August 2016

Testing Tensions: The Use of English Language Proficiency Tests for the Admission of Ontario High School Applicants to One Ontario University

Samia Lahib
The University of Western Ontario

Supervisor
Dr. Pam Bishop
The University of Western Ontario

Graduate Program in Education

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Education

Follow this and additional works at: https://ir.lib.uwo.ca/etd

⬤ Part of the Educational Assessment, Evaluation, and Research Commons, Educational Leadership Commons, and the Higher Education Administration Commons

Abstract

In today's globalized world, students have the opportunity to apply to universities around the world. At most universities where the dominant language of instruction is English, applicants whose first language is not English have to prove their English language proficiency by submitting test scores such as the internet-based Test of English as a Foreign Language (TOEFL iBT) or the International English Language Testing System (IELTS) as part of the application process. This study investigated the ability of TOEFL iBT and IELTS to identify applicants with sufficient English language proficiency to succeed academically during undergraduate study at one Ontario university between 2012 and 2014. A specific focus was placed on applicants who were studying the Ontario high school curriculum at the time of their application.

With appropriate institutional approval, this exploratory quantitative study probed University-owned data and conducted correlation and multiple regression analyses to determine the relationships between a sample of participants' ($n = 200$) TOEFL iBT or IELTS test scores with their academic success (as measured by first-year grade point average). Similar analyses were conducted between each participants' ($n = 2,810$) TOEFL iBT or IELTS test score with their grade 12 university level English (ENG4U) grade to determine if ENG4U grade could be used as a marker for identifying potentially fraudulent test scores.

The results of this study did not support the use of TOEFL iBT or IELTS as being the sole determining factor when identifying applicants with sufficient English language skills to succeed academically, nor did they provide statistically reliable support for increasing the current minimum test score required for admission. Although significant, the weak zero-order and partial correlations representing the relationship between TOEFL iBT or IELTS and ENG4U grade raised questions about the potential use of ENG4U grade as a marker for identifying fraud in

practice. Based on those findings, it is recommended that diagnostic assessments be used in combination with existing standardized tests, such as TOEFL iBT and IELTS, to aid in the academic success of future students. Although the results of this study may help inform this Ontario university's policies and procedures, further studies that employ qualitative designs may be useful by providing a rich understanding of first-year academic experiences of students who were required to submit proof of English language proficiency.

Acknowledgments

My appreciation and thanks for the accomplishment of this study are directed to Dr. Pam Bishop (Supervisor) for her patience and guidance, and to Dr. Perry Klein for his crucial insight. Without your contributions this would not have been possible. Throughout this process I have had the unconditional support of my family whom I am eternally grateful to. To my son Malik, my desire to make you proud has driven me to what, at times, seemed like an impossible challenge. I am here now because you inspire me. A special thank you goes to my husband who spent many long nights encouraging me to strive and reach my goals. Without his ongoing emotional and intellectual support I would not have been able to get to where I am today.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

The concept of a global economy is not one that is isolated to finance, politics or technology. In fact, in today's globalized world, knowledge is a commodity of intellectual property that is widely used and widely available. In order to attract prospective students to Canadian universities and propel Canada forward in an increasingly competitive globalized world, Canada must find ways to attract 'the best and brightest' minds to study, live and work here by recruiting talent in the areas of scientific and research development, economic development opportunities, and cultural diversification (Government of Canada, 2012, p. 1). Universities have adopted the challenge of diversifying their intellectual and social cultures by recruiting international students who potentially bring a wide array of intellectual capacities, knowledge, and cultural experiences. The inclusion of those international students is intended to add a global dimension to campus life and enrich the knowledge available within the intellectual culture of a university. Indirectly, exposure to more diverse populations builds a more 'culturally-aware' campus with culturally accepting students—a disposition which is usually necessary to succeed in today's global workplace.

Although it may be argued that under the pressures of the global marketplace universities have become less self-governed and autonomous and have moved toward a consumer-based framework of administration (Blackmore, Brennan & Zipin, 2010), the goal to internationalize education extends beyond the desire to increase possible financial means not previously available to many universities (Darling-Hammond, 1998). Public universities within Canada provide the opportunity for increasing numbers of citizens to pursue higher education. It is those universities

which diversify their intellectual and social cultures that are able to educate students to, for example, be flexible and adapt in a fast-changing world. As emphasized by Darling-Hammond (1998) the inability of citizens to adapt will result in a society divided by the access to knowledge and the ability to learn. In order to graduate students who demonstrate these life-long capacities universities must be able to expand the critical thinking and global understanding of the student population through the inclusion of intercultural and global perspectives; all of which are possible through the internationalization of education (Darling-Hammond, 1998).

The desire to have international students study in Canada is one shared by institutions and prospective students. The Canadian Bureau of International Education's (CBIE) 2009 survey "Canada First" evaluated the experiences of 5,925 international students from 22 universities and 4 colleges across Canada (p.1). Although not all students shared the same reasons for selecting a Canadian institution, over half of the students (53%) included in this study identified Canada as their first choice in which to study abroad (CBIE, 2009, p. 1). In fact, almost all university students (95%) indicated that they viewed Canada as "a place to reach their educational potential" and almost three quarters considered the quality of a Canadian education to be the main factor in making their decision (CBIE, 2009, p.1). Specifically, this survey identified Ontario as consistently attracting the largest number of international students in comparison to all other Canadian provinces and territories (Government of Canada, 2012, p. 18). The desire of those students to study in Canada is evident by the 78% increase in international undergraduate and graduate student enrolment at a Canadian university from 92, 727 students in 2005 to 116,890 students in 2010 (Government of Canada, 2012, p. 16). Specifically, international students studying at the university level in Canada accounted for approximately 53% of all international students in Canada in 2010 (Government of Canada, 2012, p. 16).

Although a majority of international students in Canada are choosing university level education, 1 in 3 international university students are first choosing to study at a Canadian or Canadian-curriculum based school before applying to university (CBIE, 2009, p. 1). For example, students wishing to apply to Ontario Universities may study at a high school located in Ontario which teaches the Ontario curriculum, or an Ontario curriculum-based school aboard. Having studied the provincial curriculum of the university they wish to apply to subsequently allow applicants to be reviewed based on their provincial requirements. Specifically, applicants currently studying in an Ontario high school (or the Ontario curriculum abroad) are eligible to apply to Ontario universities using the Ontario Universities' Applicant Center (OUAC) 101 application. Although OUAC 101 applications are assessed for admission based on the Ontario curriculum requirements, prospective students submitting this application may still be required to submit additional supporting documents at the request of the university. One common document required by Ontario universities is proof of English language proficiency for those applicants whose first language is not English. Some acceptable English language proficiency tests include the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), the Pearson Test of English Academic, the Michigan English Language Assessment Battery (MELAB), the CanTEST, or the Canadian Academic English Language Assessment (CAEL). Although OUAC 101 applicants study the Ontario curriculum, they may still be required to submit such proof if English is not their first language.

The Ontario university used within this study is a Tier 1 Canadian research intensive university which assess the applications of international studies and currently has upwards of 3,700 international students from 117 countries (University X, as in References, 2015, p. 6). This university has made it their Mission and Vision to attract the world's brightest minds while

maintaining the best student experience at a Canadian research intensive university. First-year

undergraduate cohorts arriving at this university are admitted with grades well above the Ontario

average for undergraduate admission, with this university requiring amongst the highest

admission averages in Canada (University X, as in References, 2014). It is important to

universities to continue to attract the best and brightest students in order to facilitate a highly

academic intellectual culture, while attempting to achieve high retention rates, all of which

promote the university in a positive light; and in turn hopefully attract more talented students.

Despite the high admission requirements set by this Ontario university as a Tier 1 institution, this

university is able to maintain an astounding first to second year retention rate of 93% (University

X, as in References, 2015, p. 6).

To ensure high retention rates, one strategy of universities is to find ways to choose

students who will likely succeed academically, and once enrolled contribute to the intellectual

and social cultures of the institution. For most universities, an admission average is used in one

way or another to loosely determine an applicant's potential to succeed academically once

enrolled as a student. Students whose first language is not English may be required as part of

their application to provide addition documentation which support their ability to communicate

effectively in an English-speaking academic institution. For most students, an English language

proficiency test may be required as such proof. It is important to public universities in Canada

that those tests are able to broadly, but effectively, predict applicants' potential to use the English

language sufficiently in order to succeed academically during their time in an undergraduate

program.

To understand the ability of English language proficiency tests to predict the future

academic success of an applicant, predictive validity studies are employed which determine the

relationship between those variables. The majority of current literature which measure the predictive validity of English language proficiency tests on academic success study international undergraduate or graduate school students as the sample. As previously mentioned, some students may choose to study at a Canadian or Canadian curriculum-based high school prior to attending university. In the case of the Ontario university used within this study, some OUAC 101 applicants may have recently arrived in Canada to study the Ontario high school curriculum or may currently be studying the Ontario high school curriculum abroad at an offshore school. For those students whose first language is not English and have studied for fewer than three years in an Ontario high school, the Ontario university used within this study requires that applicants submit an English language proficiency test, such as TOEFL iBT or IELTS, as part of their application for undergraduate admissions.

It is important that the distinction be made between those applicants required to submit proof of English language proficiency and being an international student. The term "international student" or "international applicant" is typically associated with the applicant's status in Canada as being other than permanent resident or Canadian citizen. In the event that an applicant holds either of those statuses, they are viewed as a "domestic applicant". As the requirement to submit proof of English language proficiency is based on the applicant's first language and the number of years they have spent in an Ontario high school, there is the potential for both domestic and international students to be required to submit such proof. For example, a Canadian citizen, whose first language is not English, and who lived outside of Canada before moving to Ontario for one year may be required, as a domestic applicant, to submit proof of English language proficiency. As demonstrated by this example, it is important to study the relationship between English language proficiency and the academic success of all applicants required to submit such

proof and not just those applicants classified as international. Although a review of the current

literature provides sufficient predictive validity studies on English language proficiency and the

academic success of international undergraduate and graduate school students, scant research

exists which depicts the predictive validity of TOEFL iBT or IELTS in relation to the academic

success of OUAC 101 applicants. Specifically, this literature neglects OUAC 101 applicants

who are from families or nations where the dominant language is other than English and who

have been required to submit proof of English language proficiency as part of their undergraduate

application to an Ontario university.

Understanding the authenticity of English language proficiency tests may help

universities develop their policies and procedures to utilize the strength of those tests while also

developing practices to help account for any limitations or weaknesses in the tests. One notable

weakness of English language proficiency tests, like many other standardized tests or documents

required for admission, is the potential for fraud. As no university is immune to credential or

document fraud, and all universities must do their due diligence to identify fraud whenever

possible. Although commonly recognized as a growing problem among the academic

community and in the media, academic literature in support of such fraud is few and far between.

With the limited amount of research available, it is difficult for institutions to develop methods to

identify the elusive issue that is English language proficiency test fraud. To better understand

English language proficiency test fraud, this exploratory study first aims to understand the

relationship between English language proficiency test scores and academic factors such as

academic success and prior English ability.

Nearly three decades ago, Graham's (1987) analysis of English language proficiency and

academic success emphasized that "while the research clearly shows that many factors other than

English proficiency are important to academic success, there may be for each university, or even for each program within an institution, a minimum level below which lack of sufficiency proficiency in English contributes significantly to lack of academic success" (p. 1). As a result, Graham (1987) recommended that each institution complete an analysis of all policies and procedures in order to determine the impact of English language proficiency on academic success.

For that initial reason, and the fact that the value or potential usefulness of English language proficiency tests is still not entirely clear, this study aimed to review the effect of English language proficiency test score on the academic success of OUAC 101 applicants whose first language is not English and who had studied for fewer than three years in an Ontario high school. It was also an aim of this study to determine whether Grade 12 University level English (ENG4U) grade, which is a required prerequisite for all OUAC 101 applicants to the Ontario university used in this study, could be used as a marker that when compared with TOEFL iBT or IELTS could help identify English language proficiency test scores which may not be representative of the applicant's abilities. In such cases, a test score may require further investigation to determine its authenticity and may subsequently be investigated as potentially fraudulent.

1.1   Statement of Problem

To identify applicants whose first language is not English but who possess the English language proficiency skills required to succeed within an academic English-speaking university, the university used within this study, and many other universities around the world are required to utilize standardized English language proficiency tests such as TOEFL iBT and IELTS. Although the current literature regarding TOEFL iBT or IELTS and academic success show

mixed results in support of and against the use of such test, additional resources to identify such potentially successful applicants are limited. Prior to making decisions regarding the usefulness of TOEFL iBT or IELTS, this exploratory study aimed to provide university administrators with a quantitative understanding of the predictive validity of such tests as well as a method for identifying potentially fraudulent test scores.

1.1.1   Predictive Validity

As universities attempt to internationalize their education system, the number of applicants who are required to submit proof of English language proficiency also increases. As previously mentioned, current predictive validity studies relating TOEFL iBT or IELTS to the academic success of international undergraduate or graduate school students show mixed results both in support of and against the reliability of such tests (Al-Musawi & Al-Ansari, 1999; Arcuino, 2013; Cho & Bridgeman, 2012; Cotton & Conrow, 1998; Dooey & Oliver, 2002; Feast, 2002; Fu, 2012; Ng, 2007; Simner, 1998; Wait & Gressel, 2009; Woodrow, 2006; Yen & Kuzma, 2009). Those varying results are concerning to test users, such as university admissions offices, as limited resources are available to identify applicants with sufficient English language proficiency to succeed once enrolled as a student other than standardized English language proficiency tests. Subsequently, the lack of resources available has made it difficult to instil ethical and objective practices for reviewing applications.

Those results are also concerning for senior administrators and educational leaders who are required to set the minimum English language proficiency test score at which applicants are viewed as having sufficient English language skills to succeed in an English-speaking academic institution. In the event that the minimum score set by universities is too high potentially successful applicants may be found ineligible based on their test score alone. Conversely, a low

minimum score could result in applicants being selected for admission who do not have the language skills to succeed and who may subsequently be required to withdraw from the university. As previously mentioned, maintaining high retention rates is critical to the ongoing and future success of the university and the intellectual culture that exists within it. Based on the fact that the value or potential usefulness of TOEFL iBT or IELTS is still not entirely clear and that the optimal minimum test score required for admission is not known, it is important that all public universities understand how much they can rely on English language proficiency tests to identify applicants with the ability to succeed academically once enrolled as a student at an English-speaking academic institution.

Although sufficient research has been done to understand the use of TOEFL iBT or IELTS, to the researcher's knowledge, none of those studies discuss the predictive validity of TOEFL iBT or IELTS on the academic success of OUAC 101 applicants. As the OUAC 101 applicant pool is unique to the 21 publically-funded universities in Ontario and cannot be found as a specific applicant pool of any other institution, predictive validity studies of this sort are virtually non-existent. Therefore, Ontario universities are required to generalize from the results of existing studies which use samples of international undergraduate or graduate students in order to make decisions regarding the policies and procedures governing the admission of OUAC 101 applicants. As there are a number of external factors influencing academic success, such as student background, social adaptability, financial need and motivation (Light, Xu, & Mossop, 1987; Telbis, Helgeson, & Kingsbury, 2014), predictive validity studies of academic success should not be generalized. Instead, this exploratory study aimed to provide an understanding for the relationship between TOEFL iBT or IELTS and the academic success of OUAC 101 applicants with the anticipation that those results would provide the Ontario university used in

this study with the resources required to make procedural decisions regarding the admission of OUAC 101 applicants.

1.1.2 English Proficiency Fraud

Although causes of educational fraud in relation to English language proficiency tests in Western nations are not well documented, its existence raises a concern for university admission offices and senior educational leaders. It is suspected that the pressure placed on applicants to meet high minimum English language proficiency requirements set by some institutions has contributed to the decision by some applicants to employ unethical methods, such as hiring an impersonator to write their test, in order to ensure they meet university admission requirements. This Ontario university, like many other Ontario Universities, has received, investigated, and subsequently declined fraudulent test scores submitted by OUAC 101 applicants whose first language is not English and who had studied for fewer than three years in an Ontario high school. The ability of those tests to be falsified or counterfeit is a concern for university admission offices which key test developers have neglected to effectively address. Upon receiving suspected fraudulent test scores during the admission process, the Educational Testing Service (ETS) (the creator of TOEFL) and the British Council (the creator of IELTS), have agreed to launch formal investigations as to the accuracy of those test scores. This process can be time-consuming and universities must often make decisions regarding the acceptability of those tests before receiving the formal review. While the decision to deem a test score as potentially fraudulent and subsequently reject it is at the discretion of the university, the time needed to examine and investigate those test scores has required a large commitment by administrative staff which universities are rarely structured to easily accommodate.

Research addressing fraudulent English language proficiency test scores is lacking at least in part to the proprietary nature of such test scores and data (Guhr, & Furtado, 2013). As varying types of fraud exist, such as impersonation or counterfeit documents, universities are generally not able to determine the type of fraud committed nor are they easily able to prove that fraud exists. Although testing companies may have data that point to possible fraud, they are not eager to release such proprietary data. Although the issue of English language proficiency test fraud has made its way into the media (eg. "The China Conundrum", 2011), the scant amount of academic literature and availability of proprietary data has limited the resources available for senior educational leaders within universities who require an understanding of the current issue in order to develop effective policies and procedures for combating this concern.

Indeed, this increasingly prevalent issue has brought into question the current practices and the use of English language proficiency tests for admission purposes. Using the resources currently available, this Ontario university has been trying to develop consistent and ethical methods for reviewing English language proficiency tests in which identifiable markers on each test can be compared with the prospective student's application in order to identify any inconsistencies. For example, bio-demographic data available on the TOEFL iBT or IELTS test scores may be compared with bio-demographic data provided by the applicant on their application to ensure that all data are consistent. The difficulty with this method is that many available bio-demographic markers which are found on both the English language proficiency test score and the prospective students' application can also be easily produced in counterfeit form, invariably resulting in the inconsistent identification of fraudulent test scores. This study aimed to use a grade component (ENG4U) of the prospective students' application which is submitted directly by their school as a marker that can be compared with TOEFL iBT or IELTS subsection scores to identify potentially fraudulent test scores.

