International Graduate Students' Perspectives on High-Stakes English Tests and the Language Demands of Higher Education

SHAKINA RAJENDRAM
University of Toronto

ELIZABETH JEAN LARSON
University of Toronto

JEANNE SINCLAIR
University of the Incarnate Word

Abstract

The internationalization of higher education in Canada has given rise to the increased use of standardized English language proficiency tests as gatekeeping measures in university admission policies. However, many international students who are successful on these tests still struggle with the academic and language demands of their programs. Drawing on a thematic analysis of life story interviews with five international graduate students at a major Canadian university, this study examines students' perceptions on the skills elicited by the IELTS and TOEFL, the language demands and pragmatic norms of their graduate program in language education, and the university's language support programs.

Keywords

language assessment; language proficiency; international students; higher education; university supports

Introduction

Recent decades have brought a rapid expansion of international students in higher education. In 2016, more than 4.8 million students were estimated to be enrolled in higher education institutions outside their country of origin (UNESCO, 2018). Concerns about the English language proficiency of international students who speak English as an additional language have given rise to the increased use of standardized English language proficiency tests such as the International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) in Canadian university admissions. The language requirements of universities are often based on the assumption that if students demonstrate English proficiency prior to admission, they will be able to function successfully in their new English-language environment and succeed in their program (Benzie, 2010).

However, students' performance on proficiency tests may not reflect the actual proficiency and skills they need to succeed in higher education, particularly in graduate-level coursework (Friedenberg, 2002; Schmidt & Gannaway, 2007). International students who have been successful on the TOEFL or IELTS tests have been shown in various studies to face difficulties post-admission in their listening and oral communication skills, academic writing, knowledge of local contextual references, pragmatic competence, and

language and discipline-specific knowledge and discourse (Mori, 2000; Sawir, Marginson, Forbes-Mewett, Nyland & Ramia, 2012; Trice, 2003; Zhang & Beck, 2014). Furthermore, despite their diverse linguistic and academic needs, international graduate students are often treated as a homogenous group in higher education institutions (Zhang, 2016) without personalized resources and support to navigate the challenges they face in their graduate programs (Ambrósio, Marques, Santos & Doutor, 2017).

The studies above engage the potential mismatch between the language skills international graduate students have upon applying or enrolling in their programs, and the skills they actually need for success in those programs, as well as the potential mismatch between the language supports international students need, vis-a-vis what higher education institutions provide. However, to date, there is scant research examining international graduate students' perspectives on these as interwoven, highly connected matters. Thus, this paper explores their perspectives on the skills elicited by the language proficiency tests and the actual language and pragmatic demands of higher education institutions, and how they navigate these demands both with and without university support. Using data gathered from in-depth interviews, the research questions guiding this study ask:

1. What are international graduate students' perceptions about the relationship between the skills elicited by large-scale language proficiency tests and the language demands of their graduate program?
2. How do they perceive the pragmatic norms of their program, and how do these differ from their previous educational experiences?
3. What are their perceptions of the effectiveness of the university's language support programs in helping them navigate these language demands, and what additional supports do they seek?

Theoretical Framework

This study is framed by Elena Shohamy's (2001a, 2001b, 2017) theory of Critical Language Testing (CTL). Essentially, CLT is a paradigm for research on language testing and assessment that focuses on the consequences (intended and otherwise) and uses of tests in education as well as society at large. Specifically, CLT focuses on the power that language tests exert which grants people status in society so that they change their behavior so they succeed on the tests. Test-developers and test-score users control the standards, content, format, scoring, reporting, and use of high-stakes language tests with little to no input from test-takers; nor is there a clear path for recourse for test-takers' concerns despite the significant financial cost of such tests -- which some test-takers pay multiple times. Thus, power is allocated unevenly -- and often opaquely -- throughout the high-stakes decision-making process (i.e., higher-education admission) that languages tests validate. Validity is the core question and issue of CLT. A test that is valid, that is, appropriate and fair inferences can be generated from its scores (Messick, 1989), would present little cause for concern to test-takers and scholars and activists concerned about this issue. However, validity is very challenging to achieve. In decades past, a test was considered valid if it correlated with another test or "real-world performance" that purportedly measured the same construct (Shepard, 1993). The focus of validation then turned to test content as the arbiter of validity: how thoroughly the test represented, and was relevant to, the target domain became the prime basis for test validation (Sireci, 1998). While representation and

fidelity to the target domain remain important dimensions of validity, contemporary test validation has evolved toward a unified theory centering on the construct(s) elicited by the measure, and the adequacy and appropriateness of the measure's score interpretation, including the personal and social consequences of score use (Messick, 1989; Shepard, 2016). This model forms the foundation for test validation in the current *Standards for Educational and Psychological Testing* (2014), co-developed by the American Educational Research Foundation, the American Psychological Association, and the National Council on Measurement and Evaluation. Contemporary frameworks emphasize that one cannot validate a *test*; rather, the process of validation evaluates *in situ* test score interpretation and use (Mislevy, Almond & Lukas, 2003; Mislevy & Haertel, 2006).

Bachman (1990) is widely cited as bringing the measurement community's developments in validation to the forefront of language testing research and practice (Chapelle, 1999; Spolsky, 2008). Bachman and Purpura (2008) posit that language testing programs can operate as both gatekeepers and door-openers, but the perception of the test as one or the other will depend on one's position vis-à-vis the test. Stakeholders who experience success are more likely to feel the test opens doors, and vice versa for those who experience failure. Bachman and Purpura (2008) note that fundamentally, there is a question of claiming responsibility for deciding how scores will be interpreted and used:

> … [L]anguage tests are used to provide score-based information for making a wide range of decisions… As a consequence of these decisions, some people are rewarded and some are not.… Who decides to use a language assessment, as opposed to other types of information, for allocating these rewards? Who decides where to set the cut-score that divides those who will receive the rewards from those who will not?… Ultimately, the issue of who decides is, in our view, one that involves societal, cultural, and community values that are beyond the control of the language test developer. Nevertheless, these values need to be carefully considered as we design, develop, and use language assessments. (p. 465-466)

Bachman and Purpura posit that test-developers are not responsible for decision-making around score interpretation. Yet, there is no suggestion as to who *can* be held responsible. In absence of a responsible party, it remains unclear where test-takers can seek recourse regarding inappropriate interpretations or other unintended consequences of test score use. Within the framework of CLT, the present study aims to contribute to the conversation around language tests' validity by examining international graduate students' perspectives on their relevance and impact to their graduate education.

Literature Review

O'Loughlin (2011) directly engages the question of responsibility for the appropriate interpretation and use of test scores, which he argues should be shared by test developers and score users. The study in which this recommendation is situated concerns the decision-making processes and language-assessment literacy of one Australian university's admissions office, with a focus on the IELTS exam. O'Loughlin found that the minimum entry score was set without a principled basis, there were no internal studies to validate the minimum score requirement, nor were additional relevant criteria considered in admissions decisions, as suggested by the IELTS handbook. IELTS scores have been

shown to be significantly correlated with GPA in some studies (i.e., Hill, Storch, & Lynch, 1999; Schoepp, 2018), but not in others (Cotton & Conrow, 1998), although this may have been due to sample size. Hill et al. also found TOEFL scores to not be significantly correlated with GPA.

