now follows in, contextualized by the literature reviewed on high-stakes, large-scale AELLs and test validity.
Chapter 5
Discussion

In this chapter, I will describe my interpretation of the data based on relevant literature. By connecting my findings to the literature in the field of assessment and evaluation, I hope to provide new insight on the CELBAN and its role as a high-stakes large-scale professional English language assessment. This chapter is organized by the four research questions which have guided this study with reference to the data collected and analysed here, and its connection to the literature of the field. The last section will offer recommendations for future CELBAN administrations, as well as, a discussion of the limitations of this study and directions for future research.

Sochan & Singh (2007) comment on the limited research to date on the nature of Internationally Educated Nurses (IENs) experiences in Canada (p. 131). They argue that a significant gap exists in the current understanding of “IEN acculturation, and socialization as they gain entry into practice outside of their homeland” (Sochan & Singh p. 131). Although the province of Ontario has taken steps to address concerns regarding the economic integration of IENs into the province through bridging programs (Ministry of Citizenship and Immigration, 2013) and created policy recommendations to continually improve the quality of these types of programs (Austin, 2008, p. 3), there remains limited research on the perceptions of IENs in completing certification exams either at the provincial or at the national level. This study thus contributed to this emergent research area of professional language testing validation through the analysis and inclusion of test-taker feedback to construct representation, and the interpretation of CELBAN test scores as evidence of IENs’ effective communication in, and
comprehension of, English in a professional nursing context. This study also addressed
the importance of including IENs and documenting their experiences as key stakeholders
in the certification exams and programs as highlighted by Austin (2006), Cheng &
DeLuca (2011), and Couser (2012).

In the following discussion, I will answer my research questions by synthesizing
the results. I will also identify the key findings and provide future recommendations for
professional language testing generally and the CELBAN specifically, as it pertains to the
evaluative judgement “of the degree to which empirical evidence…supports
the…appropriateness of interpretations and actions based on [CELBAN] test scores…”
(Messick, 1991, p. 1), and the validity of these test scores as a marker of professional
language competency.

**How do IENs characterize their test experience?**

IENs who participated in this study consistently described their appreciation for the
CELBAN’s professional language context which they felt provided them with an
opportunity to more accurately demonstrate their language skills as compared with other
academic language assessments. This type of feedback from test-takers suggests that
generally, a professional language assessment offers a closer approximation of the Target
Language Use (TLU) tasks, and thus greater authenticity of task (Grabowski & Dakin
2014), consistent with the intention of the CELBAN’s development in regard to face
validity (Epp & Lewis, p. 296). These findings concur with the findings in Cheng &
DeLuca (2011) that construct underrepresentation may occur when a test designed for
English academic purposes, or university entrance is used for professional certification
purposes (p. 106). This finding also confirms prior research in the area of validation of
large-scale assessment of AELLs. Where positive construct representation has been noted it has, in the past, been tied to the perception of authenticity of language tasks, as was noted in DeLuca et al. (2013) in the description of “positive construct representation” (p. 668). Although this was not a sensitizing theme of my study, it does fall more broadly within the sensitizing concept “test structure and content” (see Table 4).

Although participants indicated a preference for a professional language specific assessment, many indicated challenges in navigating the complex licensure requirements, including identifying the CELBAN as an option available to them, and identifying the assessment early enough in their certification process to allow them to prepare, register and complete the assessment within their safe practice. Where participants were able to identify and complete the CELBAN as a means to demonstrate their language competency within the span of time allotted for safe practice, they noted a general lack of preparatory resources available, especially those endorsed by a professional or regulating authority. Cheng et al. (2013) have noted that “non-recognition of international credentials and challenges in satisfying the licensure and certification testing requirements have been identified as the largest barriers to successful integration into the workforce” (p. 734). The experiences shared in this study highlight then a deeper concern regarding the costs to successful integration where Internationally Educated Professionals seek out opportunities to show equivalency or competency but are unable to determine the options available to them to achieve licensure, or find out too late to make those options viable for their application. Recent changes to the CNO requirements for IENs seeking membership has meant that applicants must now show language competency by completing either the CELBAN or IELTS exams (CNO, 2013). While the CNO has made
efforts to develop their website to make it easier for IENs to understand the registration process, what was once an eight step process demonstrated by Walton-Roberts & Hennebry (2012) is now a ten step overview which includes two steps for examinations; however the details of this section do not include a reference to the CELBAN or IELTS, neither is there a step specific to language competency or language exam (CNO, 2014). CELBAN preparatory materials have increased in number since 2015, and as of March 2016, seven official CELBAN tip sheets have been posted to the CELBAN Centre website (CELBAN Centre). Additionally, a CELBAN Readiness Self-Assessment and documents describing “Test-Taking Strategies,” and “What to Expect on Test Day” are also available at no cost to the IEN (The CELBAN Centre).

While individual comments on the CELBAN were as diverse as the participants to whom I spoke, central themes discussed in the results chapter allowed me to understand the aspects of the CELBAN described positively and most significantly, and those described negatively. Participants consistently highlighted their appreciation for CELBAN as an assessment, which focused the testing constructs on authentic areas of professional practice. Participants who contrasted the CELBAN testing experience with other tests they had experienced in the past, inevitably indicated how the professional nursing context of the CELBAN provided a sense of confidence in their own skills, as well as a confidence in the assessment to identify their skill level. In contrast, test-takers vocalized concerns with other language testing experiences which focused on general academic language skills such as essay writing skills which is consistent with the research finding that construct underrepresentation may occur when a test designed for English academic purposes, or university entrance is used for professional certification
purposes (Cheng & DeLuca, 2011, p. 106). This positive feedback on the CELBAN in contrast of other general English tests from test-takers is extremely important in terms of validating the original intention and development of the CELBAN which aimed to include IEN perceptions of the face validity of the assessment (Epp & Lewis, p. 296). In this way the validity evidence may be said to have been enhanced by these consistent findings derived from various studies including the current study (Messick, 1991, p. 2).

Although many test-takers described the CELBAN in positive terms for its focus on nursing specific language as opposed to generalized academic English, there were significant concerns raised in regard to how the CELBAN fits within a wider certification process, and its appropriateness as a measure of professional competency when this test is not identified as a prerequisite to the NCLEX-RN or CPRNE (See Penelope Theme 1).

Although certification to work as a nurse in Ontario for IENs is a process governed by a series of assessments, as of April 2016, there is no policy which requires an applicant to show evidence of English language proficiency prior to completing the NCLEX-RN or CPRNE. All applicants must apply first to the National Nursing Assessment Service to be eligible to engage with the certification process to work as a nurse in Ontario (CNO, 2015 Q&As About the NCLEX-RN for exam writers). This means, theoretically, that an IEN may complete and pass the certification exam which assesses professional competency to practice but fail the English language assessment. As in the case of Penelope, this certification experience may lead to a decrease in the perception of face validity of the wider certification process. For some, like Penelope, language skills may allow them to read and successfully respond to the professional competency exam, but at the same stage
of language development, these IENs may be unable to demonstrate a level of mastery with English language to permit them to pass the language assessment.