1.2 Purpose and Significance

It can be argued that one purpose of education within Canada is to create and develop global citizens who are able to contribute to society by educating for intellectual and social democratic development (Darling-Hammond, 1998). Specifically, Darling-Hammond (1998) suggests that "education for democracy requires not only experiences that develop serious thinking, but also access to social understanding developed by actually participating in a democratic community and developing multiple perspectives" (p. 87). In order to create this high quality education, publically funded universities try to internationalize their education systems by recruiting staff, faculty and students from around the world. The intention of recruiting globally is to attract a wide array of applicants who may possibly hold different perspectives and have different lived experiences and knowledge. By diversifying the university population in such a way, it is the hope that universities will be able to continue to grow in the competitive global market while developing global citizens.

During the internationalization of the university community, public universities in Canada must uphold their responsibility to the stakeholders contributing to their funding. In doing so, those publically funded universities must ensure that they have objective procedures underlined by ethical policies which are put in place to identify the best and brightest students. As part of the admission process senior administrators and educational leaders set admission requirements to increase the likelihood of applicants being selected for admission that are able to succeed academically once enrolled as a student. As previously mentioned, it is important that applicants selected based on those admissions criteria are able to succeed academically in order to benefit the university by allowing the intellectual and social cultures to prosper while attempting to achieve high retention rates; both of which contribute to the reputation of the university. For the

university used in this study, this means retaining the reputation of a Tier 1 research intensive university in Canada while subsequently attracting more prospective students. In order to ensure that English language proficiency tests are being appropriately used and that the appropriate minimum requirements have been set, it is critical that universities conduct studies to analyze the ability of TOEFL iBT and IELTS to identify applicants with the potential to succeed once enrolled as students.

Predictive validity studies are common in the field of education specifically when determining the relationship between TOEFL iBT or IELTS and academic success. Although IELTS has been receiving more attention in the last fifteen years, TOEFL is still widely accepted as the primary instrument for making admission decisions (Jamieson, Jones, Kirsch, Mosenthal, & Taylor, 2000; Zareva, 2005). For this reason, and based on the observation that TOEFL iBT and IELTS are the two most widely submitted test scores to this Ontario university both TOEFL iBT and IELTS have been selected for use during this study.

This study also came in response to a gap in the current literature which neglects to address the use of TOEFL iBT or IELTS by OUAC 101 applicants. Although the OUAC 101 applicant pool is unique to the 21 publically funded universities in Ontario, research regarding English language proficiency tests also neglect to differentiate between international applicants and those applicants required to submit proof of English language proficiency as discussed earlier. As domestic applicants may also be required to submit proof of English language proficiency, it is important that this study looked at all applicants who submitted TOEFL iBT or IELTS within a given applicant pool, which in this study included OUAC 101 applicants.

In addition to the concerns regarding the predictive validity of TOEFL iBT or IELTS, the issue of English language proficiency test fraud has become an increasing concern to many universities around the world. Although academic literature supporting the existence of such fraud is limited, its occurrence is known within the academic community and the media. As the concern for fraudulent test scores perpetuates within universities, the efficacy of particular policies and procedures currently used to vet applicants are called into question. Due to the proprietary nature of data in support of such test fraud, universities are limited in the resources available to combat this issue. In order to identify fraudulent test scores, it is thought that identifiable markers on the prospective student's application and test score may be compared to determine if any discrepancies arise which may suggest the existence of a potentially fraudulent test score. In such cases, strong significant correlations must be found between those comparable markers and test scores, which consequently identify strong predictive relationships. It is deviations from those strong predictive linear relationships that may suggest inconsistencies in the applicants' abilities, subsequently recommending that a more in-depth review be completed. It was the goal of this study to understand the relationship between TOEFL iBT and IELTS with ENG4U grade, which was used as a marker representing the applicants' prior English ability.

This study was intended to complement the existing body of knowledge examining the relationship between TOEFL iBT and IELTS test scores and academic success while also offering a method for identifying fraudulent TOEFL iBT or IELTS test scores. The results of this study will be distributed to the University Registrar and Vice-Provost of Academic Affairs at this Ontario university. Although the results of this study are primarily intended to inform the admission policies and procedures set out by this specific Ontario university, the information

from this study may lead other universities, especially those in Ontario, to re-examine and perhaps re-shape their admission processes.

1.3   Research Questions

The following research questions were developed in order to explore the predictive validity of TOEFL iBT and IELTS at the current English language proficiency requirements for Ontario high school applicants (OUAC 101) whose first language is not English, have studied for fewer than three years in an Ontario high school and who have applied for undergraduate admission at one Ontario university to address the problems associated with suspected fraudulent TOEFL iBT and IELTS test scores. Within each of the two research questions described below, several sub-questions were used to guide the methodology of this study (as multiple correlation and multiple regression analyses were required to answer the two major research questions addressed by this study). Independent hypotheses were also developed for each of the two research questions.

1.3.1   Research Question 1

Are either TOEFL iBT or IELTS suitable for identifying Ontario high school applicants applying for undergraduate education at one Ontario university, whose first language is not English, have studied for fewer than three years in an Ontario high school, and who are likely to succeed in an English-speaking academic environment?

1.3.1.1 Sub-questions

1.1 Are the TOEFL iBT and/or IELTS Overall test scores, of all participants, able to predict participants' academic success, as measured by first-year GPA?

1.2 Are the TOEFL iBT and/or IELTS Overall test scores, meeting the 'minimum score' or the 'minimum discretionary range' requirements, able to predict participants' academic success, as measured by first-year GPA?

1.3 Are any of the TOEFL iBT and/or IELTS subsection scores able to predict participants' academic success, as measured by first-year GPA?

1.4 Does the average GPA of participants submitting TOEFL iBT or IELTS differ significantly amongst the three subsample groups for each test: 1) 'all participants', 2) those meeting the 'minimum score' and 3) those meeting the 'minimum discretionary range?

1.5 Are TOEFL iBT and/or IELTS able to determine if an applicant will succeed, as determined by meeting their first-year GPA progression requirement of 55%, in their first year of undergraduate education at one Ontario university?

1.3.1.2 Hypothesis

It can be hypothesized that each TOEFL iBT and IELTS are able to predict applicants' future academic success, as measured by first-year GPA. The rationale for this hypothesis is rooted in the notion that English language proficiency tests, such are TOEFL iBT or IELTS, are able to determine an applicant's ability to succeed in an English-speaking academic institution (Cambridge ESOL, 2003; Educational Testing Services, 2011). Based on this notion, a hypothesis can be made to support a strong significant linear relationship between test score and academic success. Subsequently it can be hypothesized that based on this linear relation, a higher test score should indicate a stronger level of English language proficiency and essentially suggest an increased probability of academic success.

It should be noted that many factors in addition to English language proficiency, such as student background, motivation, social adaptability, academic ability and financial need, may

contribute to an applicant's ability to succeed academically as reflected by GPA (Light, Xu, & Mossop, 1987; Telbis, Helgeson, & Kingsbury, 2014). Essentially, if language proficiency were the only factor influencing academic success as measured by GPA, then it would be expected that no native speaker of English should fail within an English-speaking academic institution. For this reason the question may be raised as to the strength of the relationship between TOEFL iBT or IELTS with academic success.

1.3.2 Research Question 2

Can ENG4U grade, which is a prerequisite for all Ontario high school applicants applying for undergraduate education at one Ontario university, be used as a marker which in comparison to TOEFL iBT or IELTS test scores can be used to help identify potentially fraudulent test scores?

1.3.2.1 Sub-questions

2.1 Are the TOEFL iBT and/or IELTS Overall test scores, of all participants, able to predict participants' prior English ability, as measured by ENG4U grade?

2.2 Are the TOEFL iBT and/or IELTS Overall test scores, meeting the 'minimum score' or 'minimum discretionary range' requirements, able to predict participants' prior English ability, as measured by ENG4U grade?

 2.3 Are any of the TOEFL iBT and/or IELTS subsection scores able to predict participants' prior English ability, as measured by ENG4U grade?

1.3.2.2 Hypothesis

It can be hypothesized that a significant linear relationship will exist between TOEFL iBT and IELTS test scores and ENG4U grade. Specifically, it is hypothesized that a strong

significant linear relationship will exist between ENG4U grade and both the Reading and Writing subsections of TOEFL iBT and IELTS.

This hypothesis is rooted in the notion that ENG4U grade is the sum of assessments which evaluate a student's writing ability and often their reading comprehension by way of written assessments (Ontario Ministry of Education, 2007). As TOEFL iBT and IELTS claim to assess examinee's reading and writing ability, it is expected that a significant linear relationship would exist between these variables. In this study, a linear relationship is represented by a strong correlation coefficient.

1.4   Limitations

1.4.1   Sample Size

Sample size was a significant limitation of this study. Although three admission cycles were used within this study (2012 – 2014), the sample sizes were still relatively small, especially for those analyses which incorporated academic success. This is in part due to the fact that in order to generate a GPA for each relevant participant, the participant must submit a TOEFL iBT and IELTS test, receive an offer of admission, accept their offer of admission, and complete their first year of study. In some cases where academic success was used, specifically within the 'minimum score group' samples sizes were too low to reliably infer any interpretations. As an exploratory study, it was the decision of the research to present data for those groups, although not statistically reliable, in the event that questions arose in response to those data which could be examined in further studies.

1.4.2   Fraudulent Test Scores

Since not all fraudulent test scores may be identified, the possibility exists that fraudulent test scores may have been used in this study. Therefore, correlations analyses may be confounded by unidentified fraudulent test scores. Unfortunately, it is impossible to eliminate all fraudulent test scores or ensure that all results are truly reflective of all applicants' individual abilities.

1.4.3   Range Restriction

The use of GPA as a measure of academic success places a restriction on the range of data available to be used within this study. For applicants to be included in the academic success portion of this study, they must have submitted either a TOEFL iBT or IELTS test score, been admitted to the Ontario university used within this study, accepted their offer of admission, and subsequently completed their first year of undergraduate study. Additionally, in order to have received an offer of admission, applicants must have provided a test score above the 'minimum discretionary range' requirement, theoretically limiting the lowest test score to the base of the 'minimum discretionary range'. Unlike the practices of the university used within this study (which uses the most recent test score when assessing applications for OUAC 101 applicants), the first test score submitted to this Ontario university, acceptable or not, was used within this study. Using the first test score received was intended to limit the effect of any learned experience the applicant may have endured by writing the test multiple times and subsequently resulted in a wider range of usable data decreasing the amount of range restriction placed on this study (as those test scores initially submitted which did not meet the 'minimum discretionary range' requirement were not restricted by the lower end of the 'minimum discretionary range' requirement). Therefore, TOEFL iBT and IELTS test scores used in this study may fall below

the 'minimum discretionary range' requirement. It should be noted that eventually those applicants who did not meet the 'minimum discretionary range' requirement on their first attempt, provided an acceptable test score suggesting that range restriction may still act as a limiting factor to this study.

1.4.4   Generalizability

It is important to note that the quantitative nature of this study does not allow the correlative effects of TOEFL iBT or IELTS on academic success to be generalized beyond the one Ontario university used in this study. The composition of students applying to other universities may vary depending on the programs offered, the admission cut-off averages or any number of external factors. Therefore this information should only be used by other universities as a basis for further internal considerations.

1.5   Delimitations

1.5.1   OUAC 101 Applicants

Throughout the construction of this study many restrictions were set in order to provide valid results. The choice to include only OUAC 101 applicants allowed for all applicants to be reviewed with an assumed common level of knowledge as all applicants were taught the Ontario high school curriculum, at least at the Grade 12 level. This assumption is not one that can be made for any other applicant group at this Ontario university. Although it can be argued that the level of knowledge gained under the same curriculum may still vary, this theoretical assumption allowed this study to have somewhat of a measurable level of prior English ability. Since all OUAC 101 applicants are required to complete ENG4U and have this grade used in the calculation of their admission average, ENG4U grade was selected to measure prior English

ability. It is important to note that although the selection of OUAC 101 applicants allowed for certain delimitations, including common prior education, applicants submitting the OUAC 101 application must have at least six Grade 12 university and/or mixed level courses including program prerequisites. Since only this limited number of courses was required, it is possible that applicants may have studied only one year in an Ontario curriculum-based school while still being considered an OUAC 101 applicant. Although those factors allowed for the working assumption that all OUAC 101 applicants have some common prior education (at least six Grade 12 university and/or mixed level courses including ENG4U) the number and type of courses comprising this common prior education cannot be controlled, with the exception of ENG4U.

As previously mentioned, the notable amount of fraud among TOEFL iBT and IELTS test scores observed by admissions officers at one Ontario university, where the researcher is currently employed, resulted in the identification of bio-demographic markers which may help identify potentially fraudulent test scores which should be put forward to ETS or the British Council for review. As no statistical data are available to support or discourage the use of those markers, no definitive decisions are able to be made based on this review. When trying to identify markers, the use of OUAC 101 applicants allowed for the working assumptions that OUAC 101 applicants should be enrolled in an Ontario high school from September to June and that all TOEFL iBT or IELTS test scores written within this time should also be written within Ontario. If a test score was written outside of Ontario within the time an applicant should be at their Ontario high school, then the admissions office could conduct further investigations which may include contacting the applicant's high school for attendance records to verify their presence at their high school in Ontario on the days surrounding the test score. In the event that the applicant was in Ontario while claiming to write a test in another country, it is suggested that the test score be put forward to ETS or the British Council for formal investigation.

1.5.2   Application Period

In order to increase sample size in this study, applicants to the 2012, 2013, and 2014 admission cycles were used. The sample was limited to those three admission cycles as the admission requirements during those years were very similar. Such a delimitation was intended to control for some of the factors influencing the admission process other than English language proficiency.

1.5.3   Multiple Submissions

Since secondary data were used during this analysis, this study was confined to the information provided in those data outputs. Although applicants could submit multiple English language proficiency tests scores or ENG4U grade, only the first attempt received was used in this study.   As previously mentioned, using the first attempt was intended to control for any learning which may have taken place during the initial attempt that could have impacted subsequent attempts. It should be noted that additional data such as the reason for retaking ENG4U was not provided. Therefore, any special circumstances which may have accounted for a low first ENG4U grade was not taken into consideration and may have skewed the data. For examples, a student undergoing an extensive medical procedure may have been required to miss a substantial amount of school which subsequently influenced his/her grade. In this case, the second attempted ENG4U grade may be more representative of the participant's abilities. Additionally, it cannot be guaranteed that the first TOEFL iBT or IELTS test score attempt received by the university in this study was in fact the applicant's first attempt at writing an English language proficiency test as only those test submitted could be identified.

1.6   Researcher's Perspective

No *a priori* theoretical perspective was identified to analyze the data. Instead the

quantitative data from this study was discussed as data, and the implications of such data were

identified based, in part, on the researcher's current experience as an Admission Officer. For this

reason it is important to outline the relationship between the researcher and this study.

The Ontario university in which the study was nested was one familiar to the researcher.

Having held the position of the Admission Assistant and then subsequently the Admissions

Officer for the Ontario high school admission team within an Office of the Registrar assisted my

implementation of the study. It is with the benefit of this background that the researcher noted a

gap within the current literature and the impetus for this study to be created. To the researcher's

knowledge, there is currently no research literature published relating TOEFL iBT or IELTS with

OUAC 101 applicants. Despite attempts to reduce bias on the part of the researcher, the

researcher believes that it is important for universities to understand their policies with relation to

OUAC 101 applicants as research literature pertaining to this group is essentially non-existent. It

is important to evaluate current policies and procedures in order to ensure that the Ontario

university within this study is using the appropriate admissions criteria when evaluating OUAC

101 applications. As a Tier 1 publically-funded university, the public purposes of this Ontario

university are upheld by their loyalty and commitment to tax-paying members of society. As

such, this university takes pride in their ability to establish and implement the most ethical

procedures possible. This study may act as a resource to inform and assist senior administrative

leaders when making difficult admission or procedural decisions. For this reason, the researcher

set out to analyze the current policies and procedures for admitting OUAC 101 applicants who

are required to submit proof of English proficiency and to also investigate the ability of TOEFL

iBT and IELTS to identify applicants with sufficient English language proficiency to potentially succeed academically once enrolled as an undergraduate student.

Although the view of the researcher is embedded in a positivist perspective in which deductive reasoning and the scientific method have led the methodology of this study, it was the experiences of the researcher which guided the discussion and interpretation of those results. It can be argued that the positivist perspective does not easily lend itself well to research addressing social science question, as many theories in social science are the result of multiple factors, some of which cannot be controlled within the scientific method (Cohen & Morrison, 2007). Although the scientific method was used to develop the methodology of this study, it is acknowledged that many factors which could not be controlled in this study impact academic success. For this reason, the positivist approach was complemented by the qualitative experience of the researcher when discussing the implications and future recommendations of this study.

1.7   Definitions

Academic Success: Academic success was measured by the first-year cumulative average (referred to in this study as first-year GPA) of all first-year courses at one Ontario university between September and May. Typically academic success is a measure of how a student performs academically by way of measured course assessments. Therefore any GPA, regardless of the minimum GPA required by an institution for progression within a program, can be considered an applicant's academic success.

Applicant: Any individual who submitted an application to university.

Education Testing Service (ETS): A non-profit organization focused on design, assessment, and research to provide impartiality in educational evaluations worldwide (Arcuino, 2013). ETS is responsible for creating and administering TOEFL.

Fraud: Although possible in many forms, fraud is the deception of one's abilities intended to result in personal gain. For this study, English language proficiency test fraud may be committed in the form of impersonation, plagiarism, cheating, the creation of counterfeit documents, or any method which resulted in a test score not representing to the participant's ability.

Grade 12 University level English (ENG4U): The Grade 12 university level English course required by all Ontario high school applicants applying to the Ontario University used within this study.

Grade Point Average (GPA): Grade point average (GPA), in the traditional sense, is a measured scale from 0.0 to 4.0. Most predictive validity studies use the term GPA to represent participants' academic achievement at a given point in time. For the purpose of this study, GPA was calculated as the participants' first-year cumulative average between September to May. Summer courses were not used as part of the GPA in this study.

International English Language Testing System (IELTS): An internationally recognized English language proficiency test developed through a collaborative effort between the British Council, International Development Program of Australian Universities and Colleges, and the Cambridge English Language Assessment.

Minimum Discretionary Range: The lower of two minimum English language proficiency requirements set out by the Ontario university used within this study. Although used in practice as the minimum English language proficiency test score required by applicants to prove they have sufficient English language proficiency to succeed academically as an undergraduate student, this requirement was originally set by University Senate to be used in combination with other evidence of English language proficiency (University X, as in References, 2013, p.1).