One of the first steps in developing a test is to choose a construct or constructs to be measured. Both IELTS Academic and TOEFL iBT purport to measure academic English language proficiency, albeit with slightly different construct definitions. The test developers for IELTS have chosen the construct of, "English language proficiency needed for an academic, higher learning environment" (IDP Education Canada, 2019), while those who develop TOEFL pick, "your ability to use and understand English at the university level" (Educational Testing Service, 2019). These constructs appear to be quite similar, but their approach to measuring them is rather different. Both TOEFL and IELTS use lectures and academic writing for the listening and reading sections, but their item types and task timings are quite different. Table 1 below compares some aspects of these tests.

Table 1

Comparison of TOEFL iBT and IELTS Academic

	IELTS Academic	**TOEFL iBT**
Sections	Reading(R), Listening (L), Writing (W), Speaking (S)	Reading(R), Listening (L), Writing (W), Speaking (S)
Timing per section	Reading – 60 minutes Writing – 60 minutes Listening – 30 minutes Speaking – 11-14 minutes	Reading – 60-80 minutes Writing – 50 minutes Listening – 60-90 minutes Speaking – 20 minutes
Items/tasks per section	Reading – 40 multiple choice items Writing – 2 writing tasks Listening – 4 sections Speaking – 3 tasks	Reading – 36-56 questions Writing – 2 writing tasks Listening – 34-51 questions Speaking – 6 tasks
Scores	Each section: band score of 0-9, with half-points allowed	Each section: 0-30
Format	Computer-based or paper-based for 3 sections and face-to-face speaking	Computer-based only
Maximum score	9 (averaged overall band score)	Combined score of 120

Large-scale language-test developers such as TOEFL and IELTS have led substantial research programs to establish the validity of their measures' use in higher education settings. The validation of TOEFL has been published in book form (Chapelle, Enright, & Jamieson, 2008), following Kane's (1992) interpretative/use and validity

arguments based on Toulmin's (1958) argumentation protocol. The argument for TOEFL's validity begins, per Kane's (1992) protocol, by making a case that TOEFL's tasks elicit performances that reveal "relevant knowledge, skills, and abilities" (p. 347) representative of those needed in English-medium institutions of higher education (domain description). Targeted language abilities in these performances are scored (evaluation), scores are assumed to be parallel across tasks, test forms, and raters (generalization). The authors describe qualitative and quantitative research that connect the constructs of language tested by the TOEFL, the TOEFL test itself, and the target domain (explanation and extrapolation). In essence, these argue that TOEFL scores are attributable to certain language proficiency constructs, and that those constructs are the ones that constitute academic language proficiency as operationalized in English-medium universities, with TOEFL linking the two. These arguments in turn are used to support the idea that TOEFL scores are useful for admissions and curriculum decisions (utilization). IELTS' validity arguments take the form of multiple such investigations, as well as one book-length validity study of the IELTS listening test (Aryadoust, 2013). While largely comprehensive, Chapelle et al. acknowledge that although some research support is included for test score utilization, "additional backing is needed through studies of score use and the consequences of TOEFL score use" (p. 346).

Despite substantial work toward validating TOEFL and IELTS as useful assessments for international students applying to higher education, many international students who obtain the required admission scores on English language proficiency tests continue to struggle with the academic language demands of their program of study (Andrade, 2009). In addition to the need for strong general academic language skills, international *graduate* students also need to learn the disciplinary discourse, discipline-specific content and conceptual structures of their graduate courses, and adjust to the heavy reading requirements of graduate classes (Beaven, Calderisi, & Tantral, 1998; Lin & Scherz, 2014; Lin & Yi, 1997). Furthermore, as graduate programs usually have a significant writing component such as a thesis or major research project, international students in graduate programs are faced with demanding academic writing requirements. However, studies such as Riazi's (2016) have found that although the textual features of the writing produced by graduate students on English language proficiency tests may be similar to their written academic assignments, the tests do not account for discipline-specific content knowledge, vocabulary, argumentation, and writing genres.

Other academic difficulties reported by international students include learning how to participate in tutorials, critiquing peers' ideas, engaging in debates, delivering oral presentations, and communicating with professors (Andrade, 2006, 2009; Berman & Cheng, 2001; Hellsten & Prescott, 2004; Holmes, 2004; Robertson, Line, Jones & Thomas, 2000; Schmidt & Gannaway, 2007; Storch & Hill, 2008; Wong, 2004).

International graduate students are also faced with the challenge of adjusting to an education system that differs from what they are accustomed to, within a new linguistic and cultural environment. As Spurling (2007) states, there is a tendency for universities to see international students only as "subject learners," overlooking their needs as "culture and language learners" (p.114). International graduate students have reported experiencing culture shock and homesickness, and have expressed difficulties forming social relationships with local students because of different cultural expectations for relationship-building (Lin & Scherz, 2014; Marr, 2005). Ravichandran, Kretovics, Kirby and Ghosh

(2017) argue that the existing language proficiency tests and university supports available to international graduate students are inadequate for helping them navigate these difficulties. They maintain that obtaining the minimum score in the tests required for admission into university does not ensure that students have the actual proficiency that they need to be successful in their program.

Consistent with this, research has suggested that IELTS and TOEFL scores are not accurate indicators that students have the language competencies that are required in higher education. For example, international students who have done well on these tests have reported encountering significant difficulties post-admission in language domains such as listening, oral communication, vocabulary, knowledge of local contextual references, mastery of discipline-specific discourse, and academic writing (Feast, 2002; Mori, 2000; Ryan & Zuber-Skerritt, 1999; Sawir et al., 2012; Trice, 2003; Zeegers and Barron, 2008). In a study exploring her own journey as an international student in Canada, Liu (2011) writes that although she had succeeded in both the IELTS and the TOEFL, she had a very difficult time adjusting linguistically, psychologically and socially to life as a student in a Canadian university: "I could not understand what my instructors and classmates were talking about… I thought that I looked stupid in class because during discussions I could not express myself clearly in English…I felt frustrated with my English." (p. 79-80). As a result of these language struggles, Liu experienced symptoms of culture shock, loss of identity, lack of confidence, and low self-esteem. Expressed in her own words: "I lost my sense of achievement and comfort" (p. 79).

Other studies have found that just like Liu (2011), many other international students who were successful in pre-admission language proficiency tests faced difficulties coping with the English language demands of their programs, and these difficulties resulted in negative academic, psychological, social and emotional effects such as self-doubt, low self-confidence, segregation, anxiety, loneliness, ineffective group work, and hostility towards the host country (Andrade, 2009; Brown, 2008; Burnett & Gardner, 2006; Friedenberg, 2002; Khan, 2009; Kodama, 2007; Trice, 2003). In a study of 900 international students in Australia, it was found that 41% of them experienced high stress levels caused by several factors including cultural shocks, as a result of not being able to understand and adjust easily to the implicit and explicit culture and customs of their host institution and country (Rienties, Beausaert, Grohnert, Niemantsverdriet & Kommers, 2012).