**How do IENs describe the constructs measured by the CELBAN?**

Perhaps as a result of the challenges related to identifying individual pathways to licensure and preparatory resources, test-takers who shared their CELBAN experiences with me often described the test constructs using the words of the test questions, as opposed to the competencies or language skills being tested by the assessment. Malone and Montee (2014) explain that, “stakeholder beliefs about the correspondence between test tasks and the skills they purport to test provide important support for test validity” (p. 2). Participants in this study did not reference the Canadian Language Benchmarks website as a test preparation resource, or reference the language of the website which defines the language levels as they map onto language competencies assessed by the CELBAN (The CELBAN Centre B, 2015). The passing or acceptable scores for each CELBAN section are not addressed in the section of the CELBAN Centre website dealing with test score (The CELBAN Centre C, 2015)—but are described on the CNO website (CNO, 2013). Perhaps not surprisingly then, participants frequently noted a desire for more preparatory information that would outline what the CELBAN is, how to prepare for it, and an overview of structure and types of questions they would be asked to respond to. The generality of the descriptions offered by test-takers on the language competencies tested by the CELBAN may suggest a lack of clarity or understanding of the specific language skills which must be demonstrated, as well as those specific skills which may be a strength or weakness of a given applicant.
While participants in this study described test constructs generally, and without reference to the documents that outline the language competencies that must be demonstrated to show evidence of language proficiency, what warrants continued attention are the instances in which test-takers described testing constructs not tested by the assessment, namely professional nursing skills. Participants described both specific areas of the assessment where they felt they were being given an opportunity to demonstrate some of their nursing knowledge, and instances where they used their nursing knowledge as a strategy to select the correct answer to a test question (Sandy & Rob). These reflections on the test taking experience are troubling as the CELBAN is: “an occupation-specific language test that assesses these core competencies” (The CELBAN Centre A, 2015), but not occupation-specific competencies related to medical practice.

Language assessment is a highly technical field, with a host of specialized terms developed to describe its particularities and nuances. Many of the test-takers that I spoke with described the constructs tested by the assessment using the words of the assessment, as opposed to the competencies and objectives assessed by each section of the assessment. This may be due, in part to the challenges associated with determining the specifics of what each section of the assessment is assessing test-takers.

The CELBAN Centre website explains that “the CELBAN is an occupation-specific language test that assesses these core competencies” (The CELBAN Centre A, 2015). However “these core competencies” (The CELBAN Centre A, 2015) are not described or defined elsewhere on this page. The Test Taking Strategies workbook posted to the site does provide more specific information on what to expect in the test and how
to approach each test section, but not all sections describe the key constructs being assessed. For example, the CELBAN Listening section explains that test-takers will answer multiple choice questions in two different formats, but does not describe in detail the types of listening skills being assessed (The CELBAN Centre A, 2015 p. 4).

However, the Reading section of the Test Taking Strategies booklet offers a table focussed on skill areas test-takers can self-assess prior to taking the assessment e.g., “Do you use point form effectively? Do you omit unnecessary words and include necessary words to convey information clearly?” (p. 5). The CELBAN Centre explains that “CELBAN scores align with the Canadian Language Benchmarks” (The CELBAN Centre B, 2015). A link is provided on the CELBAN Centre Website to the Centre for Canadian Language Benchmarks website where a person can download the 235 page CLB document which outlines the competencies associated with each language benchmark, however these are not described in detail on the CELBAN Centre Website.

The passing or acceptable scores for each CELBAN section are not addressed in the section of the CELBAN Centre website dealing with test score (The CELBAN Centre C, 2015) - but are described on the CNO website (CNO, 2013).

When participants in this study described the constructs being tested they often spoke more generally about the types, or formats of questions they were responding to. Jake, for example, spoke about the skimming and scanning section as an area which was assessing his reading comprehension skills, and specifically required him to “look for specific information… [and] extract information from the passage.” This is a good summary of the types of objectives laid out for achievement of a CLB level 8 as referenced by The Canadian Language Benchmarks: English as a Second Language for
Adults document, but lacks detail on the specific skills that must be demonstrated to meet the requirements for this section. The CLB document specifies that a level 8, the minimum acceptable score for reading (CNO, 2013), requires an individual to: “identify purpose, main ideas, specific details and implied meanings, as well as mood, attitude and register” (p. 86). While Jake’s explanation offered a summary of the expectations of the task, neither he, nor any other test-taker with whom I spoke mentioned identification of implied meanings, or register as are identified in the CLB document. Although Jake and others’ contributions to my questions about which skills were being assessed by the CELBAN reveal that most participants believed each section to be testing their language skills in each respective area of the English language, there was a lack of specificity in their responses which may reveal a lack of clarity generally on what to specifically prepare or work to improve upon, prior to writing the CELBAN, or after receiving one’s results.

Some participants described instances in which they felt the CELBAN assessed their professional competency skills uniquely from their language skills. The CELBAN Centre website refers many times to the language focus of the assessment (The CELBAN CENTRE A, 2015). However, as reported in the results chapter, Sandy described the benefits of the test structure focussing on generic nursing practice as opposed to specific areas of nursing, as a good way to assess professional knowledge (see Skills Tested by the CELBAN). Rob also explained that he felt the language component may be weighted more heavily than professional knowledge by saying: “I think English, like the language assessment is more than the knowledge about nursing.” He then discussed his concerns with that approach. Specifically, he was concerned that some of his professional
knowledge might be missed in his writing section if the assessor did not have the medical knowledge to appreciate the attention to formatting required of these types of reports. Comments like those made by Sandy and Rob reveal a perception that some components of the CELBAN may be assessing some professional nursing knowledge tied to nursing practice and performance of job duties, which is not the case.

In other instances, participants described their reliance on their nursing skills to answer CELBAN questions and cited this as a benefit of the CELBAN. This was described by Sandy and Rob who selected answers to the cloze reading questions by relying on their knowledge of the medical condition described in the reading passage, as opposed to their knowledge of sentence structure, prepositions and nouns. These types of responses reveal that some CELBAN test-takers are not aware of all of the testing constructs, and in some cases believe some aspects of their practical, nursing knowledge, or job performance skills are being assessed by this assessment which is intended to test language skills specifically.

These findings concur with observations made in Cheng & DeLuca (2011), regarding “test purpose” (p. 114), a sensitizing concept in this study. Cheng & DeLuca noted instances where test-takers recognized “incongruence between the test purpose and test practice” (p. 114). In the case of this 2011 study, it was noted that “the aim of the test and the scoring protocol” of a test separate from the CELBAN were “misaligned and thus contributes to the invalidity of score meaning” (p. 114). Not all participants that I spoke to indicated that they perceived constructs not assessed by the CELBAN or measured by the CLB to be components of their test-taking experience. However, that participants in this study did not described the key measures of success as defined by the CLB when
asked about how the CELBAN assessed their skills, indicated the possibility of score variance where test-takers may not have been aware of all of the constructs being measured by the assessment, and therefore misalignment between the aim of the test and the scoring protocol.