Minimum Score: The University Senate published minimum English language proficiency test score which an applicant must meet in order to be considered as having sufficient English

language proficiency to succeed academically once enrolled as an undergraduate student (University X, as in References, 2013, p.1).

Ontario Universities' Application Center (OUAC): The applicant center used when applying to any Ontario University for full-time undergraduate admission.

OUAC 101 Applicant: Applicants applying to undergraduate education through the OUAC 101 application. All prospective students using this application must be current Ontario high school students or studying the Ontario high school curriculum in another country.

Predictive Validity: The ability of an instrument, such as a test, to predict a criterion variable.

Prior English Ability: For the purpose of this study, prior English ability is measured by grade 12 University level English (ENG4U) grade. It is important to note that ENG4U is not the sole indicator of prior English ability but has been selected in this study due to the common curriculum among all OUAC 101 applicants.

Registrant: Any participant who following the submission of their application was admitted, accepted their offer of admission, and subsequently registered in courses and completed their first year of undergraduate study.

Reliability: The degree to which an assessment instrument can be reproduced and produce stable and consistent results.

Test of English as a Foreign Language (TOEFL): Administered by ETS, the TOEFL aims to evaluate a student's ability to succeed in an English median academic environment as well as a student's ability to understand and communicate English effectively (Darnell, 1968).

Validity: the degree to which an instrument accurately measures what it claims to measure.

Chapter 2

2 Literature Review

Since the introduction of the internet to the public, globalization and internationalization of education have begun to take a form that was not previously thought possible. The internet has made it easier to connect with others and learn about education and employment opportunities that are available almost anywhere around the world. The idea of reaching potential applicants across the world is attractive to educational institutions as it symbolized the possibility for student enrolment growth and diversity. One goal of internationalizing an institution extends beyond increasing the international population which may provide a financial means not previously tapped for many universities, but it allowed for the creation of a new organizational culture which supports a culturally rich and diverse education and global presence. In today's consumer-driven society, some may argue that students' are merely consumers of the educational opportunities universities have to offer (Blackmore, Brennan & Zipin, 2010). Based on such ideology, the notion of recruiting international students is seen solely as a financial gain. For most publically-funded universities in Ontario, the internationalization of education allows universities to expand the critical thinking and global understanding of the student population through the inclusion of intercultural and global perspectives (Darling-Hammond, 1998). The Ontario university used within this study makes a priority of providing students with a pluralistic multicultural and diverse education and aims to attract the world's brightest minds (University X, as in References, 2014, p. 5). In addition, this Ontario university has made it their mission to educate students with the knowledge, skills, values and dispositions required to benefit society while graduating global citizens whose education and leadership will one day serve the public (University X, as in References, 2014, p. 5).

Internationalization and globalization are not only a benefit to institutions, but they create possibilities for many students to pursue a different future than would have originally been possible in their own community. In the hopes of attending a Canadian university, many students from around the world begin to prepare themselves to meet the stringent admission requirements set out by these top tier universities. In some cases, students attend Canadian-based curriculum private schools in their home country, or leave their families to attend a Canadian high school (CBIE, 2009). In all cases, however, students whose first language is not English—be they international students or domestic (i.e. students studying on a study permit/student visa or those holding Canadian Citizenship or Permanent Residency status)—are required to submit institution-accepted proof of their English language proficiency to Canadian educational institutions as part of the application process.

2.1   University Admission Procedures

Attending a Canadian University is a goal for many prospective students around the world (CBIE, 2009). Once applicants, these prospective students fall into two major groups based on the application submitted through the Ontario Universities' Application Center (OUAC) website: "OUAC 101" applicants who are currently studying in an Ontario high school or Ontario curriculum-based school abroad, and "OUAC 105" applicants who are applying with any other type of educational background. This study focused on a sample comprised of OUAC 101 applicants whose first language was not English, who have studied for fewer than three years in an Ontario high school and who applied for admission to any undergraduate first-entry program at this Ontario university with the exception of the Nursing program (where the English language proficiency requirements are different due to the program's practical placement component).

Upon the completion of the OUAC 101 application the university used within this study automatically requests proof of English proficiency from any OUAC 101 applicant whose first language is not English and who has studied for fewer than three year in an Ontario high school. The need to request proof of English language proficiency from any applicant who has studied for fewer than three years in an Ontario high school is in line with Cummins (1981) notion that it takes approximately four to seven years to develop 'cognitive-academic language proficiency' (CALP) in another language. Therefore incoming students who have fewer than three years of English language experience within Ontario would be expected to have lower CALP, allowing the prediction to be made that those students may have difficulty keeping up with the reading load and gaining a comprehensive understanding of their reading; understanding fast paced lectures; and interacting with and/or contributing to conversations amongst classmates and colleagues.

To satisfy the English language proficiency requirement, applicants must submit and meet the minimum requirement of an institution approved English language proficiency test. Although the tests accepted and the minimum requirements of each test may vary by institution, this Ontario university accepts six English language proficiency tests: TOEFL, IELTS, Pearson Test of English Academic, Michigan English Language Assessment Battery (MELAB), the CanTEST, or the Canadian Academic English Language Assessment (CAEL). Of the tests accepted by this university, TOEFL iBT and IELTS are by far the most commonly provided test scores received.

Once a test score is received the Admissions Assistant and Admissions Officer would compare the bio-demographic data markers on the test to the applicant's profile to ensure that the test score was properly identified. Following this verification, the Admissions Assistant and Admissions Officer coded each test score submitted as part of the applicant's profile. This allows

29

for an automated process to identify which applicants have met the minimum English language proficiency requirement and can therefore be reviewed for their academic admissibility. All English language proficiency test scores are valid for two years from the time they were written. In the event that multiple test scores are received for the same applicant, this Ontario university uses the most recent test score to meet the minimum English language proficiency requirement. It is important to note that if an unacceptable test score is received after an applicant has submitted an acceptable test score and received an offer of admission, a condition is placed on the applicant's offer which requires them to provide a new acceptable test score. If this condition is not met then the applicant's offer of admission will be withdrawn.

In 2009 University Senate at the Ontario university in this study approved the publishing of minimum English language proficiency test scores required for admission (as no requirement was previously published). In doing so, University Senate published a 'minimum score' and a 'minimum discretionary range' requirement for each acceptable test (see Table 1). For admission, the 'minimum score' requirement requires applicants to achieve a higher English language proficiency test score than does the 'minimum discretionary range' requirement. Based on the published Senate policy for English language proficiency requirements, the lower range ('the minimum discretionary range') requirement was intended to be used at the discretion of the university's Admission Committee in combination with supplementary documentation. In practice, it is the 'minimum discretionary range' requirement that is published in all recruitment communications and is used when evaluating the English language proficiency of OUAC 101 applicants. Although it is not currently used in practice, the higher range ('minimum score') requirement is in line with the high English proficiency requirements of many other top tier Ontario universities.

Once an applicant has satisfied their English language proficiency requirement, their application will be evaluated based on an academic admission average. The admission average is calculated using the applicant's top six Grade 12, university- and/or mixed-level courses, including program prerequisites. The Ontario university in this study accepts grades from all university or mixed-level courses taken through a school approved by the Ontario Ministry of Education with the exception of co-op courses. If a course is repeated, this Ontario university will use the highest grade received in the average calculation. A core admission requirement which must be used in the average calculation for all programs at this Ontario university is Grade 12 university-level English (ENG4U) (which was used as a measure of prior English ability within this study).

Table 1

University English Language Proficiency Admission Requirements for Undergraduate Applicants

Test	Minimum Score	Minimum Discretionary Range
Test of English as a Second Language (TOEFL) – Internet	88 overall with a score of 22 in speaking, reading and writing, and 20 in listening	83 overall with no score less than 20
International English Language Testing System (IELTS)	7.0 overall with no part less than 6.5	6.5 overall with no part less than 6.0
Michigan English Language Assessment Battery (MELAB)	85 overall with no score less than 80	80 overall with no score less than 78
CanTEST	4.5 overall with no part less than 4.0	4.0 overall with no part less than 4.0
Canadian Academic English Language Assessment (CAEL)	70 overall with no part less than 60	60 overall with no part less than 60

Grade point average (GPA) is a measurement used in the majority of current literature to represent academic success (Al-Musawi & Al-Ansari, 1999; Arcuino, 2013; Cho & Bridgeman, 2012; Cotton & Conrow, 1998; Dooey & Oliver, 2002; Feast, 2002; Fu, 2012; Ng, 2007; Simner, 1998; Wait & Gressel, 2009; Woodrow, 2006; Yen & Kuzma, 2009). Although the definition of what GPA measures may vary, the term GPA has come to represent student achievement (expressed numerically) at a given point in time. The Ontario university used within this study does not report grades as part of the formal GPA scale (0.0 to 4.0), but instead represents grades as a percentage value. In order to be consistent with current literature, this study defines academic success as 'first-year GPA', which is equivalent to first-year cumulative average. GPA at the end of first year was selected for use in this study as it identified the most recent indicator of academic success since the test score was written; decreasing the number of external factors able to impact the students' English language abilities from the time they have written their English language proficiency test to the time their academic success is calculated (Simner & Mitchel, 2007).

2.2   Academic Success

The majority of research considering the relationship between English language proficiency and academic success can be identified as academic prediction studies or predictive validity studies which investigate the relationship between a predictor variable (TOEFL iBT or IELTS), and a criterion variable (academic success). Graham's (1987) pivotal research reviewed the use of predictive validity studies with regards to English language proficiency and identified four major problem which should be taken into consideration with any predictive validity study; a) the criterion for judging academic success, b) the validity of TOEFL iBT and IELTS, c) the interpretation of any relationships found, and d) the large number of uncontrolled variables

32

influencing academic success. Although other studies like those found within this literature review, focus on the relationship between English language proficiency and academic success, Graham's (1987) research was one of the first major studies to outline the inconsistency in results across such studies.

Graham (1987) noted that there are a number of moderating variables that could impact a student's performance on a language proficiency test, or academic success. The concern for the complexity of language proficiency and academic success has been echoed by many other researchers who have been able to identify factors which influence language proficiency or academic success. For example, Light et al. (1987) determined that the relationship between TOEFL and academic success varied depending upon the students' area of study. In addition, Light et al. (1987) attributed some of the variance in these results to other unexplained factors such as the student's background and motivation. In another study, Telbis, Helgeson, and Kingsbury (2014) identified social adaptability, academic ability, and financial need as valid difficulties other than language proficiency for international students. Cho and Bridgeman (2012) echo the finding by Telbis et al. by discussing that language is just but one crucial factor in learning and that linguistic proficiency does not guarantee academic success.

Although many factors contribute to the variance seen in support of or against English language proficiency tests for academic purposes, Arcuino (2013) attributes some of this variance to the definition of academic success. The use of GPA is often criticized in regard to academic success due to the wide range of academic and non-academic factors influencing academic success (Graham, 1987). Light et al. (1987) clearly stated that, in the face of all criticisms, they continue to use GPA as a measure of academic success because GPA frequently determined whether or not a student was able to remain in graduate school and was therefore criteria for academic success. Despite the debate as to the accuracy of GPA to represent past and the

prospect of continued academic success, the majority of current studies consider GPA a measure of academic success. In all cases, a definition of academic success and justification for the use of GPA based on contemporary evidence is warranted.

Once GPA is accepted as a measure of academic success, consideration must be given to the type of GPA used. Arcunio (2013) defines academic success as cumulative GPA at degree completion, whereas Ng (2007) views academic success as a function of GPA, course completion, and retention. Manganello (2011), on the other hand, nominally uses GPA of a mandatory English composition class as representing academic success. Al-Musawi and Al-Ansari (1999) took an approach similar to Arcunio (2013) by using the cumulative GPA. Each type of GPA may produce different results. Simner and Mitchel (2007) suggest that using students' GPA at the end of their first year will produce the most accurate predictive validity study, as identifying the most recent indicator of academic success will allow for less external factors to impact the students' English language ability from the time they have written a test of English language proficiency. Essentially GPA acts as a proxy that is nominally used as an indicator of academic success and therefore must be defined for each situation and purpose to allow for accurate interpretations of results.

2.3   Test of English as a Foreign Language (TOEFL)

TOEFL is considered the primary instrument for making admission decision (Jamieson, Jones, Kirsch, Mosenthal, & Taylor, 2000). It is accepted as a credible instrument in more 130 countries and over 9,000 universities, colleges, agencies and other institutions including nearly every university in Canada and the United States (Educational Testing Services, 2011; Manganello, 2011; Simner & Mitchell, 2007; Zareva, 2005). The Educational Testing Service (ETS), the creator of TOEFL, is a non-profit organization focused on design, assessment, and

research to provide impartiality in educational evaluations worldwide (Arcuino, 2013). In the 1960's and 1970's, studies of English language proficiency, specifically TOEFL, were a main focus of educational research (Darnell, 1968). Clearly, to determine a student's ability to succeed in an English medium academic environment, students must be able to understand and communicate English effectively (Darnell, 1968; Zareva, 2005). TOEFL aims to evaluate this while providing test takers with the confidence of knowing they have the skills and knowledge to meet any subsequent demands of academic learning in universities (Zareva, 2005).

2.3.1 TOEFL History

Since the introduction of TOEFL in 1964, TOEFL has adapted from their paper-based test (TOEFL PBT) to a computer-based test (TOEFL CBT), and most recently an internet-based test (TOEFL iBT) (Educational Testing Services, 2011). These changes have been accompanied by changes in test content and were made in response to concerns from scoring users, language testing specialists, English as a second language teachers, and as a result of the findings produced by the ETS staff and outside experts. Although the TOEFL PBT was modified several times between 1964 and 2005, the TOEFL PBT continued to be the dominant English language proficiency test accepted worldwide (Zareva, 2005). For this reason, the majority of admission procedures have been developed with TOEFL PBT as a model. Surveys of United States and Ontario universities both revealed that the minimum TOEFL PBT admissions requirements are typically set at 550 (Zareva, 2005; Simner & Mitchell, 2007). It is suggested that this requirement is the minimum score required for an international student to be proficient in English (Zareva, 2005).

The TOEFL CBT was introduced in 1998 just seven short years before the introduction of the TOEFL iBT (Stoynoff, 2009). Based on the short lifespan of TOEFL CBT, during which the

TOEFL PBT was still accepted, there is very limited research on the TOEFL CBT (Manganello, 2011). Although the majority of literature on TOEFL focuses on TOEFL PBT, the TOEFL iBT has become the preferred test among test takers and universities. Despite being the preferred measure of English language proficiency for academic purposes, TOEFL iBT remains under-researched in comparison to TOEFL PBT (DeLuca, Cheng, Fox, Doe, & Li, 2013). In addition, it has not been determined if the results of TOEFL PBT studies and the link with academic success apply in the same way to the TOEFL iBT. Arcuino (2013) found that students typically score higher on the TOEFL iBT than the TOEFL PBT. Therefore, there is a pronounced need for research investigating the relationship of TOEFL iBT to academic success in order to modify any current admission procedures (which have been developed based on the TOEFL PBT).

Although the introduction of the TOEFL iBT came as a surprise to many, Zareva (2005) attributes the introduction of the TOEFL iBT to the need for testing multiple skills and abilities used to communicate in an academic setting. For this reason, the TOEFL iBT has been structured to measure students' ability to use English effectively in an academic setting (Stoynoff, 2009). While it may be assumed that English language proficiency and the ability to use English effectively are synonymous, and therefore both represent the students' ability to succeed in an academic setting, ETS advises that TOEFL should not be used "as the sole criterion for admission or as a predictor of academic success" (Education Testing Service, 1994, p. 8).

TOEFL iBT was first launched in North America in 2005 (Education Testing Service, 2011) and distributed to the rest of the world in 2006 (Zareva, 2005). TOEFL iBT consists of four major subsections (Listening, Reading, Writing, and Speaking) that once assessed individually contribute to an Overall score (Sawaki, Stricker, & Oranje, 2009) and takes approximately four hours to complete on a computer via the internet at a secure test center (Educational Testing Service, 2010; Stoynoff, 2009). The introduction of TOEFL iBT brings

with it many notable developments including the addition of the Speaking section; the elimination of the grammatical component; the expansion of the Writing sections; the inclusion of integrated tasks (specifically for the Speaking and Writing sections); and the ability to take notes (DeLuca et al., 2013; Educational Testing Service, 2010; Stoynoff, 2009). ETS test developers' intentions for the TOEFL iBT were a) "to stimulate university communication", b) "to help test takers determine their academic readiness" and c) to help "institutions identify and select students with the English communication skills required to succeed" in a post-secondary education (Educational Testing Service, 2010, p. 5). Although TOEFL iBT continues to be dominated by multiple-choice questions across the four subsections, the extended integrated and independent Speaking and Writing sections facilitate a more accurate representation of English for academic purposes (Educational Testing Service, 2010). With all of these changes to TOEFL iBT, it is important to develop an understanding of student performance within this new testing context (DeLuca et al., 2013). ETS (2010) described this new context in the following way:

> *The TOEFL test gives students the opportunity to prove they can communicate ideas effectively by simulating university classroom and student life communication. The language used in the test reflects real-life English-language usage in university lectures, classes, and laboratories. It is the same language professors use when they discuss coursework or concepts with students. It is the language students use in study groups and everyday university situations, such as buying books at the bookstore. The reading passages are from real textbooks and course materials* (p. 4).

Based on this new understanding and context, the assumption was again introduced that TOEFL iBT score should be able to act as an indicator of students' future success in an academic institution (Brooks & Swain, 2014).

2.3.2　Research: TOEFL and Academic Success

Arcuino's (2013) analysis of TOEFL and IELTS and their relationship to academic success as defined by final cumulative GPA for international students who graduated with a master's degree within 2006-2011 suggested that TOEFL is significantly correlated to the academic success of international master's level students. No significant relationship was found between IELTS and academic success of international master's graduates. Notably, it was seen that students scored higher on the TOEFL iBT than the TOEFL PBT.

Manganelle's (2011) study tested the strength of the relationship between TOEFL scores, scores on the Iowa State University's own English Placement Tests, and grades obtained in English classes for international students. Weak positive correlations were observed when scores from each TOEFL subsection were correlated with grades in all four levels of English language classes. Notably, positive relationships were observed when Writing and Speaking subsections were correlated against English grades whereas negative relationships were observed when Reading and Listening subsections were correlated with the same grades.