Schoepp's (2018) research on the predictive validity of IELTS scores as they relate to undergraduate students in an English-medium university in the United Arab Emirates found that IELTS scores were meaningful predictors of students' academic success, as determined by their grade point average (GPA). However, Andrade (2006) and Zhou, Frey, and Bang (2011) argue that the success of international students should be determined not just through their GPA scores, and recommend that universities should conduct more interviews, surveys, and focus groups with international students to better understand the challenges they face and their academic needs. Our study responds to this call by conducting in-depth interviews with international graduate students to gain a better understanding of their experiences with taking language proficiency tests for admission into their graduate programs, the challenges they face post-admission, and how they navigate these challenges.

Methodology

Research Setting and Participants

Our research took place in a major Canadian English-medium university (hereafter known as Metro University). The majority of graduate programs in Canadian universities require proof of English proficiency for any student who is applying from a university outside Canada where English is not the primary language of instruction. TOEFL and IELTS (amongst other tests) scores are one way of providing such proof. However, the required minimum scores vary somewhat across, and sometimes within, institutions. Table 2 contains a snapshot of the types of scores required by major Canadian institutions. The information in Table 2 comes directly from each university's graduate school website.

Table 2

Required Scores on IELTS Academic and TOEFL iBT for Graduate Studies at 6 English-Medium Canadian Universities

School	**Minimum score on IELTS Academic**		**Minimum score on TOEFL iBT**	
University of Toronto*	Overall: 7 Each component: 6.5		Overall: 93 Writing/speaking: 22	
York University	Fine Arts, Humanities and Social Sciences	7.5	Fine Arts, Humanities and Social Sciences	100
	Biology, Computer Science & Engineering, and Economics	7	Biology, Computer Science & Engineering, and Economics,	90-91
	All other science-based programs (except Biology and Computer Science & Engineering) and Mathematics & Statistics	6.5	All other science-based programs (except Biology and Computer Science & Engineering) and Mathematics & Statistics	79-80
University of British Columbia*	Overall: 6.5 Each component: 6		All programs	Overall: 90 Reading/listening: 22 Speaking/writing: 21
			Many programs	Overall 92-108
McGill	Overall: 6.5		Overall: 86 Each component: 20	

McMaster	Overall:6.5 Each component: 5.5	Most faculties Business Engineering	Overall: 92 Overall: 90 Reading/Listening:22 Speaking: 26 Writing: 24 Overall: 88
University of Alberta*	Overall: 6.5 Each component: 5.5	Overall:88 Each component: 20	

* Some programs have higher minimum scores, no master list available

The participants for our study were five international graduate (Master's or PhD level) students in a language education program at Metro University. All five of our participants were born in a country outside Canada, and also taught or worked internationally before moving to Canada. We chose to focus on international students because all of us had at one point in our own undergraduate or graduate studies been international students ourselves. Thus, we identified with the academic, linguistic, and cultural challenges that our participants faced.

All five of our participants had taken the IELTS or TOEFL more than once for the purpose of gaining admission into a Master's or PhD program at Metro University, and had all achieved at least the minimum required score on the IELTS and TOEFL before being offered a place in the program. As all our participants took the IELTS or TOEFL more than once, a limitation of our study was that we were not able to make comparisons between the experiences of graduate students who had to take high-stakes proficiency tests more than once before gaining admission into their program, and those who obtained the required test score on their first attempt. Table 3 below provides more information about the background of each participant and the tests taken:

Table 3

Information about the Participants in Our Study and the Tests Taken

Participant (Self-selected pseudonym)	**Country of Origin**	**Degree**	**Tests Taken**
Adam	China	Master of Education	IELTS (once), TOEFL (twice)
Andrea	Italy	Master of Arts	TOEFL (twice)
Jane	Thailand	PhD	IELTS (twice), TOEFL (once)

Julie	China	Master of Education	IELTS (twice)
Pedro	Colombia	PhD	TOEFL (three times)

Data Collection and Analysis

As part of a broader study on international students' experiences with navigating English language proficiency tests, we conducted three interviews with each participant using a life story and phenomenological approach (Seidman, 2013). Our phenomena of interest were the participants' experiences taking the language proficiency tests and their experiences in the first few months of their graduate program at Metro University, but we obtained this information by eliciting authentic stories of their lives both before and after becoming graduate students at the university.

We used a three-part interview structure (Seidman, 2013), and each interview lasted 1 to 1.5 hours long. In the first interview, we asked the participants about their childhood, educational background, and early language learning experiences. In the second interview, we asked them about their experiences preparing for the TOEFL or IELTS, and their experiences taking these tests. We did not ask our participants to give us any information about the questions in the tests, as test-takers are usually required to keep the content of the tests confidential. Rather, we asked them to reflect on the test format, and the skills and modalities that they were tested on. In the final interview, participants spoke about their experiences post-admission into Metro University. Specifically, we asked them about the challenges and successes they faced in their graduate program. Participants also talked about the match between the tests they took and the types of tasks and assignments they completed in their program, and reflected on whether or not taking the tests helped them with the challenges they faced in their program. Participants also described how they dealt with these challenges, and the types of support that helped them become successful in their program.

After completing all the interviews, we transcribed all of them using rough transcription, meaning our transcription did not focus on details such as prosody, pauses, intonation, etc. (Moore & Llompart, 2017). We carried out this transcription by listening to the interview data and writing out the verbal content of our participants' interview responses. The reason for using rough transcription was because our research questions required an analysis of the content rather than the linguistic features of our participants' responses.

In the first paper from this study (Sinclair, Larson & Rajendram, 2019), we present each participant's life story through narrative portraits, and discuss how their early language learning and educational experiences, and their academic and career aspirations intersect with their test-preparation and test-taking experiences. For this present study, we were interested in examining themes related to participants' perceptions of the tests, and the challenges and successes they faced post-admission into their graduate programs. Thus, in line with this focus, we conducted a thematic analysis of the data from the second and third interviews. We did not code the data from the interviews, but rather focused on

identifying, analysing and presenting pertinent patterns or themes in the data that related to each of our research questions (Braun & Clarke, 2006).

We used both deductive and inductive thematic analysis (Nowell, Norris, White & Moules, 2017) to identify the significant themes in our findings. This analysis occurred in several stages. Firstly, to begin our thematic analysis process, each of us read through the full transcripts of the participants, and highlighted specific interview responses that we believed to be pertinent to the focus of this particular study. The deductive part of this analysis involved identifying interview responses in the transcripts that related to each of the four components of the tests: listening, speaking, reading and writing. The inductive part of the analysis involved identifying emergent themes that related to the three research questions. In the next stage of the analysis, we gathered as a research team and presented these themes to each other. This allowed us to identify commonalities and differences in the responses across our five participants, and to develop broader themes based on these responses. In the third stage of the analysis, we went back to the transcripts to identify examples and quotes from each participant that would illustrate these broader themes. In keeping with the qualitative phenomenological approach used in this study, our aim in presenting these themes is not to make generalizable claims about the content of the tests, but rather to provide rich descriptions and excerpts from the interviews that highlight our participants' own perspectives and experiences.