**What, if any, potential sources of Construct Irrelevant Variance do the test-takers describe based on their experience?**

Messick has suggested that the key issues of test validity are: “meaning, relevance, utility of scores, the import or value implications of scores as a basis for action, and the functional worth of scores in terms of the social consequences of their use” (Messick, 1991, p. 4). Similarly, on the topic of validity Denzin & Lincoln (2008) have said “We know a thing only through its representations” (p. 5). The CELBAN provides a representation of language skill which is intended to guide our understanding of the occupational specific language both receptive and productive available to an IEN at the time of their assessment. Because the outcome of the assessment is intended to be utilized for the purpose of identifying those who have met the language competency requirements to proceed with their licensure, and gating those who have not, it is incredibly important to ensure that the meaning of the score refers to language skill and not to other factors which may impact our interpretation of skill level e.g., possible sources of construct irrelevant variance.

Prior research in the area of large-scale AELL testing have observed sources of construct irrelevant variance at play such as: “testing environment, test design, and score reports” (DeLuca et al., 2011, p. 670). These were used as sensitizing concepts in this study, however as score reports were not addressed within this study they did not form a
code or impact the thematic analysis of the results described here. Testing environment and test design are construct irrelevant score variance factors described in DeLuca et al. (2011) which were also noted in this study. Similarly, Cheng & DeLuca (2013) addressed similar issues concerning “test administration and testing conditions” (p. 110).

Participants to whom I spoke, described challenges in completing the CELBAN related to the availability of test supplies such as pencils and scrap paper for the written sections of the assessments, as well as the demands placed on a single assessor when listening, responding and recording feedback on the oral English section of the CELBAN. Some of the concerns raised by participants in this study related to their personal experience completing the CELBAN indicate opportunities to improve the test taking experience for future IENs as it pertains to Messick’s (1991) definition of test validation: “essentially a matter of making the most reasonable case, on the basis of the balance of evidence available both to justify current use of the test and to guide current research needed to advance understanding of what the test scores mean and how they function in the applied context” (Messick, 1991, p. 2).

The CNO reviews completed CELBAN score reports to “determine if [an IEN can] demonstrate the ability to read, write, listen, and speak at a level that permits [them] to practice nursing safely in Ontario” (CNO, 2015). Because CELBAN test scores guide the decision to progress or hold back an IEN in their certification process, it is very important to ensure that the reported score accurately communicates information about their language skill level, and no other aspects which may impact the reported CELBAN score ex. exhaustion, misinformation about the test, access to test supplies, and reliable assessment protocols on the day of the test.
While exhaustion may seem like a typical reaction to any high stakes assessment, in some cases I spoke with individuals who may have been more significantly disadvantaged in their test writing experience as a result of the challenges they experienced in trying to get to the test centre for their test date. In part this may have resulted from the very specific time in which I collected my data. In the months preceding data collection the CELBAN administrative centre moved from Manitoba to Ontario. As a result, the test was unavailable for the summer months in 2014, leaving some IENs unsure of when their next CELBAN test-taking opportunity may be. These changes impacted some IENs in terms of their certification timelines which dictate the allowable time period for IENs to demonstrate all required competencies within a time period deemed safe for them to resume practice within the nursing profession. Some participants with whom I spoke indicated that as a result of their need to demonstrate language competency before expiration of this wider licensure and certification timeline they had travelled a significant distance to complete their language assessment on their given testing day and time. Emily, for example, left her home in Bellville at 4:30 in the morning in an effort to make her assessment which began at 9 in the morning and was completed by around 4 in the afternoon. I also met Jake who flew into Toronto from Manitoba on the day before his test and took inter-city transit to travel from Toronto to Hamilton for his test. This represents an facet of the test-taking experience is categorized as an external factor which may have impacted test-takers’ abilities to demonstrate their full range of their skills as compared to test-takers who were not required to travel for several hours, or for days, to travel from another city or province to complete their assessment. Although, the majority of test-takers that I spoke with travelled to the test
centre from a location less than three hours away, the significance of the sense of urgency which compelled some test-takers to travel great distances on the day preceding, or day of, their assessment is an aspect of the assessment experience that warrants further attention.

While travel related exhaustion may have been an aspect of the test taking experience which affected only a subset of test-takers, the majority of participants in this study referenced the challenges they faced in their test preparation, specifically in regard to accessing reliable information about how and what to prepare for the CELBAN. Rob explained that he had located courses online which claimed to offer CELBAN preparation, but that these were offered at a cost in excess of one thousand dollars. Mary took a CELBAN preparation course and found that it was not a reliable source of information on the test, its structure, or the types of skills she would be expected to demonstrate. She explained “what I studied in the class and what I did today was totally, entirely different.” The challenge in locating resources may help to explain the generality with which many participants described the constructs, language skills, being tested by each section of the CELBAN. Noted in Theme 5, Preparing for the CELBAN, as of March 2016 seven CELBAN tip sheets have been posted to the CELBAN Centre website (CELBAN Centre) however as of April 2016 there are no CELBAN preparation books published by the CELBAN Centre which offer detailed information on the structure of the exam or opportunities for practicing question types. What is especially surprising about this is that there is a self-readiness exam posted on the CELBAN Centre website (The CELBAN Centre D, 2015), but the CELBAN itself is a paper and pencil test. As of April 2016, the CELBAN Centre website described the self-assessment in the following
way: “The self-assessment also gives you an overview of the test format and content, so you will have a better idea of what to expect during the test” (The CELBAN Centre D, 2015).

This description contrasts with Amanda’s experience, as she explained “I was expecting that the exam was on [the] computer…like [I] did on the readiness exam.” Test-taker responses like this indicate a willingness to self-assess skill level and work toward self-improvement prior to attempting the CELBAN. What this test-taker feedback indicates is that it would be additionally useful for the area of the website which hosts the self-readiness assessment to indicate that the CELBAN is not an online test, and thus varies in both the question and response format from the assessment offered to an IEN on their given CELBAN date.

The CELBAN by its very nature is intended to support a very specific demographic of test-takers: IENs, and specifically IENs who hope to work in Canada. As such it is unfair to compare the quantity of CELBAN preparatory materials to general English language skill tests like the IELTS. However, for the purposes of supporting test-takers of the CELBAN it is less important that there is a high volume of preparatory materials, and much more important that there are accurate, clear descriptions of the assessment made widely available to help IENs understand the assessment’s purpose and structure. It is very important to ensure that these existing, or subsequent materials consistently describe the test format and structure so that test-takers understand clearly what they are being assessed on, and how they are expected to demonstrate what they know. Without this critical information, test-takers like Rob, Mary, and Amanda are left surprised by their test taking experience wishing they had known more about what would be expected of
them prior to completing their assessment. Where test-takers are misinformed in regard to what they are being tested on, and how they will be assessed, they are unlikely to be able to best demonstrate their level skill in the given area, thus leading to potential score variance.

The last potential source of construct irrelevant variance which may have effected some test-takers was a topic related to test information, access to the physical test taking supplies required to complete the assessment. Access to note taking paper and pencils challenged various test-takers in their ability to complete the assessment, and to do so without disturbing others around them, particularly at critical assessment stages such as listening where the recording is only played once.

Although some participants indicated they struggled with keeping track of their responses from their question booklet to their answer booklet, some, like Mary indicated they would have appreciated the ability to take notes prior to formulating their final response for their test, which is not allowed for CELBAN (see Theme 6 Test Conditions and Administration). While the provision of scrap paper for note taking would have been appreciated, Mary indicated she was concerned that the notes that she did take may affect the scoring of her answer sheet because she took notes on one of the few pieces of paper provided to her, her Scantron sheet. For Mary, the concern is that her desire to take notes may have inadvertently lowered her score if her notes on her scantron sheet interfere with the computer’s ability to read her recorded answers.