Wait and Gressel (2009) investigated the relationship between TOEFL score and academic performance of 6,516 international engineering students at an American university in the United Arab Emirates. Although the relationship between TOEFL score and GPA was found to demonstrate positive statistical significance, this relationship was weaker for engineering students as opposed to students in other fields.

Al-Musawi and Al-Ansari (1999) analyzed whether students' score on TOEFL were a predictor of their academic success as measured by overall GPA for first and second year students enrolled in English Language and Literature programs. The model presented by Al-Musawi et al. suggested that TOEFL subsection scores did not contribute enough to academic

success as no significant linear predictor model was established. In addition, Al-Musawi et al. suggested using other tests, such as the First Certificate of English test, rather than using TOEFL.

Abunawas (2014) conducted a meta-analysis of international students' TOEFL scores and academic success as measured by GPA across 40 studies including 47 independent effect size values. This meta-analysis found an overall statistically significant positive relationship between international students' TOEFL score and their academic success as measured by GPA. As the effect size of this relationship is relatively small (.21), these results suggest that TOEFL score may be associated with GPA but does not imply causation. Notably, the studies used within this meta-analysis focused on different geographical populations of which it was found that studies conducted outside the U.S. had a higher effect size than those conducted within the U.S.

To determine if TOEFL was a good indicator of academic success for international students at community college, Ng (2007) used GPA, course completion, and retention as measures of academic success. Data were collected from 433 international students' academic records and found no significant relationship was found between TOEFL score and the number of ESL courses in relation to GPA, course completion, and retention rates.

In another earlier study, Light et al. (1987) found that TOEFL was not an effective predictor of academic success, as measured by GPA, for international graduate students. Specifically, students with low TOEFL scores were seen to have higher GPAs than those with high TOEFL scores. Although the authors attribute these findings to other factors such as student background and motivation, Lo (2002) also found results which supported this study in his review of TOEFL score and first-year GPA of freshmen international students.

Most of the studies reviewed in this section used the TOEFL PBT as only a small number of researchers have examined the TOEFL iBT subsections which include Listening, Speaking, Reading and Writing. In an investigation between these two versions of TOEFL (PBT and iBT),

Fu (2012) found that the two versions were similarly correlated to GPA with the exception of the Listening section. The Listening section of the TOEFL iBT was seen to be a more effective predictor of GPA than the TOEFL PBT, which supports the changes made to the TOEFL iBT. Fu (2012) furthered those results by examining the correlation of each subsection on GPA at multiple education levels. For example, the Reading section had the highest correlation with GPA for undergraduate students, while the Writing section had the highest correlation with GPA for students in a graduate level program. Fu's study emphasized the need for further analysis to be undertaken based on specific admit types and TOEFL subsection scores. Unlike many other studies discussed here, Fu (2012) broke down the overall score by subsection, which helped determine which portions of the overall score were in fact found to be correlated with GPA. In addition to correlating TOEFL with GPA, Fu (2012) also correlated high school GPA with first year GPA for undergraduate students and found that correlation to be the strongest predictor of academic success.

Eighteen years ago, Simner (1998) analyzed the use of TOEFL as an admission requirement; at which time he attributed the poor predictive validity of TOEFL to inadequate sample size or other methodological flaws such as range restriction. Simner (1998) focused his discussion on criticisms of how universities used TOEFL scores to make admission decisions rather than criticizing the TOEFL test itself. That view was unique compared to most other studies that investigated predictive validity. As mentioned above, the use of GPA as a measure of academic success remains limited by the fact that students must meet a minimum TOEFL score in order to be admitted, and subsequently register within a given institution and receive a GPA. This range restriction has a confounding effect on the results of all predictive validity studies that use GPA as a measure of academic success. Hence such a methodology should be noted as a major limitation of any such research. Unfortunately, the measures of academic

success that exist are limited at this point in time and no alternative method has been established or viewed as acceptable within the field of educational research.

Cho and Bridgeman's (2012) correlational research analysis of 2,594 undergraduate and graduate students of different disciplines investigated the relationship between TOEFL and academic success in higher education as defined by GPA. As one of the most extensive studies conducted to date, Cho et al., found a small correlation between GPA and TOEFL iBT scores of graduate and undergraduate students ($r = .16$, and.18, respectively) (p. 438). Although the predictive validity of TOEFL on academic success was weak, those correlations suggest that there is some validity in the use of TOEFL iBT as a predictor of academic success. The amount of variance in GPA explained by this model is only 3% suggesting that external factors, other than TOEFL iBT, account for approximately 97% of the variance in GPA (Cho & Bridgeman, 2012, p. 438). Unique to other studies, Cho and Bridgeman's (2012) research expanded on their results and separated participants by top, middle, and bottom TOEFL scores and then correlated each group with GPA. Their analysis was a necessary step to validate their study as only small correlations were originally seen between undergraduate and graduate students. This tiered analysis which reviewed the top, middle, and bottom TOEFL score groups supported the predictive validity of TOEFL score as students with low TOEFL scores were seen to have low GPAs, and students with high TOEFL scores demonstrated high GPAs. The study by Cho et al., like many other studies using GPA as a measure, was limited by range restriction in which the minimum TOEFL score requirement was the lowest possible score used in this analysis.

2.3.3   TOEFL Reliability

To ensure reliability and comparability of TOEFL iBT, ETS conducts research on: the implementation of standardized administrative and test security procedures; the use of detailed

test specifications to guide test development; the monitoring of score reliability and generalizability; the employment of appropriate scorers; and the ability to maintain comparable scores across test formats (Educational Testing Service, 2011). When developed, the Reading and Listening subsections of TOEFL iBT were established based on item response theory (Lord, 1980), while the Speaking and Writing subsection were developed based on generalizability theory (Brennan, 1983). An analysis by the ETS of data from 2007 revealed that the reliability estimates for the Reading (.85), Listening (.85), and Speaking (.88) subsections are relatively high, while the reliability estimate of the Writing score (.74) is somewhat lower (Educational Testing Service, 2011, p. 5). Based on the results of this study it was recommended by ETS that when making high stakes admission decisions, post-secondary institution use the total score as it provides the best information, both in its ability to reflect all four language skills, and because it is the most reliable (.94) (Educational Testing Service, p. 5). ETS also encourages institutions to consider a number of other factors such as GPA, score on other admissions exams, teacher recommendations, and interviews when determining a student's admissibility.

The majority of research regarding the reliability of TOEFL iBT is associated with the ETS. For example, Zhang (2009) conducted a "Reporter Analysis for TOEFL iBT" which was published in the *ETS Research Memorandum, 2008*. In this analysis, Zhang (2009) used the fact that test takers quite often repeated the TOEFL iBT as a basis to examine its reliability. Test takers who repeated TOEFL iBT within one month were used by Zhang as an assumption that no sufficient learning should have occurred which could change the test results significantly. Therefore Zhang argued that any significant changes should be due to measurement error and the quality of the test scores. Although the scores of students improved slightly on the second attempt, the correlation between the two test attempts appeared to be very high for the total score and moderately high for the four subsections, suggesting consistency in repeater scores. Of the

changes exhibited between the first and second attempts, the Reading subsection had the highest score changes, whereas the Speaking subsection had the lowest score changes. The distribution of those data resembles a symmetric bell-shaped distribution, suggesting that the majority of repeaters did not have a major change in score. One reason for the larger change in subsection score compared to the overall score was a student's desire to improve the subsection they previously performed poorly on.

2.3.4   TOEFL Validity

According to ETS, TOEFL iBT was designed to measure 'the student's ability to use English effectively in academic settings' (Educational Testing Service, 2006, p.4). To validate the use of TOEFL as an English language proficiency test, ETS has produced more than 40 monographs and technical papers arising from its study's findings between 1997 and 2007 (Stoynoff, 2009). Sawaki et al. (2008) validated the use of TOEFL iBT through the use of confirmatory factor analysis. This analysis determined a high-order factor model with a single higher-order general factor (English as a second language ability) and four first-order factors corresponding to the four TOEFL iBT subsections. Therefore, Sawaki et al. claimed that this model broadly supported the current TOEFL iBT structure with one score for each subsection and a single overall score. Although those results do not support previous studies, most previous studies used the TOEFL PBT. This variation may be attributed to the new addition of integrated Speaking and Writing sections, which were not included in the TOEFL PBT.

2.4   International English Language Testing System (IELTS)

Although TOEFL is known as the leading English language proficiency test, IELTS is another common English language test accepted worldwide (Stoynoff, 2009). The *IELTS Handbook* (Cambridge ESOL, 2003) indicates that the test is "designed to assess the language

ability of candidates who need to study or work where English is the language of communication" (p. 2). To ensure accurate representation of results, IELTS has two test forms: an academic form intended for examinees applying for admission to English-medium post-secondary institutions, and a general training form designed for people wanting to study in secondary school education settings, work, or take part in training in English speaking countries. All examinees take the same Listening and Speaking subsections but specialized academic or general training Reading and Writing subsections (Stoynoff, 2009).

The current research on IELTS and academic success or simply IELTS for academic purposes is limited in comparison to TOEFL. This may be in part due to the idea that TOEFL is the leading English language proficiency test, and is therefore more desired or respected by relevant sections of the research community based on previous studies. The consistency among subsections and academic intent makes TOEFL and IELTS ideal for comparison. There is insufficient research however comparing TOEFL and IELTS and even less research comparing the use of these tests for admission or academic purposes.

2.4.1 IELTS History

IELTS is an internationally recognized English language proficiency test accepted by over 9,000 organizations worldwide with over 2.2 million test takers in a given year (Cambridge ESOL, 2003). IELTS was first developed as a collaborative effort between the British Council, International Development Program of Australian Universities and Colleges and Cambridge English Language Assessment (Stoynoff, 2009). In the mid 1960's, the English Language Testing system (ELTS), as it was originally known, was developed to replace the English Proficiency Test Battery. ELTS was introduced in response to advances in teaching and learning theories but more importantly it was influenced by the growth in 'communicative' language

learning and 'English for specific purposes'. In 1987, there was a desire to broaden the international participation or take-up of this test and the International English Language Testing System (IELTS) was introduced. Since that time, IELTS has undergone several revisions in order to keep up with emerging educational research (Cambridge ESOL, 2003). From 1995 to 2004, the IELTS partners sponsored over 55 studies. During that time, both the Speaking and Writing subsections were revised (Cambridge ESOL, 2003). Those revisions were common during the early 2000's as educational research supported the need for strong written and verbal communication skills. Currently, IELTS consists of four subsections (Reading, Writing, Listening, and Speaking), and an overall test score and takes approximately four hours to complete (Cambridge ESOL, 2003).

2.4.2   Research: IELTS and Academic Success

In a mixed methods study of English language proficiency, as defined by TOEFL and IELTS, Hill, Storch, & Lynch (1999) compared each test to academic success as defined by GPA. This quantitative data was also supplemented by questionnaire and interview data. The findings revealed a moderately strong correlation between IELTS and GPA but a relatively weak correlation between TOEFL and GPA. Results based on qualitative data revealed that most studies identified non-linguistic factors as influencing their academic performance. Hill et al. (1999) attributed those factors to the weak correlations between test score and academic success.

In their study of the effect of background discipline of IELTS score, Celestine and Ming (1999) drew on the Malaysian education system, which required students to focus on the Arts or Sciences before attending post-secondary education. The participants' background as Arts or Science students was used to determine if background disciplines influenced IELTS score outcomes. In most cases where English proficiency skills were high, the background discipline

did not significantly influence IELTS score. When reviewing participants averages to low English proficiency levels, the background discipline did significantly influence IELTS score. Although the results were illuminating, the Malaysian education system provides a unique research site/context that is not the same as what is found in the Ontario education system and hence the study has limited relevance locally.

To examine the relationship between IELTS and GPA over the first semester of undergraduate study, Humphreys, Haugh, Fenton-Smith, Lobo, Michael, & Walkinshaw, (2012) had students write the IELTS test at the beginning and end of Semester 1. Of all four subsections, Speaking appeared to be the most improved as measured by IELTS, with little shift exhibited in Writing, and only marginal gains exhibited in Listening and Reading. Of all participants, those who began Semester 1 with a low IELTS score showed the greatest improvement. Although Listening and Reading demonstrated a strong significant correlation with GPA, the Speaking and Writing subsections did not. This strong correlation was maintained in Semester 2 but not during Semester 3.

Yen and Kuzma's (2009) study of the predictive validity of IELTS on the academic performance of Chinese students revealed a significant positive correlation between GPA and IELTS score. Low GPA was seen for students who had low IELTS scores especially when those low scores were in the Listening and Writing subsections. The Speaking score was the only subsection to not exhibit a significant relationship with student GPA, which could have been due to factors such as the structure of university and the lecture style of teaching. This analysis was completed at the end of first semester and then again after second semester. When compared, the predictive ability of IELTS was stronger when correlated to student GPA after first semester than after second semester. Those results could suggest that over time IELTS' predictive power on student academic performance may become weaker. It can be argued that the weakening of

predictive power over time may occur as other factors are able to influence the student's English ability and/or their academic success. The only subsection to demonstrate a significant relationship with GPA after second semester is the Listening subsection. Again as noted above, those results are consistent with a lecture-style teaching format which may offer consistency in presentation and talking style. Since all of the significant relationships determined in this study were moderate Yen and Kuzma (2009) suggested that other factors such as students' adaptability to a new learning system, speed of acculturation, and personal goals and ambitions may have been influencing a student's GPA.

Cotton and Conrow (1998) investigated the predictive validity of IELTS amongst international students studying at the University of Tasmania. For that study, IELTS was correlated to three measures of academic achievement including GPA, academic staff ratings of student performance, and students' self-ratings of performance. When correlated with GPA, the overall IELTS score demonstrated no significant correlation suggesting no predictive ability. Not only did the results of Cotton and Conrow's (1998) study suggested that a high IELTS score of 7.0 or above did not predict academic success, but they also found that a low score of 5.5 did not predict failure.

Feast (2002) researched the relationship between IELTS scores on entry to university and academic success as determined by the GPA of international students. This study found a significant weak positive relationship between English language proficiency as measured by IELTS, and academic performance, as measured by GPA. The study by Feat (2002) also reviewed the impact those results had on a university as an operational unit. Loss of international students numbers by increasing English language proficiency requirements is a common concern to senior administration within educational institutions. Feast (2002) suggested keeping the overall score requirement at 6.0, while implementing a minimum score of at least 6.0 in Reading

and Writing for all undergraduate students. Such an increase would result in the identification of applicants with a 0.9% greater GPA than those currently identified; however it would result in a 40% decrease in the number of eligible international students for admission. Similar to Kerstjens and Nery (2000), Feast recommended increasing the support levels for those students, as the loss via student attrition may not be beneficial to the operations of a university.

In a study of 55 IELTS students, Woodrow (2006) revealed a moderate positive correlation between IELTS and GPA. In contrast to Feast (2002), when broken down by subsection, the Speaking and Listening subsections showed the strongest relationship to GPA. Frequency tables for each band score showed very low numbers at the extremes. Therefore, although the data appeared to span a larger IELTS range (minimum overall 5.5, maximum overall 8.5) than most other studies, the low frequencies at 5.5 and 8.5 revealed otherwise. To determine the impact of increasing IELTS test scores, Woodrow divided the original IELTS group into two subgroups with the intention of capturing everyone with an overall score of 6.5 and below or 7.0 and above. Each of those groups was correlated to GPA. A relatively strong significant relationship was seen between the 6.5 and below group but no significant relationship was seen with the 7.0 and above group. These results suggest that IELTS is a better predictor of academic success for applicants who are near the minimum requirements than for those with a higher level of English proficiency. Woodrow also suggested that many other variables influence personal experience, which could impact IELTS such as professional experience, years learning English, and time in predominantly an English-speaking country.

To investigate the predictive validity of IELTS on academic performance and language difficulties of international students at an Australian university, Kerstjens and Nery (2000) conducted a study of 113 first-year international students. From that investigation, a weak to moderate correlation explained the predictive effect of IELTS on academic performance of their

Higher Education group. Within that group, Reading and Writing subsections demonstrated the strongest correlations. Therefore, Kerstjens and Nery recommended that there be special consideration for these two subsections. Based on the fact that significant correlations were only seen among the Higher Education group Kerstjens and Nery suggested that institutions invest in educational supports to help students' whose first language is not English succeed once enrolled as students.   Although increased educational support is a common recommendation among the field of international students, the study did not provide enough background on the two programs to thoroughly justify this recommendation. The University of Melbourne has create a Diagnostic English Language Assessment (DELA) which assesses students' English abilities in order to help them determine which educational supports offered by the institution could aid in their academic success ("Academic Skills", 2016). Shifting the focus away from finding an assessment to guarantee sufficient English language proficiency of applicants to assisting applicants once they are enrolled as students could ideally aid in university retention rates and student success.

Dooey and Oliver (2002) revealed that other factors, besides English language proficiency, influenced the academic success of undergraduate students in science, engineering, and business. Their study found little evidence to support any type of predictive validity between IELTS and future academic success, which confirmed some previous research findings (eg. Cotton & Conrow, 1998) and suggested that there are other contributing factors involved in academic success. Dooey and Oliver studied the relationship between IELTS scores and GPA for native and non-native English speakers. Native speakers obtained higher scores than their non-native counterparts, as would be expected. Although several native participants received very high scores (overall score 7 – 9), which would suggest that those students should be very high achieving, that was not identified. IELTS is a test designed for non-native speakers and does not predict the achievement of native speakers. Further analysis to determine the predictive validity

of IELTS by faculty yielded very small sample sizes, suggesting that the results may not be reliable.

2.4.3 IELTS Reliability

Literature related to the reliability and validity of IELTS is limited. To measure the internal consistency of the Listening and Reading subsections of IELTS Cronbach's alpha values were reported for 15 test versions (Cambridge ESOL, 2006). The average reliability across all versions was .89 and .86, respectively (Cambridge ESOL, 2006, p. 13). Those data suggest strong internal consistency. The reliability for the Writing and Speaking subsections cannot be reported using the same Cronbach's alpha values for the Listening and Reading subsections. The reason for this is due to the fact that the Writing and Speaking subsections are not item-based. Instead, the Writing and Speaking subsection are rated by trained and standardized examiners according to detailed descriptive criteria and rating scales.