Findings

Match between Tested Skills and the Language Demands of Graduate School

One of the common themes throughout our data was the mismatch between the tested language skills and the actual language demands in graduate school, as perceived by our participants. IELTS and TOEFL tests both contain four sections, each section focusing on a different English language skill. For clarity, we have organized our findings on this topic by test section: speaking, writing, listening, and reading.

Speaking. Our participants reported there to be a disconnect between speaking tasks, topics and delivery, and their graduate student experiences. Despite obtaining the scores she needed on the speaking section of the test, Julie felt that the IELTS speaking task types did not reflect the actual speaking she needed to do in her classes, such as giving oral presentations, debating topics, leading class discussions, and engaging in critiques of texts and ideas (Schmidt & Gannaway, 2007). Moreover, the tests did not require her to integrate her research or reading skills into her spoken discourse.

> The contents were really different from what we did in classes...For some courses…we read a passage, and then we tried to paraphrase the passage by organizing some main points and express our own ideas on the issue that the paper was talking about. And for some group presentations, we choose one perspective that we want to express with others, and we collect data from different accesses. But for the IELTS Speaking…we just try to talk, and that's it. There was nothing about reading or doing research. So, it's really different.

Jane found that the IELTS speaking tasks were irrelevant to the language demands of graduate school, and that the topics were sometimes esoteric and unfamiliar to her.

Moreover, she felt that the speaking section was not an accurate reflection of the type of speaking she needed to do in courses. The below quote illustrates her belief that knowing parts of speech is not the same as being able to contribute to class discussion, leading her to question what IELTS really measures:

> I don't know what the speaking part does to shape my grammar…if I know all the grammar, but I don't know what to say, still I cannot talk… OK, I know all the words. I know the grammar to use, but I don't really have a strong opinion about that so, I can't say anything. So, what does it measure? My accent? My pronunciation maybe.

Although our study did not focus on identifying differences between the IELTS and TOEFL tests, our findings suggest that students' test-taking experiences did differ with regard to the speaking component of the two tests. While IELTS speaking is done face-to-face, both the computer-based and internet-based versions of TOEFL speaking are monologic tasks wherein the test-taker has to speak to a computer screen. Andrea, who took the computer-based TOEFL speaking test, felt this mode of delivery was completely removed from the reality of graduate school face-to-face communication and expressed her frustration thusly: "I think it's ridiculous because you are talking to a screen[…] In university you talk with other people all the time."

Writing. Some of our participants felt that the length and content demands of language tests were not reflective of the writing they had to do as part of their coursework in graduate school. In the language education graduate program at Metro University, students are required to complete a variety of written assignments such as literature reviews, research briefs, reflective journals, research papers, article reviews, autoethnographies or duoethnographies, and case studies. Almost all of the participants described the mismatch between the essay and paragraph structures required by the tests and actual course demands. Jane described the contrasting advice given to her by her exam preparation writing teacher, and her professor at Metro University.

> When I prepared for the writing section on the exam, I took a writing course at a language school and the teacher just told me, "You just need to have five paragraphs. You know, introduction, three for body and conclusion", but once I got to school [university] the teacher [professor] say, Try to avoid writing five-paragraph essay. I was like, What?! That's what I learned.

Continuing her discussion about her language school experiences, Jane expressed that the skills she needed to successfully pass the writing test were not relevant to the writing demands of graduate school which required more analytical and critical writing. This resonates with Ravichandran et al.'s (2017) research on the writing challenges faced by international graduate students in the U.S. Julie mentioned the length of writing required by IELTS and explained how the writing prompts did not match the writing she had to do in her program.

Our study found that the level of graduate study had an impact on participants' experiences, particularly when it came to the type of writing they had to do in the program.

Jane and other participants in a thesis-stream program (PhD, Master of Arts) talked about the difficulties they faced with the research writing genre. Comparing the writing she had to do for the IELTS with the writing she now does for research papers, Jane stated:

> For the IELTS writing, it's just like three paragraphs...we don't need to analyze the data we got…we just express our thoughts, just express our ideas. But for the research papers, it's more like report, like what did you get... so I think that's not the same.

The only things related to the writing test that were useful to Jane were the transition words she memorized for the writing portion of TOEFL and the confidence passing the exam gave her. Pedro believed that the TOEFL required a specific structure and that the structure was more important to raters than the content and style of the writing. He believed the tests as "[…] are so depersonalized because they are not looking at YOU". In the following excerpt, Pedro pretends to be a writing test rater, echoing his beliefs about what they want.

> I want to see a structured paragraph, a structured piece with the introduction, with the body, and the conclusion…I'm not concerned about what you think or your opinion, or you're critical or not. I'm concerned about you give me the structure, and I'll give you the money, and the money is your points.

The above quote also illustrates Pedro's feelings about the test not properly measuring academic writing, but merely exchanging points for following the rules of structure and exchanging the points for money. Pedro's concerns are consistent with Kim's (2017) research which found there to be disparities in the writing skills tested in the TOEFL iBT and the academic writing skills that would be expected of students in their academic programs. Most of the students in Kim's study reported that they prepared for the TOEFL writing test by memorizing templates in the exam preparation material, rather than applying critical academic writing skills.

Listening. Listening did not emerge as a common topic amongst all our participants. Only Jane and Adam discussed this test modality and they had quite different opinions about it. Jane found a mismatch between the listening portions of the IELTS test and what was required for graduate school. Lectures in the listening were not relevant because in graduate school Jane mostly listened to classroom discussions.

> On the test, I listen to a lecture of a professor explaining something while in classes I hardly see that professors give a lecture. It's discussion all the time and I listen to other students…most of the time they don't sound like what I heard on the test.

Jane's comment highlights the difference between undergraduate and graduate classes, where the latter does not typically include many lectures. This highlights the need for listening tests for graduate students that are more relevant to the types of classes they attend. In contrast to Jane, Adam found that he could benefit from the listening test on the

TOEFL because the listening passages were similar to the types of conversations that happened in the North American university setting.

> The most helpful parts are listening, actually. All the content of the listening parts are campus-based. Conversation happen in campus or in the classroom, it's just two or three guys talking about some assignment or some issue happening in campus. That's almost what's happened later when I came to Metro University. That's kind of like, the knowledge I really can apply it in my future life.

Reading. Some of our participants found a mismatch between the reading passages included in the tests and the required readings in their graduate courses. Additionally, they noted that the skills they needed for the tests (e.g. skimming and scanning; reading quickly) were either unnecessary or only somewhat helpful when faced with actual course readings. Graduate students at Metro University are typically required to read academic journal articles and book chapters in preparation for their classes. Julie found that the topics of the reading passages selected in the test did not resemble the topics and types of texts she had to read during her graduate program.

> The reading passages in IELTS, they are all about, like nature, or people, or some social issues. But in Metro University we focused on study, education, so there was no connection between the readings and what we read in Metro University ...And also, all the readings we did in Metro University was really methodological things, or really academic. But for the IELTS readings, I think those kind of readings is more similar to like National Geographic, like the journal or magazine passages...So it's quite different actually.

For Andrea, the indirect style of the test reading passages was quite dissimilar from typical academic reading texts.