Test-takers are often told to be very cautious in recording answers on scantron sheets especially when erasing responses to ensure that these do not impact the automated grading of the test sheet, and to mark their responses clearly. For Mary (who wrote and
potentially erased several notes from the margins of her scantron sheet) there may be concerns that the computer may not be able to detect all of her indicated responses and that too much erasing could jeopardize the clarity of some of her indicated responses. Where a test-taker’s score is increased or decreased as a result of this technical challenge, the score variance will be irrelevant to the construct being assessed nor attributable to this interference exterior to the testing construct.

The provision of scrap paper may also play a role in providing an assessment experience which closely mimics the realities of job task performance. For nurses in particular, note taking plays an important role in assessing and triaging patient needs. Without the ability to record notes on patient behavior or responses, important pieces of information related to urgency of need, or potential health complications may not be transferred to the attending physician. Although for the purposes of test-taking scenarios where the format may include multiple choice questions, it may be presumed that there is no need to provide additional material for note taking, in the case of nursing where communications often occur in emergency care setting the job function itself requires professionals to take their time and make detailed notes to ensure the accuracy of their communications.

Other concerns related to the provision of test taking supplies include access to a sharpened pencil with which to complete the assessment questions. Penelope explained, that only one sharpened pencil is provided to each test-taker. She raised concerns that this process disadvantaged test-takers who required a sharpener, or fresh pencil as well as those around them because of the need to raise one’s hand to communicate the need for a fresh pencil. This process would delay the test-taker in recording their answers, and
would also create a disturbance for those seated around them. For the listening section in particular there exists a greater degree of concern regarding score variance as a result of a broken pencil as the listening selections are played only one time (The CELBAN Centre E, 2004). Where a test-taker is unable to record their answers to a particular played selection their score will not be a true reflection of their listening skills, but rather a reflection of what they were able to demonstrate of their skill prior to their pencil breaking.

Lastly, a possible source of construct irrelevant variance which was described by some test-takers was the reliability of the assessment protocols related to the speaking section of the CELBAN. Specifically, CELBAN documents suggest that this section of the assessment is intended to be scored by two trained CELBAN Assessors, but this was not always what was reported by test-takers. Penelope described her concern that there was only one assessor in the room which she noted was a challenge as the assessor had to complete so many tasks at once (listening, recording marks, preparing for the next response etc) that they occasionally missed the opportunity to respond to her or had to ask her to repeat herself. Her comment, “It’s very difficult to speak if somebody's not responding [to] you” is particularly noteworthy within the context of this high stakes speaking assessment, which is set up as a role play. Penelope’s observation that there was only one assessor in the room, and her suggestion for two in the future is a good example of a simple modification that can positively impact the test taking experience. Interestingly, this recommendation is also described as an expectation of the speaking assessment in the Test-Taking Strategies booklet featured on The CELBAN Centre website, which states: “During the CELBAN speaking assessment you will be
interviewed in person by two trained CELBAN speaking assessors” (The CELBAN Centre, 2004, p. 8). The provision of two assessors ensures that the one assessing the spoken language skills of the test-taker is not distracted by the added responsibility of providing an authentic role play experience, in which the conversation progresses in a smooth and natural way without delays related to the note taking of the assessor. Having to pause to take notes, or failing to hear the responses of the test-taker as a result of focusing on the scoring of their responses may lead to a variance in the reported score of the test-taker that is not entirely consistent with what was demonstrated.

Do IENs feel the exam tasks provide an accurate reflection of the types of communicative tasks required of a nurse?

CELBAN test developers, designed the assessment to incorporate the communicative tasks typical of a nurse’s day on the job in Canada, and took further steps to incorporate the perspectives of IENs in the pilot testing of the assessment. The construction of the CELBAN included “analysis of target language use, pilot testing drafts of the new assessment tool with the target population, and the use of rigorous measures of reliability and validity through statistical analysis of results” (Epp & Lewis, p. 287). This type of approach echoes the description of authenticity of test tasks offered by Grabowski & Dakin (2014) who define *authenticity of task*, as it refers to test tasks which offer a close approximation of the Target Language Use tasks. In the case of CELBAN the communicative tasks are focused on nursing language and nursing specific communication in Canada. Analysis of authenticity of test tasks differed from prior research in the area of large-scale AELL testing in that studies such as DeLuca et al. (2013) and Cheng & DeLuca (2011) focused analysis on construct representation. This
study challenged interpretation of construct representation with respect to the professional paradigm of practice as described by Austin & Dean (2006). This realization required the creation of Theme 7 Authenticity, a component not included in my original sensitizing framework. At the time of my data collection, I was interested to find out whether test-takers felt the exam tasks accurately captured the types of communicative tasks a nurse would have to complete, in other words, was the assessment perceived to be an authentic reflection of the target language use?

Joe and Mary’s comments on the CELBAN listening section provided useful information on how this assessment endeavors to authentically capture the demands placed on a nurse’s listening skills in a clinical setting. While Joe first contrasted the CELBAN listening section with other language assessment experiences, suggesting the clarity of the recordings were not as good as others he had heard, he later clarified that this was due to the CELBAN’s inclusion of sounds and noises which commonly accompany a clinical setting: “We can hear…some background noises that happens around that situation, in the general atmosphere (Joe).” Mary agreed, explaining that as a nurse, “…you need to be good with your communication, [with] your listening because this is exactly what happens in a nursing setup. You’re listening, there’s a lot of noise going on…you cannot assume things.” Test-takers did not describe a situation in which any of the background noise came into question for the listening assessment, such as an announcement of a particular hospital code which would require nurses in a particular area to report to a supervisor, rather what was described was the challenge of listening to one conversation while amidst the sounds of a bustling hospital, or other clinical setting. Some test-takers, like Amanda, raised concerns that the audio selection is played only
once, and that other background noises may interfere with a person’s ability to record an answer saying, “It [The CELBAN] should not be in the hall—50 people are sitting there making noises and they are expecting us to listen carefully, and…clearly, and once—it’s very hard.” While background noise may be considered a source of score variance in a general test of English listening skills, it is interesting to note that within a nursing specific context it is described by some participants, like Mary, as an assumed reality of the challenges of nursing life in a clinical setting thus reflecting the authenticity of the test tasks—a close approximation of the Target Language Use tasks.