2.4.4 IELTS Validity

Validity on all IELTS subsections is limited and would benefit from a confirmatory factor analysis. There is virtually no conclusive study to determine the reliability of the Listening subsection. This may be in large part due to the need for participants to be able to listen in order to complete other subsections. For example, Nakatsuhara (2011) examined the test takers' listening proficiency when performing the IELTS Speaking test. This study showed that at least in some part, listening-into-speaking ability was measured during Part 2 and Part 3 of the Speaking section.

Seedhouse, Harris, Naeb, and Ustunel (2014) studied the relationship between Speaking subsection features and their corresponding subsection scores. Although the variation among subsections has an anticipated direction, the relationship is not a clear linear progression.

Measures of speaking accuracy and fluency showed significant difference between scores of 5 and 8 in a direct proportional relationship. This same proportional relationship was not seen for all aspects of speaking, with the lowest grammatical range ability for participants being a Speaking score of 5 and the highest grammatical range ability for participants being a Speaking score of 7. For grammatical range complexity, test takers with a Speaking score of 7 performed higher than those with a score of 8 possibly suggesting non-linear progression. The study by Seedhouse et al. (2014) also performed a qualitative analysis which did not identify any single features to account for the variation among Speaking scores. Instead, those data suggested that it is a cluster of assessable speaking features within any IELTS Speaking test that ultimately contributed toward a given score.

Moore, Morton, and Price's (2011) conducted a study in Australia of 13 tests comprised of, on average, three reading samples to investigate the validity of the IELTS Academic Reading test in comparison to university study. The requirements of reading for IELTS and reading for university study were both identified by Moore et al. Although the results do not suggest a weak relationship, they do suggest that there is room for strengthening the validity of the IELTS Reading section for academic use. Both IELTS and academic reading require a basic comprehension but academic reading also requires critical evaluation of material which seems to be lacking in the IELTS Reading subsection.

The Writing subsection of IELTS is far more researched than the Listening and Reading subsections. The abundance of writing research may be due, in part, to the ability to gather concrete evidence in the form of writing samples. The essay genre of writing can be viewed as the foremost genre of university writing (Moore & Morton, 1999). Standard Task 2 of the IELTS Writing section was found to resemble the essay genre of writing more closely than any other form of university writing. Therefore Moore and Morton suggested that the current format of

IELTS Task 2 measures academic writing, as it required test takers to present a written argument or case. Moore and Morton also suggested that although this follows the essay genre, the Writing subsection could benefit from a more diverse range of rhetorical functions in order to be more representative of university writing.

2.5   Academic Fraud

For over a decade fraud and deception have become an endemic problem within education (Guhr, & Furtado, 2013). The increasing concern for fraud within education may be in part attributed to the growth of international student enrolment, the commercialization of international education, the economic value of education, advances in primary internet-based technology, and policy goals and policy mistakes (Guhr, & Furtado, 2013; McCarthy, 2015). Although the education sector is not immune from fraud, there seems to be a lack of scholarly research and evidence in support of educational or more specifically documentation fraud. Instead, such a lack of research may be due to the fact that no common framework exists to classify, collect, and analyze fraudulent information, and that companies, such as ETS and the British Council, protect their data as proprietary. Guhr and Furtado (2013) identified that fraud in education is not new, but with the growth of fraud, or the growth of the existence of fraud, it is important to address many educational questions. During their review of international fraud in education, Guhr and Furtado (2013) reported having to rely heavily on media research, online-based research, background interviews, and expertise gathered in its recruiting, admissions, and student services projects. The current study is no different as research involving educational fraud has been limited in large part by stakeholders' with little incentive to release proprietary information.

Although some may argue that the lack of proprietary data should not stop researchers from gathering their own institutional data, directly measuring English language proficiency test

fraud is nearly impossible. Guhr and Furtado (2013) attribute the inability to measure fraud to a

number of factors. Firstly, Guhr and Furtago (2013) discuss the misclassification of data. Since

only detected fraud can be classified and instances of fraud may be undetected, it is possible for

fraud to be misclassified or not classified at all. Alternatively, instances of fraud which are

detected may not properly be classified as the method of fraud is not always identifiable. For

example, upon receiving confirmation of a formal investigation by ETS or the British Council,

institutions are told that the test score has been cancelled however no reason, not even the

confirmation of fraud, is provided to the institution. Classifying the test score as fraudulent, or

going further to identifying the type of fraud (e.g. impersonation or plagiarism), is difficult as no

reliable data exists to guide this classification, therefore falsifying its use during any future

studies. The second factor Guhr and Furtado (2013) identified was the issue of reporting

occurrences of fraud as not all instances of fraud are known, measurable or identifiable in a

reportable format. Lastly, Guhr and Furtado (2013) attribute the inability to measure fraud to

data availability. Even when instances of fraud are confirmed and identified, those data are

typically not made public by stakeholders (e.g. ETS or the British Council) who own such data.

For instance, a Chinese student who wrote an IELTS test on July 25th in the Shandong Province

was informed by the British Council on August 11 that her "test results [would] not be released

and a test report [would] not be issued" ("British Council reportedly withholds", 2015). When

asked for a reason, this student was told that she "breach[ed] IELTS test rules and regulation" and

was not given any further explanation.

Universities are responsible for performing due diligence and staying alert to the

possibility of fraudulent and misleading student documentation. Ethically, admitting applicants

based on fraudulent documentation, while denying admission to applicants who have integrity

and who have submitted authentic documentation is essentially rewarding fraud. Ezell's (2008) book entitled *Counterfeit Diplomas and Transcripts* outlined the countless number of vendors who sell transcripts and educational credentials from real educational institutions around the world bringing light to the comprehensive yet ambiguous and seemingly invisible battle at times that is academic fraud. English language proficiency tests, similar to transcripts and degrees, are not immune to the act of fraud. Koenig and Devlin (2012) discuss educational credentials as a commodity in the world market. In order to gain a degree and be eligible to trade within this market, applicants must first gain access to educational institutions; it is this access which is the driving force for educational fraud (Koenig & Devlin, 2012).

The issue surrounding documentation fraud, including fraudulent English proficiency test scores goes beyond the consequences of having an offer of admission revoked, but could result in the deportation or removal of a student study visa. In order to apply for a Canadian student visa, applicants must submit their letter of acceptance (Citizenship and Immigration Canada, 2016). This letter of acceptance is issued by each educational institution to applicants who have been admitted and is in a format set out by the Government of Canada. As offers of admission are condition, applicants must meet all admission conditions set out by that university including the English language proficiency requirements in order to retain the validity their letter of acceptance. Therefore, submitting a fraudulent document during the admission process can invalidate the student visa already issued, or that is in the process of being issued.

In order to be alert and responsible to academic fraud, including fraudulent TOEFL iBT and IELTS test scores, educational institutions must instil a methodical and consistent approach to document review (Koenig, & Devlin, 2012). As part of this approach, institutions must develop appropriate training for their employees which allows for consistency amongst reviews

while also developing a method of communication with applicants suspected of fraud. Before establishing an effective approach, institutions must understand the how they use their documents.

2.5.1   Case of Academic Fraud

In a case review by the New York Times a Chinese applicant discussed her experience applying to American universities ("The China Conundrum*"*, 2011). The use of an agent has become very common for Chinese applicants. For this student, the agent helped her navigate the otherwise bewildering application process and provided opportunity where possible. As this applicant's English was not sufficient to write a strong admission essay, staff members at the agency would charge approximately $4,000 to ask her questions in Chinese and write a cohesive essay in English. If applicants were interested in test preparation, this could also be arranged for an additional fee of $3,300. In the same report, the New York Times identified another case which identified agents falsifying school letterhead in order to produce doctored transcripts and counterfeit letter of recommendation ("The China Conundrum", 2011). This issue was only discovered when a parent complained to the school that their child was being charged for education documentation. Little did they know that the promise of their child's agent to get them admitted to an American university came with an ethical price tag.

In Ontario, several media reports have identified acts of fraud in higher education. For example, in 2014 London Ontario police charged three people with forgery and personation in an exam-sitting scam which charged clients up to $7,000 to have someone write their English proficiency test for them ("London police charge", 2014). Similarly, a woman from China was sentenced to two years for running a scam which had imposters write English proficiency tests

for students applying for post secondary education using fraudulent Chinese passports ("Chinese women gets two years", 2013).

The concern for English language proficiency test fraud is echoed in Australia where vice-chancellor Jeanette Hacket of Curtin University discussed the possibility of removing IELTS from their list of acceptable tests due to the fact that IELTS may be "too risky" ("Curtin University", 2011; "Tough Sentences", 2011). In additional to impersonation, in 2009 and 2010 Curtin University found nine staff members guilty of falsifying IELTS test scores for up to AUD 11,000. These staff members were prosecuted and received penalties ranging from an AUD 20,000 fine to two years jail time ("Tough Sentences", 2011). Although Curtin University raised this concern for the use of IELTS, they subsequently continued to use of IELTS to meet English proficiency requirements.

Although fraud can be seen in many forms, impersonation is a common concern. In the article *Ghost in the Machine* (2011) several Chinese companies advertising "ghost writing" and impersonation services for students wishing to pursue an education in an English-speaking country. During this process, clients were able to obtain IELTS scores high enough to meet English proficiency requirements. In the event that students were caught cheating, they were banned from re-taking IELTS in the country where the test was written. To counteract this, impersonators would travel to neighbouring countries to re-take this test until a desired and accepted score was achieved. Similar accounts of fraud were identified in India where arrests were made for impersonating candidate during the IELTS exam process by using fake passports (Guhr, & Furtado, 2013).

In response to the concern surrounding English language proficiency test fraud, some universities have begun to develop new methods for assessing applicants who have been required to submit proof of English language proficiency. For example, York University has developed an

independent English language proficiency test which for use during the admission process in addition to any external English language proficiency tests such as TOEFL iBT or IELTS (York University, 2016). In order to put forward recommendations for the future use of TOEFL iBT and IELTS, this study aimed to provide an understanding for the current use of those standardized English language tests by understanding the relationship between TOEFL iBT and IELTS with academic success as measured by first-year GPA, and prior English ability, as measured by ENG4U grade.

Chapter 3

3   Methods & Procedures

Based on the fact that the value or potential usefulness of TOEFL iBT and IELTS with respect to OUAC 101 applicants is not entirely clear this exploratory study aimed to determine if either the TOEFL iBT or IELTS are appropriate for use during the admission of OUAC 101 applicants. Additionally this study was intended to propose a method for identifying potentially fraudulent TOEFL iBT or IELTS test scores. In both cases, correlation and regression analyses were conducted using OUAC 101-specific admission criteria. Analyses were performed using IBM's SPSSv.23 software (Cleophas & Zwinderman, 2010).

3.1   Participants/Applicants

The sample consisted of 2,801 applicants to one Ontario University who applied during the 1129 (2012 – 2013), 1139 (2013 – 2014), and 1149 (2014 – 2015) admission cycles using the OUAC 101 application and who submitted either TOEFL iBT or IELTS test scores in support of their application. Those specific admission cycles were selected due to the similarity in the admission requirements in those years. Applicants ranged in age from 16 to 22. The OUAC 101 applicant pool included applicants who are currently studying at an Ontario curriculum-based school (i.e. those studying either in Ontario or abroad at a high school that uses the Ontario curriculum). Using OUAC 101 applicants in this study allowed for the working assumption that applicants have a common learning experience to the extent that they have all been taught the Ontario curriculum, at least at the grade 12 level. This assumption was not one that could be made for any other applicant pool. Similarly, those data are unique to applicants applying to this Ontario university and are not intended to represent all Ontario high school applicants, therefore

the results should not be generalized. The Ontario university used within this study requires that any applicant—domestic or international—whose first language is not English and who studied for fewer than three years in an Ontario high school submit proof of English language proficiency (University X, as in References, 2015). Studying three years in an Ontario high school or meeting the minimum test score (see Table 1), is a requirement for all undergraduate programs at this Ontario university except the Nursing program. Therefore, applicants applying to this university's Nursing program were omitted from the sample in this study.

Use of university-owned application data for the present study was approved by the Vice-Provost of Academic Affairs and Vice-Provost of International at on Ontario university. To preserve the privacy of participants' personal information and the integrity of the data as noted earlier (the researcher is a current member of Admission Department at this Ontario university), the data were collected by the Statistical Analyst and Reporting Consultant of the Office of the Registrar at one Ontario university before being provided to the researcher. As this university holds the rights to the data collected, contact with and compensation of participants' for data use, and the obtaining of participants' informed consent, was not required.

3.2   Measures

3.2.1   Demographics

The data were amalgamated into four master lists: the TOEFL iBT applicant list, the TOEFL IBT registrant list, the IELTS applicant list, and the IELTS registrant list, based on the test submitted and participant application status. The applicant lists included any participant who has applied to this Ontario university using the OUAC 101 application and submitted either TOEFL iBT or IELTS as part of their application process, whereas the registrant list included only OUAC 101 applicants who submitted either a TOEFL iBT or IELTS test score, was

admitted to this Ontario university, accepted their offer of admission, and subsequently

completed their first year of undergraduate education. Each list included applicants' term of

admittance to the university (1129, 1139, or 1149), application status, program(s) applied to, high

school admission averages, TOEFL iBT or IELTS score by subsection (Reading, Writing,

Speaking, Listening, Overall), grade 12 university-level English (ENG4U) final grade, first-year

cumulative average (GPA), and first-year program. It is important to note that a GPA and first-

year program were identified for registrants only, as only those participants who entered a

program and subsequently completed their first year of undergraduate study at the Ontario

university used within this study were able to achieve a first-year GPA.

3.2.2   Student Number

Student numbers were collected solely to track data from applicants at multiple time

points; the statistical consultant ensured that the researcher had no list connecting student

numbers to personal information.

3.2.3   English Language Proficiency Tests

When detecting fraud among English language proficiency test scores, it is deemed nearly

impossible to identify all fraudulent test scores. Although suspected fraudulent test scores were

removed from this study, it cannot be guaranteed that all fraudulent test scores were identified.

Since the results of the first TOEFL iBT or IELTS test score submitted by each student were used

in this study, and not all participants met the minimum discretionary range requirement on their

first attempt, the range of data used in this study are wider than would be expected. Using a

range of data outside those who are eligible for admission is intended to help reduce the range

restriction which is often a limiting factor of English language proficiency studies.

3.2.4 Test of English as a Foreign Language (TOEFL)

As previously mentioned, TOEFL iBT is a product of the Educational Testing Service and was developed by test developers who possessed both advanced degrees in fields such as English, language education and linguistics, and a wide array of career backgrounds (Chapelle, Jamieson, & Enright, 2008). The TOEFL iBT test consists of four scored subsections: Reading, Writing, Listening, Speaking, and an Overall score. Factor analysis studies support a four first-order factor solution with a single higher-order factor, suggesting that all subsections are valid and load onto a single higher-order factor (e.g. Stricker& Rock, 2008). Scaled scores on each subsection range from 0–30 with the summative overall score ranging from 0–120. Although the overall score will be used for all correlations, each of the four subsections will be included in the multiple regression analyses to answer the fifth and six research questions. As the primary instrument for making admission decisions (Jamieson, Jones, Kirsch, Mosenthal, & Taylor, 2000), TOEFL iBT was selected as one of the main English proficiency test used in this study.

3.2.5 International English Language Testing System (IELTS)

As indicated earlier, IELTS was first developed as an outcome of a collaborative effort between the British Council, International Development Program of Australian Universities and Colleges, and Cambridge English Language Assessment (Cambridge ESOL, 2003). The IELTS is an internationally recognized English language proficiency test accepted by over 9,000 organizations worldwide with over 2.2 million tests typically being taken in a given year (IELTS Canada, 2016). Despite the TOEFL's rank in the literature as the leading English language proficiency test (Zareva, 2005), the Admission Department (where the researcher is currently employed) observed that the number of IELTS scores submitted by students is rapidly increasing. The increased number of IELTS test scores submitted is representative of applicants' choice to

write the IELTS test and is independent of universities confidence in the reliability of IELTS. Accordingly, the IELTS was used in this study as a comparative measure to the TOEFL iBT. Like the TOEFL iBT, the IELTS also consists of four subsections: Reading, Writing, Listening, Speaking, and an Overall section. The subsections, including the overall score, range from 0–9 using whole and half point scores. The IELTS score uses a scaled scores on each subscale including the overall score rather than having the overall score being the sum of all subsections as in the case of the TOEFL iBT. Though no factor analysis has been performed to determine the model 'fit' of all IELTS subsections together, individual studies have reviewed the construct validity of each subsection independently, with positive relations between each factor and English language proficiency for academic purposes (see Stoynoff's 2009 meta-analysis). However, Stoynoff (2009) cautioned that there were mediating factors in these relationships.

3.2.6   Grade Point Average (GPA)

This study used the term GPA to represent applicants' academic success. For the purpose of this study, GPA was not measured using the traditional GPA scale, but was measured by first-year cumulative average.

3.2.7   Grade 12 University Level English (ENG4U)

All OUAC 101 applicants were required to complete ENG4U in order to obtain their Ontario Secondary School Diploma and as a prerequisite for all first-entry undergraduate programs at one Ontario university. For this reason, ENG4U was selected as a measure of common prior English ability.

3.3   Procedure

The Statistical Analyst and Reporting Consultant at the Ontario university used within this study ran queries to create the four above mentioned master participant lists based on the type of test submitted, (TOEFL iBT or IELTS) and application status (applicant or registrant). That query included all Ontario high school applicants to this Ontario university during the 1129, 1139, and 1149 admissions cycles who have submitted TOEFL iBT or IELTS results. Applicants were able to submit multiple English proficiency test scores or ENG4U grades to this Ontario university for review. As noted earlier, it is the policy at this Ontario university to use the most recent English proficiency test score and highest ENG4U grade received. For the purpose of this study, the first English proficiency test and ENG4U grade attempt were used.

3.4   Analysis Plan

Before any statistical analyses were conducted, the data were cleaned. As applicants were able to submit multiple English language proficiency test scores, and were able to retake ENG4U, the first attempt only—successful or not—was used (hence only applicants' initial/first-attempt TOEFL iBT, IELTS, and ENG4U score submissions were used for the present analyses), leaving only one set of scores per applicant to consider in this study.

Four Pearson correlations (Overall TOEFL iBT and IELTS scores with academic success [GPA] and with prior English ability [ENG4U grade]) were run to answer Research Questions 1.1 and 2.1, as the linear relationship between functionally continuous variables is a fundamental indicator of the predictive validity of one variable for the other.