> The reading was okay, like it said, it was obvious that they were trying to trick you sometimes, where actual academic papers usually don't try to trick you, they try to make you understand right away what they're trying to say.

While Pedro felt that the reading component of the test helped him to develop skills such as skimming and scanning texts, he believed that there were many reading skills such as note-taking that he could “learn down the road” through time and experience in his program. Jane found that reading fast, which was a skill that she needed to develop for the IELTS and TOEFL, was not useful in graduate school as she needed to read slowly and look up the meaning of words while reading in order to understand the texts better.

> I can read quickly and just scan for keywords, but I feel like if I read too fast, I may misunderstand something and make wrong inference. So I try to be careful...When I see some word that I'm not sure of, I need to look up and that slow down my reading. So preparing for the reading section on the exam didn't really help me.

Rather than reading fast, skimming and scanning, participants felt that the types of reading skills they needed to develop to be successful in their graduate program were reflection, analysis and synthesis - skills that instructors in graduate-level courses typically require of students (Lin & Scherz, 2014).

Encountering New Pragmatic Norms

Our participants found that the mismatch between educational cultures at home and in Canada created challenges for their successful adaptation, especially in terms of classroom discourse. This mismatch, and accompanying concerns about how they may be evaluated by peers and their teacher, has been previously described by Cheng (2000), Lee (2009), and Liu and Littlewood (1997) as a reason why international students may initially struggle to fully participate.

Adam had mixed feelings about his graduate program, which he felt had "some flaws, especially to the international students." He detailed his concerns about international students' experiences, which he ascribed to differences in the cultures of the two countries' education systems:

> The education system here is quite different from what they have in China, because in China people are more…how do you say, not as active… more inactively receive knowledge from the instructor, not like here. Most of the time the professor works as a facilitator not a lecturer. It takes a bunch of time to get used to that.

Similarly, despite Pedro's success on the test, he described the challenge of familiarizing himself with the discourse and rhetoric in graduate school in Canada. He also faced challenges adjusting to the way that class discussions occurred in graduate school. Instead of taking turns in an orderly fashion like his classroom experiences in Colombia, the discourse of the Canadian graduate school classroom seemed unstructured and even rude:

> My challenge was getting used to the academic world and the academic world of North America...there is, like, a rhetoric, or like a discourse...the way you say things, the way you write, the way you participate in the classes, the words that you use are different from the world that I was living before...In my first class...people would interrupt to each other, somebody's speaking, and then it seems like the person is kind of finishing the thought, and somebody else jumps in, and then the person is finishing their thought, somebody else jumps in...I was like, shoot, that's so rude.

Julie also found it difficult to participate in class discussions and respond to questions during class. She noted that in China, students were only required to listen and take notes but did not need to produce language or engage their peers on academic topics. Despite getting a good score on the speaking test, Julie found it challenging to express her thoughts and explain her ideas during class discussions. She explained:

> In Metro University, the professors require us to respond to all the questions, or we'll have discussions. But in China, we only just sit there and listen to teachers, and just teachers speak, and we just keep quiet and take lot of notes. So it's really

> different… Language is a restriction to international students, because even if we want to express something, we will think a lot, like "How can I express my thoughts in a very understandable way?" Because we think that if we say something and others won't understand, we will feel very stressful…So I didn't really respond a lot in the class.

Similar to the challenges that Pedro and Julie faced, Sabet and Babaei's (2017) study found that the listening skills tested on the IELTS did not prepare students for the type of academic listening skills required in university, in terms of their pragmatic understanding, the integration of listening and speaking, information literacy, and the types of topics covered. Along with Pedro and Julie, Jane also found speaking in her classes at Metro University challenging due to the differences between home and Canadian education cultures. For Jane, this resulted in her feeling a loss of confidence in her speaking ability once she began her studies. Her perceived lack of proficiency was not related to all language domains, though; she felt it was primarily related to the speaking demands related to academic coursework:

> It's just at Metro University you know, and everybody is so great, I was like, 'Oh my God! No! I wish I was good again!' I don't know, it's so hard. You know, at some point, when you have to talk about academic stuff, that's when I get stuck. I can talk about the weather, I can talk about food, and everything, But all the academic term, it was like, 'Oh my God! This is so hard!'

But academic language proficiency was not the only barrier Jane faced. She also felt cultural differences about language use, compounded by the fact that she was studying language and literacy. One of these differences was the use of the term "native speaker." Tension arose in between herself and her thesis committee members over the appropriateness of the use of 'native speaker' in her research. Operating from a position of power, the committee disregarded Jane's use of this term in her research. Responding to this, Jane commented:

> They say native speaker model shouldn't be the ideal for language learners, shouldn't be something that learners should strive for. You should accept variation and focus on communicative aspect… In my study I used the native speaker as the benchmark and my thesis committee said, "No this is the 21st century, the native speaker model is no longer valid.

This was difficult for Jane, because in the education culture she was familiar with, the notion was still very relevant. In fact, she felt the desire to communicate like a native speaker was perfectly valid and that the committee was being unreasonable. She explained, "No! It's still valid here in Thailand! ...The main reason for us to want to go study abroad is to improve the English skills. And we want to be able to communicate like native speaker." That some international students prefer "native speaker" models of communication is a complex issue, influenced by multiple personal and cultural factors (Subtirelu, 2013).

How Students Navigated the Challenges of their Graduate Program

Consistent with research showing how the interplay of personal, social and academic factors such as individual initiative, motivation, peer support networks, and pre- and post-admission university academic supports contribute towards student success (Benzie, 2010; Feast, 2002; Holdsworth, Turner & Scott-Young, 2018; O'Loughlin, 2008; Zeegers & Barron, 2008), our findings suggest that students tried to navigate the challenges they faced in their graduate programs by seeking out available supports themselves.

University support. Metro University offers a variety of free English language support programs, ranging from formal (e.g., writing centres, general conversation courses, and peer support programs) and informal (e.g., conversation cafes, workshops, online learning modules). Most participants actively participated in one or more of these supports. However, their perceptions of the effectiveness of university supports varied substantially. Upon identifying the challenges he faced in his program, Pedro self-assessed his writing ability and determined that he needed to improve it. He actively sought out writing support from the university, which he felt was effective in helping him become a better academic writer:

> When I came to Metro University the first thing I noticed was oh, it's hard for me to write the way it has to be written, I have to write here. So I took those courses... English grammar, focusing on grammar, editing your own work, focusing on transitions… All of those courses have helped me to understand and learn better this academic writing.

On the other hand, Julie felt that the writing support offered by the university was not useful to her. She reflected that she felt she needed support with her grammar, to make her writing flow like "real English." The writing centre, on the other hand, only focused on content. However, even regarding content, Julie felt the feedback was too brief to be useful:

> I wanted to make my language look as natural as possible. But they didn't help us to correct any grammar errors like that. They only focus on the content… But that was not my focus. I just want to make my language, the flow of language, kind of natural, like real English. They just responded, like very short sentences, the feedback. So I think our focus was not the same, so the help they provided was not really useful.