Similarly, other sections of the CELBAN were described with mixed feelings in regard to the way test tasks presented or required IENs to respond with the required target language. Rob acknowledged the realism of the way the listening section demonstrated a nurse-patient interaction in which the nurse indirectly asked questions about the patient’s health in an effort to have them share more openly about their lifestyle choices. This led me to consider the challenges that may arise in testing what is truly heard by an IEN on this test. For example as a nurse you must observe and listen closely to patient appearance and behaviour. In a case where a patient provides an indirect response to a question, especially where this pertains to lifestyle, which was described as a common topic in the roleplaying section of the CELBAN, such as: smoking, or exercise, it may be difficult for a nurse to present the findings of their patient interview in the format of a true or false type response. In nursing practice, they may regularly record the complexity of their interaction with a chart note to clarify whether the patient smokes or exercises regularly.
When asking test-takers about the parts of the assessment that felt like their day-to-day work as a nurse, many offered feedback on the oral assessment component which is a nurse-patient role play. Many participants connected positively to this component of the assessment with respect to how closely it felt like the work they do most often. Many commented on their appreciation for the commitment the oral assessors had to providing a good role playing experience, but some also noted the challenges of attempting to act out a scene which takes place in very unique environment. Jake noted that the goal of the nurse-patient interview is to assess where to send the patient, or how to triage them. Although, test-takers were provided with a general outline for how to advance the role play the intention of the outline it seems was to promote enough oral communication that the test-takers could be accurately scored on their interaction. This provided a feeling of confusion for Jake who noted: “In a real situation there’s a procedure on how you can end a conversation of the nurse patient interaction-but in that form [the role-play] I think it was nothing like that…I was not sure how to—where do I send the patient? What’s the next level here?” Although the oral assessment provides the test-takers with a scenario grounded in practice, the intention of the CELBAN is not to assess the performative aspects of language and linguistic exchange (cultural language competence) which are tied to nursing practice, but rather the communicative ability of the test-takers (receptive and productive language skills). For the purposes of the CELBAN, Jake could have sent the “patient” anywhere, as it was the clarity of the English language being assessed, not the comprehensiveness of the nurse-patient interaction. However, at the same time the communicative act being demonstrated by Jake in that moment was one guided by medical practice which is focused on excellent patient care. The challenge of assessing
one set of skills (language) in an authentic context without putting emphasis on nursing practice is a challenge, it seems, not only for the construction of CELBAN tasks but for the test-takers as well.

The challenge of decoupling language from practice was also described by Stephanie whose comments highlighted the challenges of attempting to authentically represent nursing contexts for IENs whose knowledge of what it means to work as a nurse may vary depending on the epistemology of practice of the country in which they were trained as a nurse. Stephanie explained that one part of her oral assessment required her to attempt to convince a patient to exercise more regularly. She explained that in the country where she had worked previously as a nurse, the concept of changing a patient’s behaviour or attempting to convince them to undertake an action, is a task expected of the physician, not the nurse engaging with the patient. For Stephanie, this task seemed inauthentic in that it required her to perform an action that would not typically be required of her prior to coming to Canada. The challenge of representing professional tasks authentically for the purposes of assessing IEPs as was highlighted through Stephanie’s experience is consistent with the findings of Austin & Dean (2006), who note that “reconceptualizing one’s professional paradigm of practice was a difficult shift for many foreign-trained pharmacists” (26). While the same medical challenge may appear in any country around the world, what is considered the appropriate, or expected manner of communicating is very much dependent upon how the roles and responsibilities of the profession are understood by the country in which the incident takes place. As a result, the type of communicative response on the part of the test-taker may be challenged in terms of length, and comprehensive expression in situations where they are asked to
respond for the purposes of demonstrating linguistic ability, to a scenario that challenges their understanding of how, and if, the professional is intended to respond at all.

**Recommendations for CELBAN administration**

The findings of this research have highlighted the following four recommendations for professional language testing generally and the CELBAN specifically, as it pertains to the evaluative judgement “of the degree to which empirical evidence…supports the…appropriateness of interpretations and actions based on [CELBAN] test scores…” (Messick, 1991, p. 1), or the validity of these test scores as a marker of professional language competency. This study recommends:

1) The licensure and certification requirements, including language assessment requirements, be listed concisely on an official primary webpage to which all applicants are directed

2) CELBAN test locations and dates should be monitored and tracked to support IEN applicants in the pursuit of licensure and certification within the period of their safe practice

3) More official CELBAN assessment preparation materials are needed to support test-takers as well as the continued growth and success of the assessment

4) Minor changes to the administrative protocols of the CELBAN

First, the licensure and certification requirements for IENs should be outlined at an official primary web to which all applicants can be directed. The current method of disseminating licensure and certification information on the CNO website appears on a variety of pages, in conjunction with a series of other sites which offer more detailed information related to specific licensure and certification steps and stages. The CELBAN
Centre website and the Centre for Canadian Language Benchmarks website invite risk of out of date information being perpetuated, as system updates are not coordinated simultaneously across these associations. To achieve this end, a general licensure and certification infographic should be created, maintained and updated on one site or webpage, and made available to partnering associations to ensure a singular message is delivered to IENs. Currently, a 10 step bulleted list describes how to *Become a Nurse in Ontario* when applying from outside of Canada, on the CNO website (2014), as well as a table which features 8 categories described under *Registration Requirements* on the same website (CNO a, 2014), as well as an 8 step bulleted listed also titled *Registration Requirements* which appears on the same website (CNO b, 2014). Although the table and bulleted list identify the same general components such as the need to provide evidence of proficiency in English or French, only the bulleted list provides IENs with additional information of what the specific pieces of acceptable evidence are.

Second, CELBAN test locations and dates must support the IEN applicants in meeting the expectation that they provide evidence of language proficiency within the time allotted for their safe practice. Although at the time of my data collection there were two functioning CELBAN test locations in the province of Ontario, as of May 2016, the CELBAN Centre website listed only one location for IENs to complete the assessment (The CELBAN Centre f, 2015). Moreover, for a period of six-months from May-October of 2016 only 4 test dates were listed for the province of Ontario and offered in one Toronto location (The CELBAN Centre f, 2015). While offering the assessment in a large city centre is certainly useful for many test-takers, this singular test location may be seen as disadvantaging others who are located further from the city centre, as they will be
forced to incur additional costs related to travel and transportation to the sole test location in the province. This is a significant factor in terms of the potential impact on CELBAN scores, where some test-takers will travel less than 5 kilometres on the day of an assessment lasting several hours, and some will travel in excess of 188 kilometers, as in the case of Emily who travelled to Toronto from Belleville.

Third, official CELBAN preparation materials are required to support the continued growth and success of the assessment. Currently, the self-assessment tool available to test-takers is described as providing “an overview of the test format and content, so you will have a better idea of what to expect during the test” (The CELBAN Centre D, 2015). However, the assessment page does not specifically reference the difference between the online format of the self-assessment tool and the paper-and-pencil format of the actual CELBAN, leading to some confusion for test-takers such as Amanda. Although the CELBAN Centre has endeavored to publish additional assessment preparation resources, in the absence of official CELBAN preparation courses or workbooks, enterprising individuals will continue to advertise themselves as experts in CELBAN preparation for a significant fee, despite their lack of understanding and accuracy of information presented as described by Mary. Official CELBAN preparation materials should clearly outline the language competencies assessed by each section of the assessment as outlined by the 235 page Canadian Language Benchmark document currently available by way of a hyperlink on the CELBAN Centre website, but not described on that site (Centre for Canadian Language Benchmarks, 2014).