To determine the predictive validity of TOEFL iBT and IELTS at the two minimum requirements on academic success (Research Questions 1.2 and 2.2), the original sample used in Research Questions 1.1 and 2.1 were divided, based on their ability to meet the 'minimum score'

and 'minimum discretionary range' requirements. Pearson correlations between TOEFL iBT

Overall score and GPA as well as IELTS Overall score and GPA were conducted for both the

'minimum score' and 'minimum discretionary range' registrant subsamples (Research Question

1.2). Similarly, Pears on correlations between TOEFL iBT Overall score and ENG4U grade as

well as IELTS Overall score and ENG4U grade were conducted for both the 'minimum score'

and 'minimum discretionary range' applicant subsamples (Research Question 2.2).

Multiple regression analyses were conducted to answer Research Questions 1.3 and 2.3.

Two simultaneous-entry regression models were tested using all participant data for each TOEFL

iBT subsections' as the predictors of GPA or IELTS subsections' as the predictors of GPA

(Research Question 1.3). Partial correlations were used to determine the unique influences of

each predictor variable (TOEFL iBT or IELTS subsection) on GPA after removing the affect of

each other predictor for both the TOEFL iBT and IELTS models. Adjusted R^2 indicated amount

of variance in GPA explained by each model's set of predictors (Reading, Writing, Listening, and

Speaking) for the TOEFL iBT and IELTS, respectively. Unstandardized B and standardized β

values indicated the effect of each predictor on GPA within each respective model. The same

regression strategy was used using participant data only that met either the 'minimum score' or

the 'minimum discretionary range' requirements for each TOEFL iBT or IELTS.

Again the same regression strategy used to answer Research Question 1.3 was employed

for Question 2.3, but with prior English ability (ENG4U grade) as the criterion of the models

rather than GPA. Similar to the previous GPA analysis, this analysis completed six

simultaneous-entry regression models in total to determine the predictive relationship between

TOEFL iBT or IELTS test scores for three subsample groups 1) all participants, 2) those

participants meeting the 'minimum score' requirement, and 3) those participants meeting the 'minimum discretionary range' requirement, with ENG4U grade.

To account for Type I error rate inflation (due to running multiple analyses), a statistical p-value of 0.01 was used for all analyses.

Independent sample t-test analyses were conducted to compare the mean GPA of five comparison groups to address Research Question 1.4: 1) all TOEFL iBT with all IELTS participants, 2) TOEFL iBT participants meeting the 'minimum score' requirement with IELTS participants meeting the 'minimum score' requirement, 3) TOEFL iBT participants meeting the 'minimum discretionary range' requirement with IELTS participants meeting the 'minimum discretionary range' requirement, 4) TOEFL iBT participants meeting the 'minimum score' requirement with TOEFL iBT participants meeting the 'minimum discretionary range' requirement, and 5) IELTS participants meeting the 'minimum score' requirement with IELTS participants meeting the 'minimum discretionary range' requirement.

Chapter 4

4   Results

4.1   Descriptive Statistics

The sample size for each TOEFL iBT and IELTS subsample group was reviewed post-hoc in order to analyze the number of applicants within each group. As well, that review was used to determine how changes in the minimum English language proficiency requirement, from the 'minimum discretionary range' requirement to the 'minimum score' requirement, could impact the number of OUAC 101 applicants applying to this Ontario university with acceptable TOEFL iBT or IELTS test score (see Table 2 for the breakdown of sample size used among the twelve subsample groups). Of all the participants included in this study who submitted TOEFL iBT scores, 50.46% met the 'minimum discretionary range' requirement and 37.27% met the 'minimum score' requirement on their first attempt. When the IELTS requirement subsample groups were reviewed, 55.10% of participants met the 'minimum discretionary range' requirement on their first attempt, whereas only 13.23% of participants met the 'minimum score' requirement on their first attempt.

Table 2

Sample Size Breakdown of Subsample Groups for OUAC 101 Applicants and Registrants of one Ontario University Who Submitted Either TOEFL iBT or IELTS as Part of Their Application for Undergraduate Admissions.

	Applicant Sample Size		Registrant Sample Size	
Participant Group	TOEFL iBT	IELTS	TOEFL iBT	IELTS
All participants	652	2147	56	144
Minimum discretionary range	329	1183	51	132
Minimum score	243	284	37	27

4.2 Academic Success

4.2.1 Pearson Correlation

Pearson correlation analyses were run to determine if the TOEFL iBT or IELTS Overall test scores were related to academic success, as measured by first-year GPA (Research Question 1.1). Results showed that the TOEFL iBT Overall score ($M = 98$, $SD = 8.62$) was not correlated with first-year GPA ($M = 75.68$, $SD = 11.01$), $r(54) = .25$, $p = .059$ (see Figure 1). Similarly, no significant relationship was observed between the IELTS Overall score ($M = 6.75$, $SD = 0.45$) and first-year GPA ($M = 73.66$, $SD = 9.92$), $r(142) = .12$, $p = .166$ (see Figure 2).
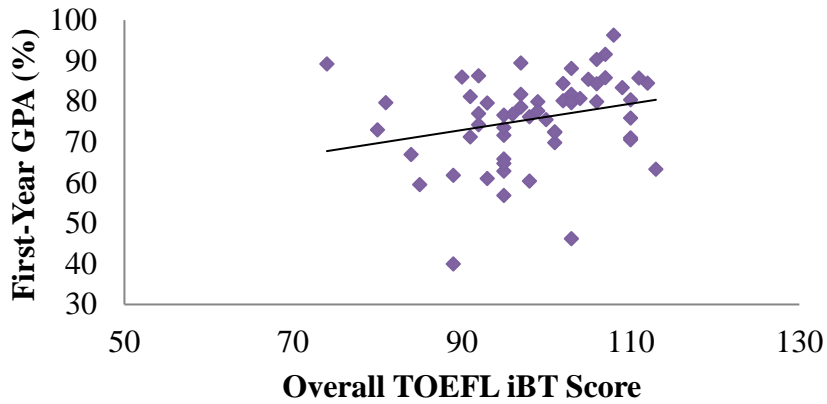
Figure 1. Correlation between Overall TOEFL iBT score for all participants and academic success, as defined by first-year GPA.
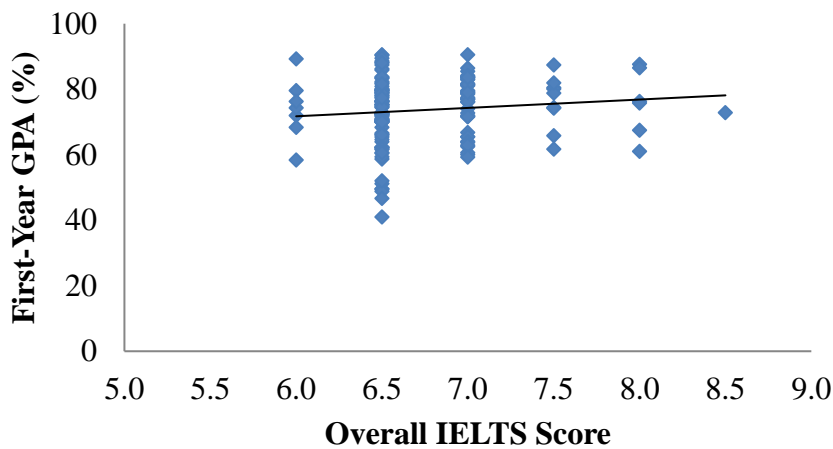
Figure 2. Correlation between Overall IELTS score for all participants and academic success, as defined by first-year GPA.

Additional Pearson correlations were conducted on subsamples of participants who met 1) the TOEFL iBT or IELTS 'minimum score' requirement and 2) the TOEFL iBT or IELTS 'minimum discretionary range' requirement in order to determine if either requirement was related to academic success (Research Question 1.2). Participants in the 'minimum score' subsample groups were also included in the 'minimum discretionary range' group as those

participants met both requirements. The inclusion of participants meeting the 'minimum score' requirement within the 'minimum discretionary range' groups was in line with admission procedures of the Ontario university used within this study. Analyses revealed that the Overall TOEFL iBT scores of participants meeting the 'minimum score' requirement ($M = 102.51$, $SD = 6.14$) were not related to first-year GPA ($M = 77.53$, $SD = 10.16$), $r (35) = .25$, $p = .134$ (see Figure 3). Conversely, the Overall TOEFL iBT scores of participants meeting the 'minimum discretionary range' requirement ($M = 99.88$, $SD = 7.31$) were moderately correlated with first-year GPA ($M = 75.50$, $SD = 11.23$), $r (49) = .41$, $p = .003$ (see Figure 4).
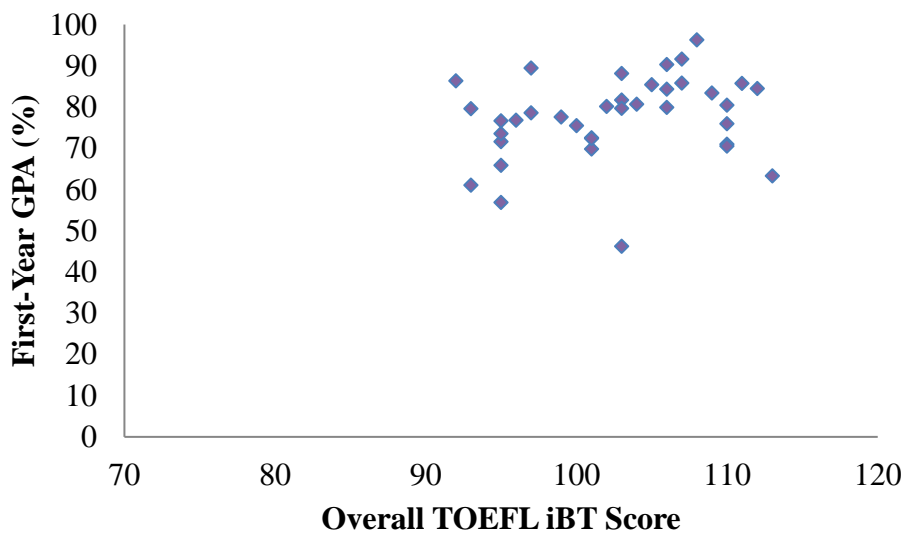
Figure 3. Correlation between Overall TOEFL iBT scores meeting the 'minimum score' requirement and first-year GPA.
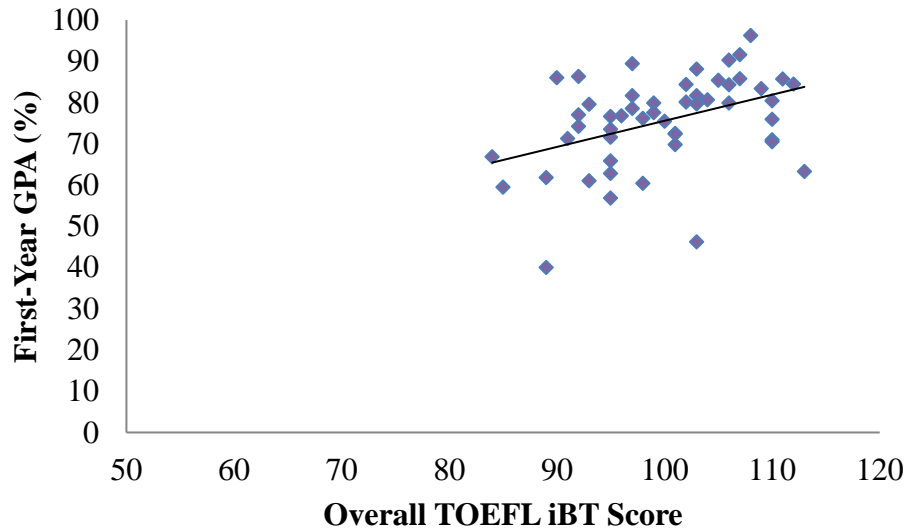
Figure 4. Correlation between Overall TOEFL iBT score and academic success, as measured by first-year GPA, of participants meeting the 'minimum discretionary range' requirement.

Similar results were seen with the IELTS participant groups as the Overall IELTS scores of participants meeting the 'minimum score' requirement ($M = 7.41$, $SD = 0.46$) were unrelated to first-year GPA ($M = 74.56$, $SD = 8.61$), $r(25) = .06$, $p = .770$ (see Figure 5), while the Overall IELTS scores of participants meeting the 'minimum discretionary range' requirement ($M = 6.65$, $SD = 0.25$) were correlated with first-year GPA ($M = 73.34$, $SD = 10.32$), $r(130) = .19$, $p = .047$ (see Figure 6).
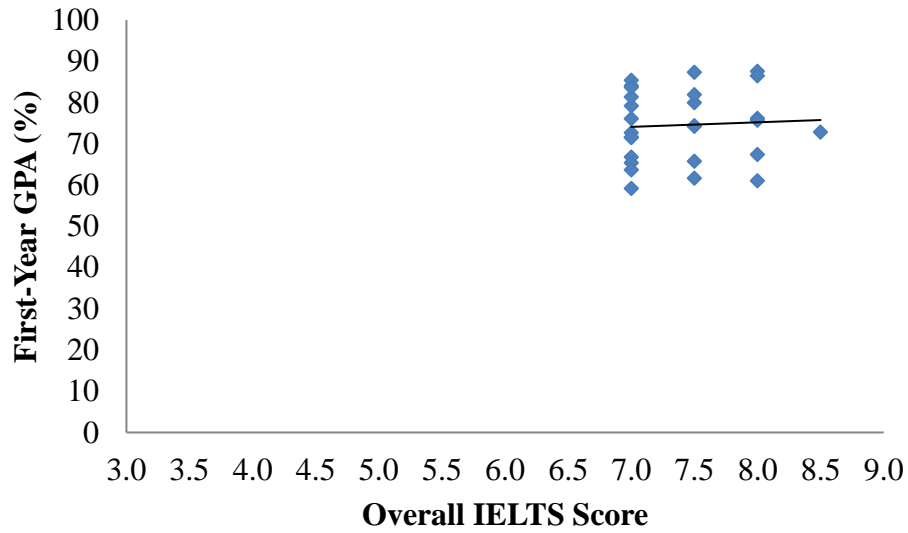
Figure 5. Correlation between Overall IELTS score and academic success, as measured by first-year GPA, of participants meeting the 'minimum score' requirement.
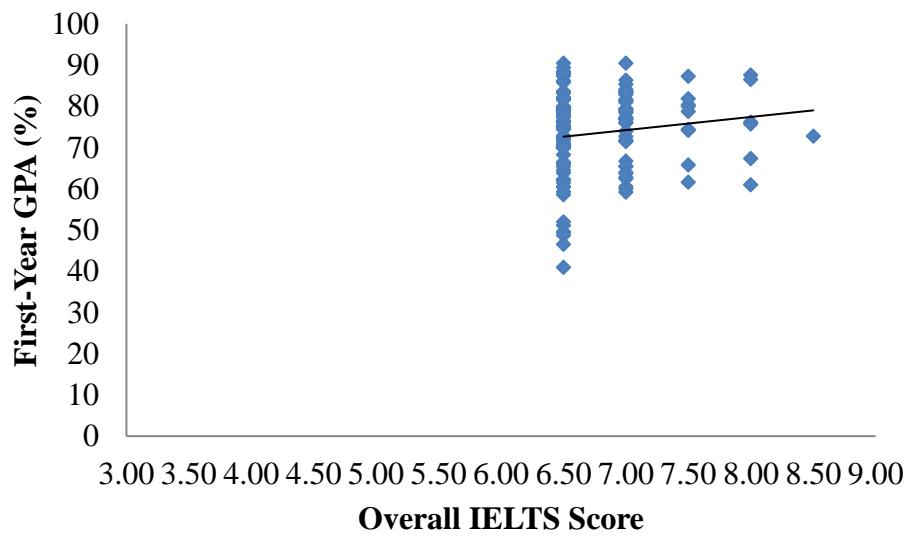
Figure 6. Correlation between Overall IELTS score and academic success, as measured by first-year GPA, of participants meeting the 'minimum discretionary range' requirement

4.2.2   Multiple Regression: All Participants

In addition to using the TOEFL iBT or IELTS Overall score, multiple regression analyses were conducted using the four subsections of the TOEFL iBT or IELTS (Listening, Reading, Writing, and Speaking) to predict first-year GPA. For those analyses, GPA was correlated with the TOEFL iBT or IELTS score of three subsample groups 1) all TOEFL iBT or IELTS participants, 2) TOEFL iBT or IELTS participants meeting the 'minimum score' requirement and 3) TOEFL iBT or IELTS participants meeting the 'minimum discretionary range' requirement (Research Question 1.3). Simultaneous entry of all participants' scores on each TOEFL iBT subsection, Listening ($M = 24.79$, $SD = 3.07$), Reading ($M = 24.63$, $SD = 3.18$), Writing ($M = 25.05$, $SD = 2.48$), and Speaking ($M = 23.89$, $SD = 3.06$), into an SSTYPE3 (or Model I) multiple regression analysis was used to determined if the four subsections together significantly predicted first-year GPA ($M = 75.68$, $SD = 11.01$). The model was not significant ($F\,(4,51) = 2.82$, $p = .035$), and the unadjusted R^2 indicated that the four subsections accounted for 18.1% of the variance in GPA (unbiased adjusted R^2 indicated 11.7% of explained GPA variance). Listening, Reading, Writing, and Speaking scores also did not individually contribute to the model equation's prediction of GPA (see Table 3; B/β = 0.85/.24, $t(51) = 1.63$, $p = .110$; B/β = 0.86/.25, $t(51) = 1.70$, $p = .094$; B/β = 0.68/.15, $t(51) = 1.04$, $p = .303$; and B/β = -0.97/-.27, $t(51) = -1.79$, $p = .080$, respectively). The inability of TOEFL iBT subsections to predict first-year GPA was supported by Listening, Reading, Writing, and Speaking's nonsignificant partial correlations with first-year GPA ($r(51) = .22$, $p = .110$; $r(51) = .23$, $p = .094$; $r(51) = .14$, $p = .303$; $r(51) = -.24$, $p = .080$; and $r(51) = -.18$, $p = .204$, respectively). Therefore, none of the TOEFL iBT subsections effectively predicted academic success as defined by first-year GPA for the all participant group.