Although there were many resources and supports available in the university, both Pedro and Adam felt that the information about these supports was not made easily accessible to students. They felt their program either lacked appropriate support systems, or poorly publicized the supports systems that did exist. Pedro explained that it required determination and effort on his part to find these supports, because the university does not go out of its way to help students locate them through their own means and channels:

> The university never tells you anything. They assume that you already pass [the test], you are here at the university, now you are on your own… So I was trying to be more independent, and in order to get that independence, to go out and see what

> support I would get from university… the university never reached out to you to say, "What do you need?" No, the resources are out there, and there are many resources… But it's up to you, up to the individual to go and find out those resources.

Adam also felt that he was not aware of supports that existed until too late. Unfortunately, he revealed that he never had high expectations about the supports his program would offer, so he accepted his situation as somewhat inevitable:

> I should get some help from the [writing support program] but I didn't know that until the program ended. I should get a notice before that, so we barely know that…I never expect too much from the program. I know it's just a 1.5 years program.

Peer and external support. Since several participants did not receive the language or academic support they needed from the university, peer support was a key factor in helping them navigate their challenges. Pedro found that his peers were an invaluable source of support in helping him overcome the specific problems he faced with academic writing. He received constructive feedback on his writing from his peers, which helped him tremendously. Like Pedro, Julie had a strong network of support which helped her develop her writing skills. She received language support from her personal contacts outside the university, such as her Bible study instructor. Her network was also composed of other graduate students from China, and sharing a common culture and language helped them tackle the difficulties they faced in their assignments:

> We have a very close community… we can share a lot of things when we do the assignments…like maybe I didn't get the point of the requirements, but my friends, they do, so they will tell me… I got a lot of peer assessment… After we write our papers, our friends will try to check it out, to see if you have any good points or something that you don't have to say in the paper.

Discussion & Recommendations

Our findings raise several concerns regarding the use of English language proficiency tests in university admissions policies, specifically, with regard to the alignment between the content of the tests, the demands of higher education, and the supports offered by the university. In documents explaining the English language proficiency testing requirements of Metro University, a high level of English proficiency, as evidenced by high test scores, is equated with the capacity to be successful in the university's degree programs. In making the assumption that high scores on the language proficiency tests will ensure that international students can function successfully at the university level (Benzie, 2010), universities are also assuming that passing a minimum threshold on a standardized language proficiency test is the most important determining factor of international student success. However, our study suggests that international students need to develop many other skills to have enriching and rewarding experiences in their program of study. However, all our participants continued to face challenges in all four language modalities despite successfully achieving the scores required for admission into Metro University. Their experiences in their respective programs suggest that the skills

assessed by tests such as the IELTS and TOEFL are not entirely aligned with the actual challenges of graduate school, because the elicited skills are not the same as the skills that students will need to acquire to be successful in the target language domains of higher education settings. Our findings suggest that preparing for and taking the tests does not necessarily help international students to develop other essential skills such as using rhetoric, understanding discipline-specific discourse, and participating in classroom discussions. Although certain tested skills were transferable to the target domain, this varied from among participants and was dependent on task type. One limitation of our study was that all of our participants were from the same department in the same university. Therefore, it was beyond the scope of our study to analyze the academic language requirements of other graduate programs at Metro University, or to determine if there were differences between international students' test-taking and learning experiences in undergraduate and graduate programs. A future area for study would be to compare the specific skills required in various undergraduate and graduate courses with the skills tested in the IELTS and TOEFL. Due to our small sample size, we were also not able to conduct a systematic analysis to determine if there were differences between participants who took the IELTS and TOEFL. Thus, future studies in this area could compare the learning experiences of international students who take a variety of tests for admission into English-medium universities. Further research could focus on determining if there are differences between these tests in terms of the skills they test and their applicability for graduate education.

Our findings revealed that the experiences, needs, and challenges faced by the five students varied according to their learning styles, previous educational experiences, and availability of peer and external support. In another paper based on this study (Sinclair, Larson & Rajendram, 2019), we discuss how their educational background, early language learning experiences, cultural and linguistic background, career trajectory, social networks, and lived experiences played a role in their negotiation of the tests, and navigation of the challenges in their graduate programs of study. Although all our participants were successful in their graduate programs, university admissions policies must not assume that all students who meet the minimum language requirements, for example by achieving a band 6.5 on the IELTS test, are equally prepared to overcome their challenges and achieve success on their own because not all of them have the same personal resilience or access to the same types of support. The various support programs offered by the university to students, for example the writing centre, may not be appropriately matched to the needs of international students. Our participants suggested that these services were not always advertised consistently, or offered at the appropriate time when they were most needed. In addition, the types of services offered may not be relevant to the students' needs and may not be tailored to their personal experiences in the program. To better help international students with the challenges they face, we recommend that instead of making the English language requirements for admission more stringent, universities should focus on providing customized supports and resources to develop the skills that the language proficiency tests may not encompass (Fass-Holmes & Vaughn, 2015). As Slethaug (2015) and van der Walt (2013) suggest, universities should also provide supports that take into account the home cultures of international students, and develop linguistically and culturally responsive and inclusive curricula to affirm the home languages and cultures of their international students. Research on international student success has shown that the

various supports provided by the university are essential in ensuring the academic, cultural and social adjustments of international students, and in improving their emotional and psychological outcomes (Cho & Yu, 2015). In addition to providing supports, universities should have a policy to ensure that students who need these supports will know about them and have access to them throughout their program.

Our study also offers important implications for university language requirements and admission criteria. Since different academic programs and contexts will require different types of language use (e.g., discipline-specific vocabulary, formal vs. informal registers), universities should implement more holistic measures of language proficiency, for example by requiring portfolios and additional interviews to understand an individual's experiences using English in various contexts. Considering the fact that international students applying to graduate school have already successfully completed an undergraduate discipline, universities should also use more contextualized measures that account for students' discipline-specific academic language skills, and capture the full repertoire of their language practices. Pedro suggested universities use a variety of measures (what he calls 'packages') for admission, which he describes this way:

> I think universities or government should look for these packages in terms of […] your background […] language. Demonstrate me with letters of recommendation, with things that you've done, and with an interview, it can be a Skype interview or a personal interview, so I can assess your ability, right? And with a writing sample.

A test of English for discipline-specific purposes (e.g. English for visual arts students, English for engineering students) would also provide necessary contextualization to an applicant's ability to use language in specific academic contexts. Although tests other than IELTS and TOEFL are accommodated in most major universities, there is a tendency to compare applicants' scores on those tests against IELTS or TOEFL descriptors, implying that scores on the IELTS or TOEFL are the best measure of proficiency (Dunworth, 2010). Dunworth stresses that universities should be flexible enough not to rely only on the IELTS and TOEFL while ensuring that they have academically sound criteria for the language proficiency measures they accept, and a good understanding of what their language measures actually mean in practice for their students. One way to make the language assessment practices of universities more democratic (Shohamy, 2001a, 2001b) is to involve test users (students taking the tests and university policy-makers) in test creation, design, and in the ways scores are used. Higher education policy-makers should work together with test-creators and test-takers to determine how test scores are used, to develop more relevant test content and holistic language proficiency requirements, and to ensure that all international students receive the language, academic and pragmatic supports they need to be successful in their programs.