Lastly, minor changes to the administration of the CELBAN should be adopted including: the provision of two pencils, scrap paper, and two assessors for the oral
English section of the assessment to decrease the opportunities for disruption during the assessment, and ensure consistency among assessment administrations. The expectation that there be two assessors to participate in the scoring of the oral English section of the CELBAN is described as the general practice for the CELBAN in the Test-Taking Strategies booklet featured on the CELBAN Centre website (The CELBAN Centre, 2004, p. 8), however it seems this protocol has at least on one occasion been variable (see the case with Penelope).

**Limitations and Directions for Future Research**

This study represents the qualitative findings of my focus groups and individual interviews with only XVI participants at two CELBAN testing locations in the Province of Ontario in the winter of 2015. Future research and future administrations of the CELBAN may consider whether or not the perceptions of test constructs, and feelings on testing conditions and administration obtained by this study are shared by test-takers in other test locations and provinces. A survey of a large sample of CELBAN test-takers may also help to determine if the results of this study are generalizable to a wider population, particularly in regard to the reported CELBAN administration and preparation experiences. Statistical analysis of specific survey question types on the congruence between what test-takers perceive to be tested and what the test questions are designed to assess, may highlight how best to communicate the specific skills IENs must demonstrate to show English language competency, as well as where and how that message is best being conveyed currently.

Another limitation of this study is related to its restricted data analysis. The data were coded primarily by myself in consultation with my committee and a colleague also
completing a Master of Education degree, and guided by my sensitizing concepts. I applied deductive codes which emerged from the literature on large-scale, high-stakes English language assessment, and created new codes where needed. A colleague in education reviewed a subset of the data and compared their coding structure with my own. This process encouraged a greater degree of inter-rater reliability, and encouraged the reduction of emergent codes to key differences in test-taker experiences and descriptions. However, the review of data coding by my colleague was limited to a small subsection of two pages of data transcription. It is entirely possible that acting from my own subject positions, I have approached the coding of this data through a number of biased lenses with which I view the world. In future research it would be valuable to collaborate with a nursing professional to review the perceptions of test constructs as well as the role that professional knowledge plays in demonstrating language competency in a language assessment presented in the format of profession specific scenarios and contexts. This type of research could be used to further refine the types of scenarios posed and the way in which assessment responses are assessed for accuracy. Future research on the experiences of test-takers who are unable to provide evidence of language proficiency in their first attempt on the CELBAN, would also be very useful in its ability to provide additional information on perceptions of the testing constructs following a first attempt with the assessment and the receipt of their score report, as well as providing further insight into the wider experience of IENs in their pathways to licensure and certification.
Conclusion

The CNO has estimated an attrition rate of forty percent among Registered Nurses educated abroad who apply for membership (Blythe et al., 2009). This estimated loss of IENs to the profession suggests a need for further research on how best to support federally recruited IENs in their pursuit of licensure and certification with the CNO, thus facilitating their subsequent economic integration in Canada. The CELBAN is intended to facilitate integration by recognizing the professional knowledge of IENs obtained in their home countries, and thus focuses language testing specifically on the language of the profession (Epp & Lewis, 2009). As a national testing instrument of professional language competency, the CELBAN judges what constitutes proficiency based on the receptive and productive language skills demonstrated by the test-takers on their given test date. This type of approach to language skill assessment concurs with Kerns (2004) critique on instructional approaches which organize language curriculum by having students work through a sequence of phrases, sentences, paragraphs, and finally extended discourse. Kern acknowledges these may be “eminently logical, but do not mesh well with the psychological needs of language learners who strive to communicate in meaningful, whole acts” (p. 10). Recognition of the motivational and psychological factors which may impact economic integration of IEPs is particularly important for Canada where research has reported only 19-40% of Internationally Educated Professionals (IEPs) obtain suitable employment in their professional field (Ngo & Este, 2006). And others have noted that “non-recognition of international credentials and challenges in satisfying the licensure and certification testing requirements… cost the Canadian economy as much as $5 billion a year” (Cheng et al., 2013 p. 734).
Empirical research on the CELBAN is primarily limited to studies conducted at the time of the development of the assessment (Epp & Lewis, 2009). Systematic searches of academic journals in Canadian nursing reveal very limited research on the topic of the CELBAN (The Canadian Nurse, 2005; Infermiere Canadienne, 2005). While this study is exploratory in nature, it offers an in-depth understanding of test-taker perceptions on the CELBAN thus contributing to our understanding of the meaning and inferences we make from its assessment score (Messick, 1996). Additional CELBAN research is needed to continue to support IENs in their attainment of licensure and certification to practice, as well as the national vision for economic and social integration of Internationally Educated Professionals into our Canadian communities of practice.
References


C. Addison Stone, Elaine R. Silliman, Barbara J, Ehren, & Kenn Apel (eds) 
Handbook of language and literacy. (pp. 72-81). New York: The Guilford Press.


Retrieved from

http://www.citizenship.gov.on.ca/english/keyinitiatives/bridgetraining.shtml


*Infirmiere Canadienne*, 6(5), 12.


121
www.celbancentre.ca/test-taking-strategies.aspx


comparative examination of Indian and Philippine internationally educated nurses
and their entry into Ontario’s nursing profession. (*CERIS Working Paper No. 92*),
content/uploads/2012/11/CWP92_Walton-Roberts_Hennebry.pdf


experienced by visible-minority social workers in mainstream social service
agencies: a community project. (*CERIS Working Paper No.57*). Retrieved from
CERIS Working Paper Series website: http://www.ceris.metropolis.net/wp-
content/uploads/pdf/research_publication/working_papers/wp57.pdf
Appendix A

Ethics Letter of Approval

December 17, 2015
Mrs. Stefanie Bojarski
Master's Student Faculty of Education Queen's University
Duncan McArthur Hall 511
Union Street West
Kingston, ON, K7M 5R7
GREB Romeo#: 6014138
Title: “GEDUC-757-14 Proving My Competency One Test at a Time: Internationally Educated Nurses and the CELBAN”

Dear Mrs. Bojarski:

The General Research Ethics Board (GREB) has reviewed and approved your request for renewal of ethics clearance for the above-named study. This renewal is valid for one year from December 2, 2015. Prior to the next renewal date you will be sent a reminder memo and the link to ROMEO to renew for another year.

You are reminded of your obligation to advise the GREB of any adverse event(s) that occur during this one year period. An adverse event includes, but is not limited to, a complaint, a change or unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours. To submit an adverse event report, access the application at https://eservices.queensu.ca/ROMEO.Researcher.Admin; click on "Events"; under "Create New Event" click on "General Research Ethics Board Adverse Event Form”.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example you must report changes in study procedures or implementation of new aspects into the study procedures. Your request for protocol changes will be forwarded to the appropriate GREB reviewers and/or the GREB Chair. To submit an amendment form, access the application at https://eservices.queensu.ca/RO/iiEO.Researcher.Admin; click on "Events"; under "Create New Event" click on "General Research Ethics Board Request for Amendment of Approved Studies”.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Yours sincerely,

Joan Stevenson
Ph.D. Chair
General Research Ethics Board

C.: Dr. Liying Cheng, Faculty Supervisor
Dr. Liying Cheng, Chair, Unit REB
Ms. Erin Wicklam, Dept. Admin.
Appendix B

Semi-Structured Focus Group Guide

I’d like to begin with a general view of your testing experience.