Table 3

Multiple Regression Values of TOEFL iBT Subsection Scores on First-Year GPA for All

Participants.

	Unstandardized B	*Standardized β*	*t*	*p*	*Partial*
Intercept	39.71		2.36	.022	
LISTN	.845	.235	1.63	.11	.222
READ	.857	.248	1.70	.094	.232
WRIT	.681	.154	1.04	.303	.144
SPEAK	-.968	-.269	-1.79	.080	-.243

$R^2 = .181$
Significant at $p < .01$

The sample multiple regression analyses were then completed using the scores from all participants who submitted an IELTS test scores. Simultaneous entry of all IELTS participants' scores on each subsection, Listening (LISTN; $M = 6.89$, $SD = .81$), Reading (READ; $M = 6.82$, $SD = .73$), Writing (WRIT; $M = 6.34$, $SD = .43$), and Speaking (SPEAK; $M = 6.46$, $SD = .62$) into an SSTYPE3 (Model I) multiple regression analysis revealed no significant model with first-year GPA ($M = 73.66$,$SD = 9.92$) using the four IELTS subsections, $F(4,139) = 0.774$, $p = .544$. Nonetheless the unadjusted R^2 implied that the predictor variables (LISTN, READ, WRIT, SPEAK) accounted for 2.2% ($R = .022$) of the variance in first-year GPA (adjusted $R^2= -.006$). As seen in Table 4, Listening, Reading, Writing, and Speaking scores did not contribute to the model equation significantly (B/β = 1.28/.10, $t(139) = 1.06$, $p = .289$; B/β = 1.00/.07, $t(139) = $ .78, $p = .435$; B/β = 0.62/.03, $t(139) = .31$, $p = .754$; and B/β = -0.40/-.03, $t(139) = -.26$, $p = .796$, respectively). Moreover, IELTS Listening, Reading, Writing, and Speaking were not related to first-year GPA after considering the other subsections (partial correlations of r (139) = .09, $p = $ .289; r (139) = .07, $p = .435$ r (139) = .03, $p = .754$; and r (139) = -.02, $p = .796$, respectively).

Table 4

Multiple Regression Values of IELTS Subsection Scores, of All Participants, on First-Year GPA.

	Unstandardized B	*Standardized β*	*T*	*p*	*Partial*
Intercept	56.61		3.849	.000	
LISTN	1.28	.104	1.06	.289	.090
READ	1.00	.074	.783	.435	.066
WRIT	.621	.027	.314	.754	.027
SPEAK	-.399	-.025	-.260	.796	-.022

$R^2 = .148$

Significant at $p < 0.01$

4.2.3   Multiple Regression: Minimum Score Group

Following this analysis, similar post hoc multiple regression analyses were performed using only scores from participants who met the 'minimum score' requirement (Research Question 1.3). The four subsections of TOEFL iBT, Listening ($M = 25.70$, $SD=2.68$), Reading ($M = 25.95$, $SD = 2.19$), Writing ($M = 25.70$, $SD = 2.21$), and Speaking ($M = 24.97$, $SD = 2.32$) which met the 'minimum score' requirement were correlated with first-year GPA ($M = 77.53$, $SD = 10.16$). However the model proved to be nonsignificant, $F (4,32) =2.38$, $p = .073$, unadjusted $R^2= .229$, adjusted $R^2 =.133$. None of the subsections (Listening, Reading, Writing, Speaking) predicted first-year GPA (B/β = 0.85/.22, $t(32) = 1.31$, $p = .20$; B/β =1.27/.27, $t(32) = 1.43$, $p = .164$; B/β = 0.41/.09, $t(32) = .45$, $p = .657$; and B/β = -0.97/-.22, $t(32) = -1.22$, $p = .232$, respectively; see Table 5). This was further supported by the subsections' nonsignificant partial correlations with first-year GPA ($r(32) = .23$, $p = .200$; $r(32) = .24$, $p = .164$; $r(32) = .08$, $p = .657$; and $r(32) = -.21$, $p = .232$ for Listening, Reading, Writing, Speaking, respectively).

Table 5

Multiple Regression Values of TOEFL iBT Subsection Scores, of Participants Meeting the

Minimum Score Requirement, on First-Year GPA.

	Unstandardized B	*Standardized β*	*T*	*p*	*Partial*
Intercept	36.40		1.34	.189	
LISTN	.847	.223	1.31	.200	.225
READ	1.27	.273	1.43	.164	.244
WR	.411	.089	.448	.657	.109
SPEAK	-.968	-.221	-1.22	.232	-.210

$R^2 = .229$

Significant at $p < 0.01$

Subsequent multiple regression analysis was completed using IELTS participants meeting the 'minimum score' requirement. When IELTS subsection, Listening (LISTN; $M = 7.57$, SD =.874), Reading (READ; $M = 7.50$, $SD = .844$), Writing (WRIT; $M = 6.81$, $SD = .371$), and Speaking (SPEAK; $M= 7.19$, $SD = .695$) were modeled against first-year GPA ($M =74.56$, $SD = 8.61$) for participants meeting the IELTS 'minimum score' requirement, the model was nonsignificant, $F(4,22) =1.06$, $p = .400$. Although the unadjusted R^2 implied that the IELTS subsections accounted for 16.2% of the first-year GPA variance (adjusted $R^2 = .009$), none of the IELTS subsections (Listening, Reading, Writing, Speaking) added to the prediction of first-year GPA (B/β = 3.10/.31, $t(22) = 1.13$, $p = .269$; B/β = 1.57/.15, $t(22) = .66$, $p = .516$; B/β = -6.56/-.28, $t(22) = -1.23$, $p = .233$; and B/β = -5.19/-.42, $t(22) = -1.78$, $p = .090$, respectively; see Table 6). Moreover, after accounting for the other IELTS subsections, none of the subsections were correlated with first-year GPA (partial correlations: $r(22) = .24$, $p = .269$; $r(22) = .14$, $p = .516$; $r(22) = -.25$, $p = .233$; and $r(22) = -.35$, $p = .090$ for Listening, Reading, Writing, and Speaking, respectively).

Table 6

Multiple Regression Values of IELTS Subsection Scores, of Participants Meeting the 'Minimum Score' Requirement, on First-Year GPA.

	Unstandardized B	*Standardized β*	*T*	*p*	*Partial*
Intercept	121.33		2.92	.008	
LISTN	3.10	.315	1.13	.269	.235
READ	1.57	.153	.660	.516	.139
WRIT	-6.56	-.283	-1.23	.233	-.253
SPEAK	-5.19	-.419	-1.78	.090	-.354

$R^2 = .162$
Significant at p < .01

4.2.4   Multiple Regression: Minimum Discretionary Range Group

A similar post hoc simultaneous entry of participants' scores who met the TOEFL iBT 'minimum discretionary range' requirement, was analyzed to determine if there was a significant model for predicting first-year GPA ($M = 75.50$,$SD = 11.23$) using the TOEFL iBT subsections, Listening (LISTN; $M = 25.02$, $SD = 2.77$), Reading (READ; $M = 25.20$, $SD = 2.65$), Writing (WRIT; $M = 25.25$, $SD = 2.45$), and Speaking (SPEAK; $M = 24.27$, $SD = 2.69$) (Research Question 1.3). Results indicated the model was significant, $F(4,46) = 4.58$, $p = .003$. Although the unadjusted R^2 implied that the predictor variables accounted for 28.5% (R = .285) of the variance in first-year GPA, (unbiased adjusted $R^2 = .223$), none of the subsections (Listening, Reading, Writing, Speaking) influenced the model significantly (B/β = 1.19/.29, $t(46) = 2.08$, $p = .043$; B/β = 1.38/.33, $t(46) = 2.33$, $p = .024$; B/β = 0.376/.08, $t(46) = .58$, $p = .566$; and B/β = -0.38/-.09, $t(46) = -.65$, $p = .518$, respectively; see Table 7). This outcome was also supported by nonsignificant partial correlations of Listening, Reading, Writing, Speaking with GPA ($r(46) =$

.29, *p* = .043; *r*(46) = .33, *p* = .024; *r*(46) = .09, *p* = .566; and *r*(46) = -.10, *p* = .518,

respectively).

Table 7

Multiple Regression Values of TOEFL iBT Subsection Scores, of Participants Meeting the

'Minimum Discretionary Range' Requirement, on First-Year GPA.

	Unstandardized B	Standardized β	t	p	Partial
Intercept	10.77		.549	.586	
LISTN	1.19	.293	2.08	.043	.293
READ	1.38	.326	2.33	.024	.325
WRIT	.376	.082	.578	.566	.085
SPEAK	-.382	-.092	-.652	.518	-.096

R^2 = .285
Significant at p < 0.01

Again simultaneous entry of participants' scores which met the IELTS 'minimum

discretionary range' on each subsection of Listening (*M* = 6.95, *SD* = .795), Reading (*M* = 6.86,

SD = .721), Writing (*M* = 6.36, *SD* = .410), and Speaking (*M* = 6.52, *SD* = .602) were correlated

with first-year GPA (*M* = 73.59, *SD* = 9.98). This multiple regression analysis also revealed a

nonsignificant GPA prediction model using the four IELTS subsections, *F*(4,128) = 1.01, *p* =

.407. The unadjusted R^2 implied that the IELTS test scores accounted for 3.1% of the first-year

GPA variance (adjusted R^2 = 0). None of the four subsections (Listening, Reading, Writing,

Speaking) were found to add significantly to the model's prediction of first-year GPA (B/β =

1.66/.13, *t*(128) = 1.32, *p* = .190; B/β = 1.02/.07, *t*(128) = .75, *p* = .456; B/β = 1.20/.05, *t*(128) =

.56, *p* = .580; and B/β = -.79/-.05, *t*(128) = -.49, *p* = .624, respectively; see Table 8). Further,

Listening, Reading, Writing, Speaking were not related to GPA after considering the other

subsections' effects (partial correlations: $r(128) = .12$, $p = .190$; $r(128) = .07$, $p = .456$; $r(128) = .05$, $p = .580$; and $r(128) = -.04$, $p = .624$, respectively).

Table 8

Multiple Regression Values of IELTS Subsection Scores, of Participants Meeting the 'Minimum Discretionary Range' Requirements, on First-Year GPA.

	Unstandardized B	*Standardized β*	*t*	*p*	*Partial*
Intercept	52.57		3.25	.001	
LISTN	1.66	.132	1.32	.190	.116
READ	1.02	.074	.747	.456	.066
WRIT	1.20	.049	.555	.580	.049
SPEAK	-.789	-.048	-.491	.624	-.044

$R^2 = .031$
Significant at $p < .01$

4.2.5 Independent Sample t-test

A comparison of descriptive statistics for academic success, as measured by first-year GPA, was conducted using three subsample participant groups (1) all participants, 2) participants meeting the 'minimum score' requirement and 3) participants meeting the 'minimum discretionary range' requirement) from each TOEFL iBT or IELTS test (Research Question 1.4). Review of the mean GPA for the TOEFL iBT and IELTS all participants group indicated that the majority of these students performed well in their first year of undergraduate studies ($M = 75.68$, $SD = 11.01$ and $M = 73.66$, $SD = 9.92$, respectively). Therefore, independent-sample t-tests were conducted to compare the first-year GPA of five subset groups submitting a TOEFL iBT or IELTS test score (see Table 9). This analysis showed no significant difference in the first-year mean GPA of all participants who submitted either the TOEFL iBT or IELTS for admission purposes, $t(196) = 1.26$, $p = .210$.

Further independent sample t-tests were conducted to compare the mean GPA of participants meeting the TOEFL iBT and IELTS 'minimum score' or 'minimum discretionary range' requirements. When the first-year GPA of TOEFL iBT applicants meeting the 'minimum score' requirement ($M = 77.53$, $SD = 10.16$) and the first-year GPA of TOEFL iBT applicants meeting only the 'minimum discretionary range' requirement (excluding applicants meeting both the 'minimum discretionary range' and 'minimum score' requirements; $M = 70.14$, $SD = 12.51$) were compared, a significant difference was seen, $t(49) = 2.17$, $p = .035$. Conversely, when the first-year GPA of IELTS applicants meeting the 'minimum score' requirement ($M = 74.56$, $SD = 8.61$) and the first-year GPA of IELTS applicants meeting only the 'minimum discretionary range' requirement (excluding applicants meeting both the 'minimum discretionary range' and 'minimum score' requirements; $M = 73.34$, $SD = 10.32$) were compared, no significant difference was seen, $t(130) = .567$, $p = .572$.

The mean first-year GPA of participants submitting TOEFL iBT or IELTS test scores were compared based on the minimum requirements of TOEFL iBT or IELTS score that was met, in order to compare the mean first-year GPA across tests. When the mean first-year GPA of participants submitting TOEFL iBT scores which met the 'minimum score' requirement ($M = 77.53$, $SD = 10.16$) were compared with participants submitting IELTS scores which met the 'minimum score' requirement ($M = 74.56$, $SD = 8.61$), no significant difference was found, $t(62) = 1.23$, $p = .224$. Similarly, when the mean first-year GPA of participants meeting the TOEFL iBT 'minimum discretionary range' requirement only ($M = 75.50$, $SD = 11.23$) were compared with participants meeting the IELTS score which met the 'minimum discretionary range' requirement only ($M = 73.59$, $SD = 9.98$), no significant difference was observed, $t(181) = 1.12$, $p = .263$. When the 'minimum discretionary range' score only was used as inclusion criteria, any

participant meeting both the 'minimum score' and the 'minimum discretionary range' was excluded.

Table 9

Independent Sample T-test Results for Five Comparison TOEFL iBT and IELTS Test Score Groups

		TOEFL iBT			IELTS		
		All	Minimum Score	Minimum Discretionary Range	All	Minimum Score	Minimum Discretionary Range
TOEFL iBT	All				1.26		
	Minimum Score			2.17*	1.23		
	Minimum Discretionary Range		2.17*				1.12
IELTS	All	1.26					
	Minimum Score		1.23				.567
	Minimum Discretionary Range			1.12		.5670	

* significant at $p = .05$

4.2.6 Success as Measured by Progression Requirements

Following the analysis of the initial research questions, post hoc analysis to review academic success as a categorical variable in which applicants either succeed or fail was performed (Research Question 1.5). Although a common assumption sometimes is that English language proficiency tests predict academic success and applicants 'real-life' academic ability once they are admitted to university (Brooks & Swain, 2014; Simner, 1998), TOEFL iBT and

IELTS claim to provide post-secondary institutions with a method of determining if applicants'

English language proficiency is sufficient to use English effectively in an English 'medium'

academic environment (Cambridge ESOL, 2003; Educational Testing Services, 2011). This

study used a tetrachoric correlation to examine academic success as a categorical variable in

which applicants who subsequently enrol as students either pass their first-year progression

requirements (GPA > 55%), or do not (GPA< 55%). Similarly, English language proficiency

was viewed as a categorical variable in which applicants' ability to meet the minimum

requirement (or not) for subsamples of TOEFL iBT or IELTS was based on participants ability to

meet 1) the 'minimum score' requirement and 2) the 'minimum discretionary range' requirement.

When viewed as categorical variables, no significant relationship was found between TOEFL

iBT, as measured by the 'minimum score' requirement, and academic success, as measured by

first-year GPA progression requirements, r (54) = .17, p = .670. Similarly, no relationship was

found between IELTS test scores meeting the 'minimum score' requirement, and academic

success, as measured by first-year GPA progression requirement, r (142) = .27, p = .153.

Having tested a notable number of relationships it emerged that no significant relationship

was found between IELTS test scores meeting the 'minimum discretionary range', and academic

success, as measured by first-year GPA progression requirements, r(142) = -.06, p = .748. By

contrast, TOEFL iBT test scores meeting the 'minimum discretionary range', were significantly

correlated with GPA, as measured by first-year GPA progression requirements, r (54) = .46, p <

.001.

4.3   Prior English Ability

4.3.1   Pearson Correlations

Pearson correlation analyses determined whether the TOEFL iBT or IELTS Overall test scores were related to prior English ability, as measured by ENG4U grade (Research Question 2.1). Results showed that ENG4U grade ($M = 78.60$, $SD = 9.87$) was moderately positively correlated with the TOEFL iBT Overall test score ($M = 94.07$, $SD = 12.65$), $r (650) = .34$, $p <$ .001 (Figure 7). Similarly, a moderate significant relationship was observed between the IELTS Overall test score ($M = 6.42$, $SD = 0.64$) and ENG4U grade ($M = 78.36$, $SD = 8.78$), $r (2145) =$ .38, $p < .001$ (Figure 8).
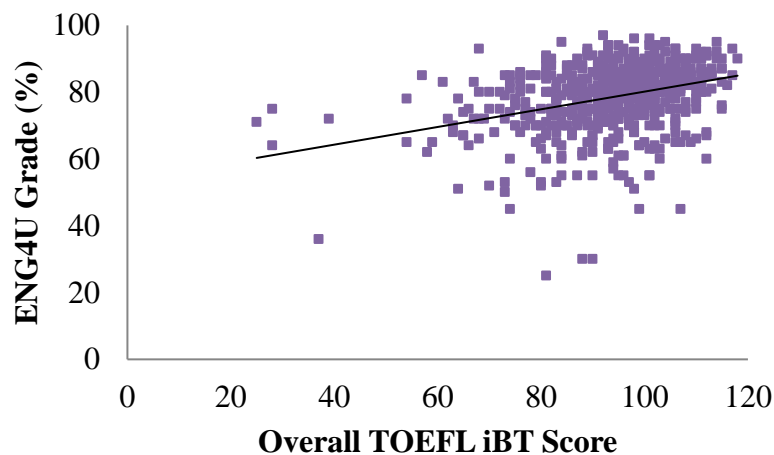
Figure 7. Correlation between Overall TOEFL iBT score and prior English ability, as measured by ENG4U grade.
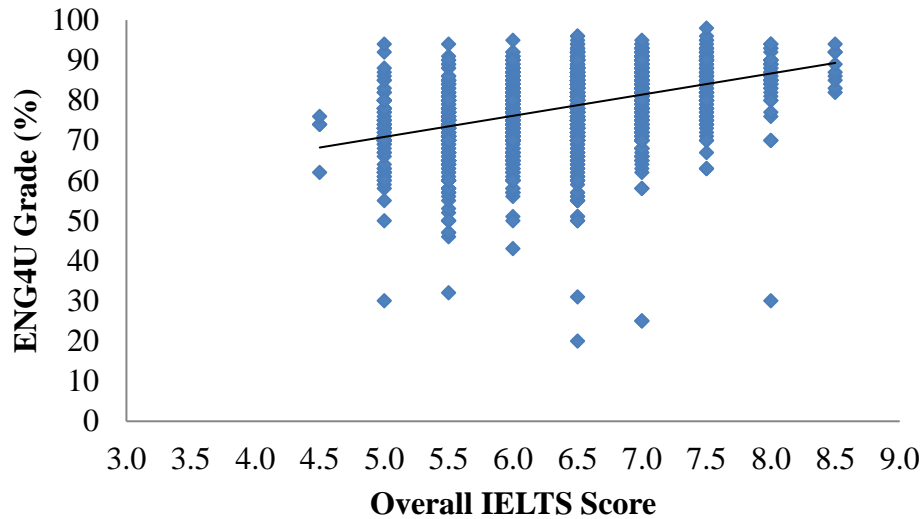
Figure 8. Correlation between Overall IELTS score and prior English ability, as measured by ENG4U grade.