References

Ambrósio, S., Marques, J. F., Santos, L., & Doutor, C. (2017). Higher education institutions and international students' hindrances: A case of students from the African Portuguese-speaking countries at two European Portuguese universities. *Journal of International Students, 7*(2), 367-394.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Andrade, M. (2006). International students in English-speaking universities: Adjustment factors. *Journal of Research in International Education, 5*(2), 131-154.

Andrade, M. S. (2009). The effects of English language proficiency on adjustment to university life. *International Multilingual Research Journal, 3*, 16-34.

Aryadoust, V. (2013). *Building a validity argument for a listening test of academic proficiency*. Cambridge: Scholars Publishing.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: University Press.

Bachman, L. F., & Purpura, J. E. (2008). Language assessments: Gate-keepers or door openers? In B. Spolsky & F. Hult (Eds.). *The handbook of educational linguistics* (pp. 456-468). Malden, MA: Blackwell Publishers.

Beaven, M., Calderisi, M., & Tantral, P. (1998). *Barriers to learning experienced by Asian students in American accounting classes.* Paper presented at the American Accounting Association MidAtlantic Regional Meeting, March 26–28.

Benzie, H. J. (2010). Graduating as a 'native speaker': International students and English language proficiency in higher education. *Higher Education Research & Development, 29*(4), 447-459.

Berman, R., & Cheng, L. (2001). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics, 4*(1), 25–40.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brown, L. (2008). The incidence of study-related stress in international students in the initial stage of the international sojourn. *Journal of Studies in International Education, 12*, 5-28.

Burnett, C., & Gardner, J. (2006). The one less travelled by…: The experience of Chinese students in a UK university. In M. Byram & A. Feng (Eds.), *Living and studying abroad: Research and practice* (pp. 64-90). North York, ON: Multilingual Matters Ltd.

Chapelle, C. A. (1999). Validity in language assessment. *Annual Review of Applied Linguistics, 19*, 254-272.

Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (2008). *Building a validity argument for the Test of English as a Foreign Language*. New York: Routledge.

Cheng, X. (2000). Asian students' reticence revisited. *System, 28*(3), 435-446.

Cho, J., & Yu, H. (2015). Roles of university support for international students in the United States: Analysis of a systematic model of university identification, university support, and psychological well-being. *Journal of Studies in International Education, 19(*1), 11-27.

Cotton, F., & Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a group of international students studying at the University of Tasmania. In S. Wood (Ed.), *International English Language Testing System (IELTS)*

Research Reports 1998: Volume 1 (p. 72). Retrieved from https://www.ielts.org/-/media/research-reports/ielts_rr_volume01_report4.ashx

Dunworth, K. (2010). Addressing the issue of English language proficiency in Australian universities. *Australian Universities' Review, 52*(2), 5-10.

Educational Testing Service. (2019). About the TOEFL iBT test. Retrieved from https://www.ets.org/toefl/ibt/about

Fass-Holmes, B., & Vaughn, A. A. (2015). Evidence that international undergraduates can succeed academically despite struggling with English. *Journal of International Students, 5*(3), 228-243.

Feast, V. (2002). The impact of IELTS scores on performance at university. *International Education Journal, 3*(4), 70-85.

Friedenberg, J. (2002). The linguistic inaccessibility of US higher education and the inherent inequity of US IEPs: An argument for multilingual higher education. *Bilingual Research Journal, 26*(2), 309–328.

Hellsten, M., & Prescott, A. (2004). Learning at university: The international student experience. *International Education Journal, 5*, 344-351

Hill, K., Storch, N., & Lynch, B. (1999). A comparison of IELTS and TOEFL as predictors of academic success. In R. Tulloh (Ed.), *International English Language Testing System (IELTS) Research Reports 1999, Volumn 2* (p. 62). Retrieved from https://www.ielts.org/-/media/research-reports/ielts_rr_volume02_report3.ashx

Holdsworth, S., Turner, M., & Scott-Young, C. M. (2018). Not drowning, waving. Resilience and university: A student perspective. *Studies in Higher Education, 43*(11), 1837-1853.

Holmes, P. (2004). Negotiating differences in learning and intercultural communication: Ethnic Chinese students in a New Zealand university. *Business Communication Quarterly, 67*, 294–307.

IDP Education Canada. (2019). *Choice of academic vs. general training*. Retrieved from https://ieltscanadatest.com/take-ielts/choice-of-two-tests/

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*(3), 527-535.

Khan, S. Z. (2009). Imperialism of international tests: An EIL perspective. In F. Sharifian (Ed.), *English as an International Language: Perspectives and pedagogical issues* (pp. 190-208). North York, ON: Multilingual Matters.

Kim, E. Y. J. (2017). The TOEFL iBT writing: Korean students' perceptions of the TOEFL iBT writing test. *Assessing Writing, 33*, 1-11.

Kodama, Y. (2007). *Japanese students' contact with native English speakers during a study abroad experience* (Unpublished Master's dissertation). University of Alberta, Canada.

Lee, G. (2009). Speaking up: Six Korean students' oral participation in class discussions in US graduate seminars. *English for Specific Purposes, 28*(3), 142-156.

Lin, S. Y., & Scherz, S. D. (2014). Challenges facing Asian international graduate students in the US: Pedagogical considerations in higher education. *Journal of International Students, 4*(1), 16-33.

Lin, J. G., & Yi, J. K. (1997). Asian international students' adjustment: Issues and program suggestions. *College Student Journal, 31*(4), 473–479.

Liu, L. (2011). An international graduate student's ESL learning experience beyond the classroom. *TESL Canada Journal, 29*(1), 77-92.
Liu, N. F., & Littlewood, W. (1997). Why do many students appear reluctant to participate in classroom learning discourse? *System, 25*(3), 371-384.
Marr, T. (2005). Language and the capital: A case study of English "Language shock" among Chinese students in London. *Language Awareness, 14,* 239-253.
McGill University Graduate and Postdoctoral Studies. (2019). *English language proficiency*. Retrieved from https://www.mcgill.ca/gradapplicants/international/apply/proficiency
McMaster University. (2019). *Academic Calendars: General regulations of the graduate school 2.1.12 English language requirements.* Retrieved from https://academiccalendars.romcmaster.ca/content.php?catoid=25&navoid=4667#2-1-12_english_language_requirements
Messick, S. (1989). Validity. In. R. Linn (Ed.). *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan Publishing.
Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. *ETS Research Report Series, 2003*(1), i-29.
Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20.
Moore, E., & Llompart, J. (2017). Collecting, transcribing, analyzing and presenting plurilingual interactional data. In E. Moore & M. Dooly (Eds.), *Qualitative approaches to research on plurilingual education* (pp. 403-417). Dublin, Ireland: Research Publishing Net.
Mori, S. (2000). Addressing the mental health concerns of international students. *Journal of Counseling & Development, 78,* 137-144.
Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*. https://doi.org/10.1177/1609406917733847
O'Loughlin, K. (2008). The use of IELTS for university selection in Australia: A case study. In J. Osborne (Ed.), *IELTS Research Reports, Vol. 8* (pp. 145–241). Canberra, Australia: IELTS Australia.
O'Loughlin, K. (2011). The interpretation and use of proficiency test scores in university selection: How valid and ethical are they? *Language Assessment Quarterly, 8*(2). https://doi.org/10.1080/15434303.2011.564698
Ravichandran, S., Kretovics, M., Kirby, K., & Ghosh, A. (2017). Strategies to address English language writing challenges faced by international graduate students in the US. *Journal of International Students, 7*(3), 764-785.
Riazi, A. M. (2016). Comparing writing performance in TOEFL-iBT and academic assignments: An exploration of textual features. *Assessing Writing, 28,* 15-27.
Rienties, B., Beausaert, S., Grohnert, T., Niemantsverdriet, S., & Kommers, P. (2012). Understanding academic performance of international students: The role of ethnicity, academic and social integration. *Higher Education, 63,* 685-700.
Roberston, M., Line, M., Jones. S., & Thomas, S. (2000). International students, learning environments and perceptions: A case study using the Delphi technique. *Higher Education Research and Development, 19,* 89-102.
Ryan, Y., & Zuber-Skerritt, O. (1999). Supervising non-English speaking background