1. What was the physical environment of the test like? What was the general atmosphere?
2. Were you aware of any observers in the room?
3. Were there any technical difficulties during the assessment?
4. How did you ‘feel’ as a result of the experience?
5. Tell me in general about your experience of the CELBAN.

Having talked in general about the test experience, I’d like to focus now on the test itself.

6. How do you feel the assessment tested your English language skills?
7. How do you feel the assessment tested your professional nursing skills?
8. Do you feel the exam tasks were similar to what you would have to do as a nurse while working in Canada?
9. Did you read the instructions before engaging with each section and task?
10. Which section(s) and or task(s) did you feel was the most difficult? Why?
   - **Writing** there are two types of writing tasks:
     11) Do you feel that the form filling section was an accurate reflection of the type of writing that you would do as a nurse working in Canada?
     12) Do you feel that the report writing section was an accurate reflection of the type of writing you would do as a nurse working in Canada?
   - **Reading** There are two types of reading tasks
     13) Do you feel that the types of skills tested by the reading section tasks are similar to those required by the work of a nurse in a Canadian hospital?
     14) Would you say each reading task was equally challenging?
15) Were there any questions that were confusing or that you did not understand?

- **Listening**

16) Do you feel that the types of scenarios portrayed in the video were realistic for an Ontario hospital setting?

17) Do you feel the types of scenarios recorded in the audio section were realistic for a Canadian hospital setting?

- **Final Thoughts**

18) Any other comments you would like to make?
Appendix C

Semi-Structured Interview Guide

Speaking

1) Describe your experience with the oral interview.
2) How do you feel the oral speaking assessment tested your English language skills?
3) How do you feel this section tested your professional nursing skills?
4) Describe your experience with the role plays.
5) Was the timing appropriate?
6) Were you aware of any observers in the room as you completed the assessment?
7) Did you feel you had enough time or should the time allowed have been longer?
8) Did you take notes?
9) Do you feel these types of tasks are realistic in terms of the interactions you have in a Canadian hospital?
10) Any other observations about the speaking section?

Appendix D

Letter of Information & Consent Form

“Proving My Competency One Test at a Time: Internationally Educated Nurses and the CELBAN”

This research is being conducted by Stefanie Bojarski under the supervision of Dr. Liying Cheng, in the Faculty of Education at Queen’s University in Kingston, Ontario.

What is this study about? The purpose of this research is to learn more about the CELBAN through your experience taking it. Specifically, this study will try to determine what factors affect test-taker’s scores on the CELBAN based on Internationally Educated Nurses responses to the different sections of the exam, as well as the environmental factors that may impact their experience. If you choose to participate in this study you will be asked to participate in a focus group and individual interview following your exam. In the focus group you will speak with other test-takers and the researchers about your experience writing the CELBAN. In the interview you will discuss your experience of the speaking assessment of the CELBAN with the researcher. There are no known physical, psychological, economic, or social risks associated with this study.

Is my participation voluntary? Yes. This research is being conducted independently of CELBAN and those that will be marking your exam. It would be greatly appreciated if you would answer all material as frankly as possible, but you should not feel obliged to answer any material that you find objectionable or that makes you feel uncomfortable. You may also withdraw at any time with no effect on your exam, or your standing with the College of Nurses Ontario.

What will happen to my responses? We will keep your responses confidential. Only researchers will have access to this information. The data may also be published in professional journals or presented at scientific conferences, but any such presentations
will be of group findings and will never breach individual confidentiality. Should you be interested, you are entitled to a copy of the findings.

**Will I be compensated for my participation?** No. However, as a token of appreciation you will receive a free pizza dinner with non-alcoholic beverages immediately following the group sitting of the CELBAN and a gift card available to use at Shoppers Drug Mart for your participation in this study.

**What if I have concerns?** Any questions about study participation may be directed to Stefanie Bojarski at stefanie.bojarski@queensu.ca. Any ethical concerns about the study may be directed to the Chair of the General Research Ethics Board at chair.GREB@queensu.ca or 613-533-6081.

Again, thank you. Your interest in participating in this research study is greatly appreciated.

*This study has been granted clearance according to the recommended principles of Canadian ethics guidelines, and Queen's policies.*
Consent Form
“Proving My Competency One Test at a Time: Internationally Educated Nurses and the CELBAN”

Name (please print clearly): ________________________________________

1. I have read the Letter of Information and have had any questions answered to my satisfaction.

2. I understand that I will be participating in the study called Proving My Competency One Test at a Time: Internationally Educated Nurses and the CELBAN. I understand that this means that I will be asked to participate in a focus group and interview immediately following the group sitting and individual sitting of my CELBAN exam.

3. I understand that my participation in this study is voluntary and I may withdraw at any time. I understand that every effort will be made to maintain the confidentiality of the data now and in the future. Only researchers in this study will have access to this area. The data may also be published in professional journals or presented at scientific conferences, but any such presentations will be of general findings and will never breach individual confidentiality. Should I be interested, I am entitled to a copy of the findings.

4. I am aware that if I have any questions, concerns, or complaints, I may contact Stefanie Bojarski; stefanie.bojarski@queensu.ca; project supervisor and director of the Assessment and Evaluation Group, Dr. Liying Cheng 613-533-6000 extension: 77431; living.cheng@queensu.ca; or the Chair of the General Research Ethics Board (613-533-6081) at Queen’s University.

I have read the above statements and freely consent to participate in this research:

Signature: ________________________________
Date: ___________________________
Appendix E

Demographic Information: Sex, Age, Country of Origin & First Language

Table 8.

Demographic Information: Sex, age, country of origin, & first language

<table>
<thead>
<tr>
<th>Sex</th>
<th>Pseudonym</th>
<th>Age</th>
<th>Country of Origin</th>
<th>First Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>Sandy</td>
<td>30-40</td>
<td>Pakistan</td>
<td>Urdu</td>
</tr>
<tr>
<td>Female</td>
<td>Mary</td>
<td>30-40</td>
<td>India</td>
<td>Thamil</td>
</tr>
<tr>
<td>Female</td>
<td>Penelope</td>
<td>20-30</td>
<td>India</td>
<td>Gujarati</td>
</tr>
<tr>
<td>Male</td>
<td>Jake</td>
<td>30-40</td>
<td>South Sudan</td>
<td>Bari &amp; Arabic</td>
</tr>
<tr>
<td>Male</td>
<td>Joe</td>
<td>20-30</td>
<td>India</td>
<td>Malayalam</td>
</tr>
<tr>
<td>Male</td>
<td>Rob</td>
<td>20-30</td>
<td>India</td>
<td>Malayalam</td>
</tr>
<tr>
<td>Female</td>
<td>Stephanie</td>
<td>30-40</td>
<td>India</td>
<td>Gujarati</td>
</tr>
<tr>
<td>Female</td>
<td>Anna</td>
<td>20-30</td>
<td>India</td>
<td>Hindi</td>
</tr>
<tr>
<td>Female</td>
<td>Nancy</td>
<td>20-30</td>
<td>Nepal</td>
<td>Nepali</td>
</tr>
<tr>
<td>Female</td>
<td>Emily</td>
<td>30-40</td>
<td>India</td>
<td>Malayalam</td>
</tr>
<tr>
<td>Female</td>
<td>Janet</td>
<td>20-30</td>
<td>Philippines</td>
<td>Filipino</td>
</tr>
<tr>
<td>Female</td>
<td>Rose</td>
<td>20-30</td>
<td>India</td>
<td>Punjabi</td>
</tr>
<tr>
<td>Female</td>
<td>Jessie</td>
<td>20-30</td>
<td>India</td>
<td>Punjabi</td>
</tr>
<tr>
<td>Female</td>
<td>Amanda</td>
<td>20-30</td>
<td>India</td>
<td>Punjabi</td>
</tr>
<tr>
<td>Gender</td>
<td>Name</td>
<td>Age</td>
<td>Country</td>
<td>Language</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Male</td>
<td>Sam</td>
<td>20-30</td>
<td>Philippines</td>
<td>Tagolog</td>
</tr>
<tr>
<td>Male</td>
<td>Syd</td>
<td>20-30</td>
<td>India</td>
<td>Malayalam</td>
</tr>
</tbody>
</table>