students in the globalized university. In Y. Ryan & O. Zuber-Skerritt (Eds.), *Supervising postgraduates from non-English speaking backgrounds* (pp. 3-11). Buckingham, UK: SRHE & Open University Press.

Sabet, M. K., & Babaei, H. R. (2017). On the relationship between the IELTS listening and listening in academic English programs. *Advances in Language and Literary Studies, 8*(2), 170-179.

Sawir, E., Marginson, S., Forbes-Mewett, H., Nyland, C., & Ramia, G. (2012). International student security and English language proficiency. *Journal of Studies in International Education, 16*(5), 434-454.

Schmidt, L., & Gannaway, D. (2007). E-Learning to aid international student transition. *International Journal of Learning, 13*(12), 55-60.

Schoepp, K. (2018). Predictive validity of the IELTS in an English as a medium of instruction environment. *Higher Education Quarterly, 72*, 271-285.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences.* (4th ed.). New York, NY: Teachers College Columbia University.

Shepard, L. A. (1993). Chapter 9: Evaluating test validity. *Review of Research in Education, 19*(1), 405-450.

Shepard, L. A. (2016). Evaluating test validity: Reprise and progress. *Assessment in Education: Principles, Policy & Practice, 23*(2), 268-280.

Shohamy, E. (2001a). Democratic assessment as an alternative. *Language Testing, 18*(4), 373–391.

Shohamy, E. (2001b). *The power of tests: The impact of language tests on teaching and learning*. Harlow, UK: Longman.

Shohamy, E. (2017). Critical language testing. *Language testing and assessment*, 1-15.

Sinclair, J., Larson, E. J., & Rajendram, S. (2019). "Be a machine": International graduate students' narratives around high-stakes English tests. *Language Assessment Quarterly, 16*(2), 236-252. doi:10.1080/15434303.2019.1628238

Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research, 45*(1-3), 83-117.

Slethaug, G. E. (2015). Conclusion: International teaching and learning at universities – Achieving equilibrium with local culture and pedagogy. In G. E. Slethaug & J. Vinther (Eds.), *International teaching and learning at universities: Achieving equilibrium with local culture and pedagogy* (pp. 159-172). New York: Palgrave Macmillan.

Spolsky, B. (2008). Language assessment in historical and future perspective. In N. Hornberger (Ed.), *Encyclopedia of language and education* (pp. 2570-2579). Boston, MA: Springer

Spurling, N. (2007). Exploring adjustment: The social situation of Chinese students in UK higher education. *Learning and Teaching in the Social Sciences, 3*(2), 95-117.

Storch, N., & K. Hill. (2008). What happens to international students' English after one semester at university? *Australian Review of Applied Linguistics, 31*, 4.1-4.17.

Trice, A. G. (2003). Faculty perceptions of graduate international students: The benefits And challenges. *Journal of Studies in International Education, 7*, 379-403.

Subtirelu, N. (2013). What (do) learners want (?): a re-examination of the issue of

learner preferences regarding the use of 'native'speaker norms in English language teaching. *Language Awareness, 22*(3), 270-291.
The University of British Columbia Graduate School. (n.d.). *English proficiency requirements.* Retrieved from https://www.grad.ubc.ca/prospective-students/application-admission/english-proficiency-requirements
Toulmin, S. E. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.
Trice, A. G. (2003). Faculty perceptions of graduate international students: The benefits and challenges. *Journal of Studies in International Education, 7*, 379-403.
UNESCO. (2018). *Global flow of tertiary-level students*. Retrieved from http://uis.unesco.org/en/uis-student-flow
University of Alberta Faculty of Graduate Studies and Research. (2019). *Graduate program manual: 5.8 English language proficiency requirement.* Retrieved from https://www.ualberta.ca/graduate-studies/about/graduate-program-manual/section-5-admissions/5-8-english-language-proficiency-requirement
University of Toronto School of Graduate Studies (2013). *English-language proficiency testing*. Retrieved from http://www.sgs.utoronto.ca/prospectivestudents/Pages/English-Language-Proficiency-Testing.aspx
Van der Walt, C. (2013). *Multilingual higher education: Beyond English medium orientations*. North York, ON: Multilingual Matters.
Wong, J. K. (2004). Are the learning styles of Asian internationals culturally or contextually based? *International Education Journal, 4*, 154-166.
York University Faculty of Graduate Studies. (n.d.). *Academic requirements.* Retrieved from https://gradstudies.yorku.ca/current-students/regulations/admission/#english
Zeegers, M., & Barron, D. (2008). Discourses of deficit in higher degree research supervisory pedagogies for international students. *Pedagogies: An International Journal, 3*, 69-84.
Zhang, Z. (2011). A nested model of academic writing approaches: Chinese international graduate students' views of English academic writing. *Language and Literacy, 13*(1), 39-59.
Zhang, Y. (2016). International students in transition: Voices of Chinese doctoral students in a U.S. research university. *Journal of International Students, 6*(1), 175-194.
Zhang, Z., & Beck, K. (2014). I came, but I'm lost: Learning stories of three Chinese international students in Canada. *Canadian and International Education, 43*(2), 1-14.
Zhou, Y., Frey, C., & Bang, H. (2011). Understanding of International graduate students' academic adaptation to a U.S. graduate school. *International Education, 41*(1), 76-94.

Author Biography

Dr. Shakina Rajendram is a lecturer in the Department of Curriculum, Teaching & Learning at the Ontario Institute for Studies in Education (OISE), University of Toronto. She completed her PhD in Language and Literacies Education/Comparative, International

and Development Education at OISE, and her research focuses on teacher education, translanguaging and plurilingual pedagogies, collaborative learning, and multiliteracies.

Elizabeth Jean Larson is a PhD student in the Department of Curriculum, Teaching, and Learning at the Ontario Institute for Studies in Education, University of Toronto. Her research focuses on oral proficiency, intercultural communication, critical literacy, and self-assessment measures.

Dr. Jeanne Sinclair is a faculty member in Teacher Education at the University of the Incarnate Word in San Antonio, Texas. Her research focuses on assessment validity, primarily in the areas of language and literacy.