Note 1. Yellow represents Hamilton
Note 2. Green represents Toronto
Appendix F

Demographic Information: Language & Nursing Experience

Table 9.

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Language Spoken at Home</th>
<th>County Studied</th>
<th>Language of Nursing Instruction</th>
<th>Years of Nursing Practice</th>
<th>Years studying English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy</td>
<td>Urdu</td>
<td>Pakistan</td>
<td>English</td>
<td>8</td>
<td>6+</td>
</tr>
<tr>
<td>Mary</td>
<td>Thamil</td>
<td>India</td>
<td>BLANK</td>
<td>3</td>
<td>Since 5yrs old</td>
</tr>
<tr>
<td>Penelope</td>
<td>Gujarati &amp; English</td>
<td>India</td>
<td>English</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Jake</td>
<td>Bari &amp; English</td>
<td>Kenya</td>
<td>English</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Joe</td>
<td>Malayalam</td>
<td>India</td>
<td>English</td>
<td>1.6</td>
<td>21</td>
</tr>
<tr>
<td>Rob</td>
<td>Malayalam</td>
<td>India</td>
<td>English</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>Stephanie</td>
<td>Gujarati</td>
<td>India</td>
<td>English</td>
<td>8</td>
<td>Since 3yrs old</td>
</tr>
<tr>
<td>Anna</td>
<td>Punjabi</td>
<td>India</td>
<td>English</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Nancy</td>
<td>Nepali</td>
<td>Nepal</td>
<td>English</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Emily</td>
<td>Malayalam</td>
<td>United Arab Emirates</td>
<td>English</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Janet</td>
<td>Filipino</td>
<td>Philippines</td>
<td>English</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>Rose</td>
<td>Punjabi &amp; English</td>
<td>India</td>
<td>English</td>
<td>3.5</td>
<td>16</td>
</tr>
<tr>
<td>Jessie</td>
<td>English</td>
<td>India</td>
<td>English</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>Amanda</td>
<td>Punjabi</td>
<td>BLANK</td>
<td>English</td>
<td>4</td>
<td>BLANK</td>
</tr>
<tr>
<td>Name</td>
<td>Language</td>
<td>Country</td>
<td>Language</td>
<td>Week 1</td>
<td>Week 2</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>-------------</td>
<td>----------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Sam</td>
<td>English</td>
<td>Philippines</td>
<td>English</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>Syd</td>
<td>Malayalam</td>
<td>India</td>
<td>English</td>
<td>8</td>
<td>20</td>
</tr>
</tbody>
</table>

Note 1. Yellow represents Hamilton
Note 2. Green represents Toronto
# Appendix G

## Demographic Information: Language Assessment & Work

Table 10.

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Prior English Assessment</th>
<th>Prior Assessment Levels</th>
<th>Currently Working in Health Care</th>
<th>Part/ Full Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy</td>
<td>Canadian Language Benchmark</td>
<td>Reading - 8 Writing - 8 Listening - 8 Speaking - 8</td>
<td>Family Physician Clinic</td>
<td>Full Time</td>
</tr>
<tr>
<td>Mary</td>
<td>Canadian Language Benchmark</td>
<td>Reading- 8 Writing - 8 Listening - 8 Speaking - 8</td>
<td>BLANK</td>
<td>BLANK</td>
</tr>
<tr>
<td>Penelope</td>
<td>Canadian Language Benchmark</td>
<td>Reading - 8 Writing - 7 Listening - 8 Speaking - 7.5</td>
<td>PSW</td>
<td>BLANK</td>
</tr>
<tr>
<td>Jake</td>
<td>Canadian Language Benchmark</td>
<td>Reading - 8 Writing - 6 Listening - 9 Speaking - 8</td>
<td>CORS Management</td>
<td>BLANK</td>
</tr>
<tr>
<td>Joe</td>
<td>CELPIP</td>
<td>Reading - 7 Writing - 7 Listening - 8 Speaking - 8</td>
<td>Lab Assistant Niagara College</td>
<td>BLANK</td>
</tr>
<tr>
<td>Rob</td>
<td>CELPIP</td>
<td>Reading- 6.5 Writing- 6.5 Listening- 7.5 Speaking- 8</td>
<td>PSW</td>
<td>Full Time</td>
</tr>
<tr>
<td>Stephanie</td>
<td>IELTS</td>
<td>Speaking -8 BLANK</td>
<td>BLANK</td>
<td>BLANK</td>
</tr>
<tr>
<td>Anna</td>
<td>BLANK</td>
<td>BLANK</td>
<td>No</td>
<td>no</td>
</tr>
<tr>
<td>Name</td>
<td>Test</td>
<td>Reading</td>
<td>Writing</td>
<td>Listening</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Nancy</td>
<td>Canadian Language Benchmark</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Emily</td>
<td>IELTS</td>
<td>7</td>
<td>7</td>
<td>7.5</td>
</tr>
<tr>
<td>Janet</td>
<td>Benchmark</td>
<td>7.5</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Rose</td>
<td>BLANK</td>
<td>BLANK</td>
<td>BLANK</td>
<td>BLANK</td>
</tr>
<tr>
<td>Jessie</td>
<td>Benchmark</td>
<td>8</td>
<td>8</td>
<td>8.5</td>
</tr>
<tr>
<td>Amanda</td>
<td>Benchmark</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Sam</td>
<td>IELTS</td>
<td>7</td>
<td>6.5</td>
<td>8</td>
</tr>
<tr>
<td>Syd</td>
<td>IELTS</td>
<td>yes - DSW</td>
<td>7</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Note 1. Yellow represents Hamilton
Note 2. Green represents Toronto
Appendix H

Demographic Survey

1) Would you describe yourself as:

   Transgender
   Female
   Male

2) Please indicate your age range:

   -20
   20-30
   30-40
   40-50
   50-60
   60+

3) What is your country of origin

   ________________________________________________

4) What is your first language

   ________________________________________________

5) What language do you speak at home

   ________________________________________________

6) In what language did you receive your nursing education

   ________________________________________________

7) In what country did you study nursing

   ________________________________________________

8) How many years did you practice as a nurse before coming to Canada

   ________________________________________________
9) How long have you been studying English


10) Are you working in the health care sector currently? (If yes, please specify the type of job)