Analysis of prior English ability, as determined by the participants' ENG4U grade, was analyzed through Pearson correlations with TOEFL iBT or IELTS subsample groups of applicants who met 1) the 'minimum score' requirements and 2) the 'minimum discretionary range' requirement (Research Question 2.2). For applicants who submitted a TOEFL iBT score that met the 'minimum score' requirement, a weak relationship was demonstrated between the Overall TOEFL iBT score ($M = 102.62$, $SD = 6.36$) and ENG4U grade ($M = 81.40$, $SD = 8.51$), $r(241) = .15$, $p = .024$ (Figure 9). Similarly, a weak significant correlation was found when the Overall score for applicants who submitted TOEFL iBT scores ($M = 100.07$, $SD = 7.33$) that met the 'minimum discretionary range' requirement were correlated with ENG4U grade ($M = 80.75$, $SD = 8.60$), $r(327) = .18$, $p = .001$ (Figure 10).
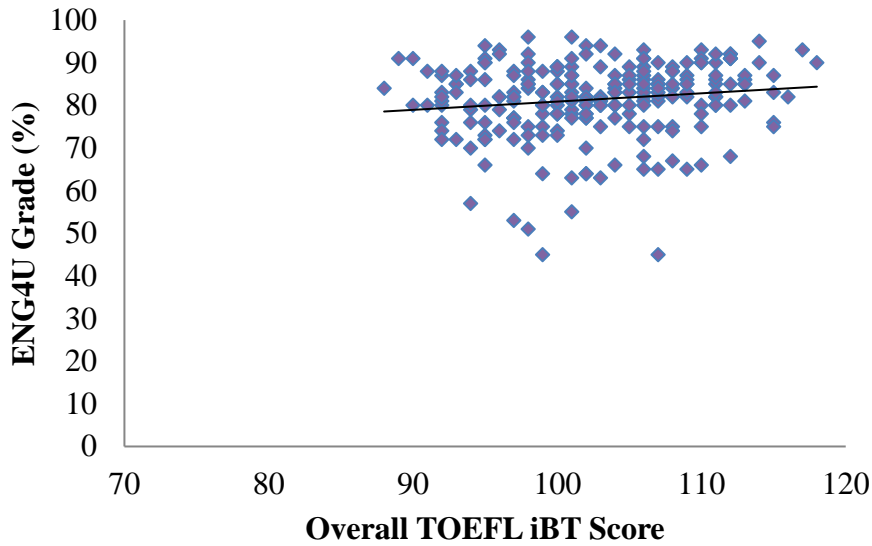
Figure 9. Correlation between the Overall TOEFL iBT score, of participants meeting the 'minimum score', and prior English ability, as measured by ENG4U grade.
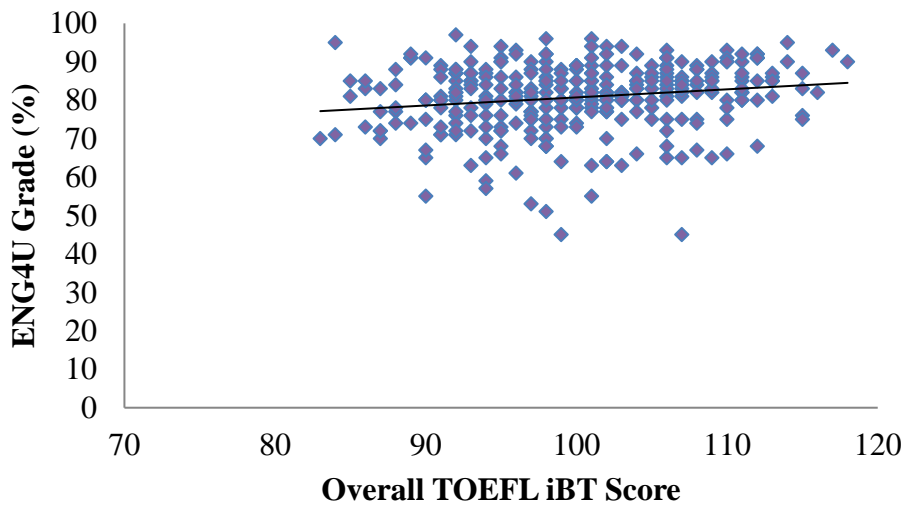
Figure 10. Correlation between the Overall TOEFL iBT score, of participants meeting the 'minimum discretionary range', and prior English ability, as measured by ENG4U grade.

When prospective students' Overall IELTS score ($M = 7.35$, $SD = .41$), which met the 'minimum score' requirement was correlated with ENG4U grade (M = 83.20, SD = 7.85) a weak

significant relationship was found, r (282) = .17, p = .005 (Figure 11). Similarly, a weak

significant correlation was seen for those variables (IELTS, M = 6.83, SD = .43; and ENG4U, M

= 81.11, SD = 7.76) when applicants who met the 'minimum discretionary range' requirement

were used, r (1181) = .21, p< .001 (Figure 12).
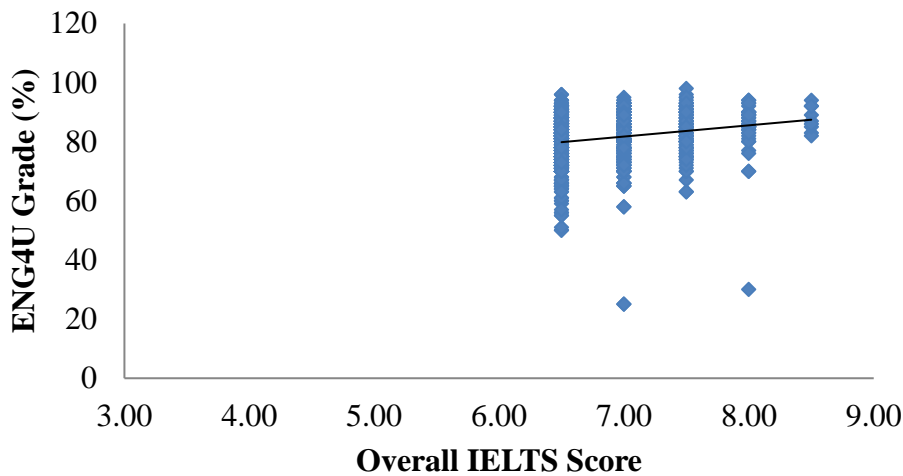
Figure 11. Correlation between Overall IELTS score, of participants meeting the 'minimum

score', and prior English ability, as measured by ENG4U grade.
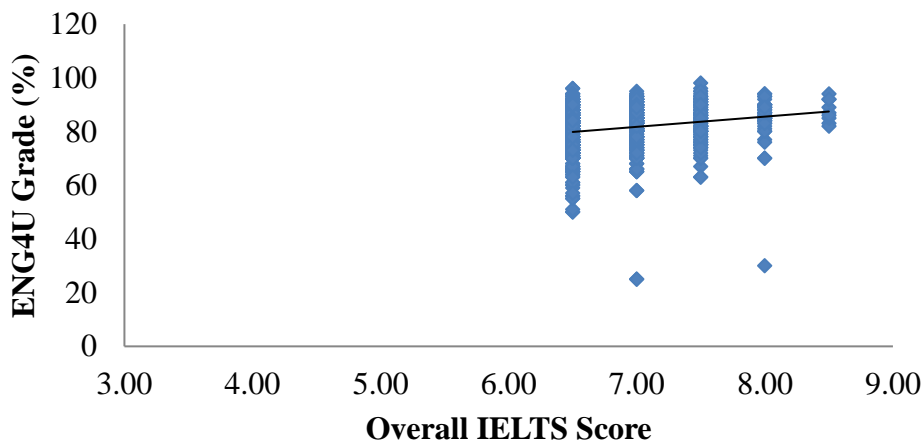
Figure 12. Correlation between Overall IELTS score, of participants meeting the 'minimum

discretionary range', and prior English ability, as measured by ENG4U grade.

4.3.2 Multiple Regression: All Participants

Additional multiple regression analyses were conducted on subsamples scores of all TOEFL iBT or IELTS participants with prior English ability (ENG4U grade). These multiple regression analyses were conducted using the four subsections of TOEFL iBT or IELTS (Listening, Reading, Writing, and Speaking) for each subsample group outlined above (Research Question 2.3). Simultaneous entry of all TOEFL iBT participants' subsection scores, Listening (LISTN; $M = 23.82$, $SD = 4.15$), Reading (READ; $M = 23.25$, $SD = 4.73$), Writing (WRIT; $M = 24.04$, $SD = 3.24$), and Speaking (SPEAK; $M = 23.25$, $SD = 3.35$) into an SSTYPE(3) (Model I) multiple regression analysis revealed a significant predictive model of ENG4U grade ($M = 78.60$, $SD = 9.87$), $F(4,647) = 23.30$, $p < .001$ (unadjusted/adjusted $R^2 = .126/.120$). As seen in Table 9, the TOEFL iBT Writing subsection contributed to the model equation significantly ($B/\beta = 0.61/.20$, $t(647) = 4.20$, $p < .001$), and was significantly correlated to ENG4U grade after accounting for the other subsections (partial correlation: $r(647) = .16$, $p < .001$). The magnitude of the zero-order correlation suggests that the predictive ability of the Writing section is moderate ($r(647) = .322$, $p < .001$). All other subsections (Listening, Reading and Speaking) did not show any significant relationships based on the selected significance values $p < .01$ (partial correlation: $r(647) = .038$, $p = .33$; $r(647) = .082$, $p = .04$; and $r(647) = .087$, $p = .03$, respectively; zero-order: $r(647) = .242$, $p = .33$; $r(647) = .247$, $p = .04$; and $r(647) = .273$, $p = .03$, respectively)

Table 10

Multiple Regression Values of TOEFL iBT Subsection Scores, of All Participants, on ENG4U

Grade.

	Unstandardized B	*Standardized β*	*t*	*p*	*Zero-order*	*Partial*
Intercept	49.26		15.70	.000		
LISTN	.117	.049	.972	.332	.242	.038
READ	.221	.106	2.10	.036	.247	.082
WRIT	.610	.200	4.20	.000	.322	.163
SPEAK	.291	.099	2.23	.026	.237	.087

$R^2 = .126$

Significant at $p < .01$

Simultaneous entry of participants' scores on each IELTS subsections, Listening ($M =$ 6.57, $SD = .97$), Reading ($M = 6.48$, $SD = .93$), Writing ($M = 6.06$, $SD = .62$), and Speaking ($M =$ 6.24, $SD = .72$) into an SSTYPE3 (Model I) multiple regression analysis revealed a significant predictor model of ENG4U grade ($M = 78.37$, $SD = 8.78$), $F(4,2142) = 96.27$, $p < .001$. As seen in Table 10, all IELTS subsections contributed significantly to the model equation's prediction of ENG4U grade (LISTN, B/β = 0.80/.09, $t(2142) = 3.14$, $p = .002$; READ, B/β = 1.87/.20, $t(2142)$ $= 7.34$, $p < .001$; WRIT, B/β = 1.87/.13, $t(2142) = 5.74$, $p < .001$; and SPEAK, B/β = 0.92/.08, $t(2142) = 3.22$, $p = .001$). The unadjusted R^2 indicated that the model accounted for 15.2% of the variance in ENG4U grade (adjusted $R^2=.15$). After accounting for the other subsections' influence, Listening, Reading, Writing and Speaking were significantly positively related to ENG4U grade (partial correlations: $r(2142) = .07$, $p = .002$; $r(2142) = .16$, $p< .001$; $r(2142) =$ .12, $p< .001$; and $r(2142) = .07$, $p = .001$, respectively). The relatively low magnitude of the zero-order correlations, however, suggest the predictive ability of the Writing and Speaking are weak overall ($r(2142) = .29$, $p < .001$; and $r(2142) = .25$, $p = .001$, respectively), while the

relations between Listening and Reading to ENG4U grade are moderate (r (2142) = .31, p = .002; and r (2142) = .35, p < .001, respectively).

Table 11

Multiple Regression Values of IELTS Subsection Scores, of All Participants, on ENG4U Grade.

	Unstandardized B	Standardized β	t	p	Zero-order	Partial
Intercept	43.92		22.06	.000		
LISTN	.800	.089	3.14	.002	.313	.068
READ	1.87	.198	7.34	.000	.345	.157
WRIT	1.87	.132	5.74	.000	.286	.123
SPEAK	.923	.075	3.22	.001	.246	.070

$R^2 = .152$

Significant at $p < .01$

4.3.3   Multiple Regression: Minimum Score Group

A similar post hoc multiple regression analysis predicting ENG4U grade (M = 81.40, SD = 8.51) was completed using only participants who met the 'minimum score' requirement (Research Question 2.3). When the TOEFL iBT test score of participants who met the 'minimum score' requirement of each TOEFL iBT subsection, Listening (M =25.81, SD = 2.56), Reading (M = 25.92, SD = 2.56), Writing (M = 25.87, SD = 2.24), and Speaking (M = 24.98, SD = 2.41) were analyzed, the resulting prediction model was nonsignificant based on the selected significance value of $p < .01$, F (4,238) = 2.96, p = .021. This model suggested that when the 'minimum score' requirement was used, TOEFL iBT subsection scores accounted for only 4.7% of the variance in ENG4U grade (unadjusted R^2 implied that the predictor variables, R^2 = .047, adjusted R^2= .031). Within the model, TOEFL iBT Writing added significantly to the prediction of ENG4U grade (B/β = 0.73/.19, t(238) =2.72, p = .007), but Listening, Reading, and Speaking did not (B/β = -0.17/-.05, t(238) = -0.76, p = .450; B/β = 0.15/.04, t(238) = 0.61, p= *.544*; and

B/β = 0.13/.04, $t(238) = 0.52$, $p = .601$, respectively; see Table 11). After accounting for the

other TOEFL iBT subsections' influence, Writing was also significantly related to ENG4U grade

(r (238) = .17, p = .007), but Listening, Reading, and Speaking were not (partial correlations: r

(238) = -.05, p = .450; r (238) = .04, p = .544; and r (238) = .03, p = .601, respectively).

Table 12

Multiple Regression Values of TOEFL iBT Subsection Scores, of Participants Meeting the

'Minimum Score' Requirement, on ENG4U Grade.

	Unstandardized B	Standardized β	T	p	Zero-order	Partial
Intercept	60.01		6.85	.000		
LISTN	-.174	-.052	-.757	.450	-.004	-.049
READ	.151	.042	.607	.544	.084	.039
WRIT	.725	.191	2.72	.007	.210	.174
SPEAK	.129	.036	.523	.601	.097	.034

R^2 = .047
Significant at p < .01

Again post hoc multiple regression analyses were completed using participants who met

the 'minimum score' requirements. When using IELTS participants, who met the 'minimum

score' requirement, each IELTS subsection, Listening (M =7.68, SD =.731), Reading (M = 7.55,

SD = .743), Writing (M = 6.82, SD = .394), and Speaking (M = 7.06, SD = .606) was entered into

a model to predict with ENG4U grade (M =83.20, SD = 7.85). The model revealed a significant

predictor model, F (4,279) = 5.93, p< .001. The unadjusted R^2 indicated that the model accounted

for 7.8% of the variance in ENG4U grade (adjusted R^2= .065). Within this model, the IELTS

Reading subsection added to the model's prediction of ENG4U grade significantly based on the

selected significance value of p < .01 (Listening, B/β = -0.03/-.003, $t(279)$ = -.04, p =

.969;Reading, B/β = 2.94/.28, $t(279)$ = 4.22, p < .001, Writing, B/β = .81/.04, $t(279)$ = .68, p =

.495; and Speaking, B/β = -1.49/-.12, $t(279)$ = -1.87, p = .062; see Table 12). After accounting

for the other IELTS subsections, Reading was related to ENG4U grade (partial correlation: r

(279) = .25, p< .001), but Listening, Writing, and Speaking were not (partial correlations: r (279)

= -.002, p = .969; r (279) = .04, p = .495; and r (279) = -.11, p = .062, respectively). Although

significant, the zero-order Reading–ENG4U grade correlation suggests relatively weak predictive

ability of the Reading score on ENG4U grade, r (279) = .26, p < .001.

Table 13

Multiple Regression Values of IELTS Subsection Scores, of Participants Meeting the 'Minimum

Score' Requirement, on ENG4U Grade.

	Unstandardized B	*Standardized β*	*t*	*p*	*Zero-order*	*Partial*
Intercept	66.23		7.03	.000		
LISTN	-.029	-.003	-.039	.969	.102	-.002
READ	2.94	.278	4.22	.000	.255	.245
WRIT	.806	.040	.684	.495	.067	.041
SPEAK	-1.49	-.115	-1.87	.062	-.044	-.111

R^2 = .078
Significant at p < .01

4.3.4 Multiple Regression: Minimum Discretionary Range Group

Each TOEFL iBT subsection for participants who met the 'minimum discretionary range'

requirement, Listening (M =25.36, SD =2.65), Reading (M = 25.08, SD = 2.87), Writing (M =

25.18, SD = 2.61), and Speaking (M = 24.43, SD = 2.59) were entered into a post hoc regression

model to predict ENG4U grade (M=80.75, SD = 8.60) (Research Question 2.3). The resulting

model was significant based on the selected significance value of p < .01, F (4,324) =3.43, p =

.009, but only accounted for 4.1% of the variance in ENG4U grade (R^2 = .041, adjusted R^2=

.029). Also, none of the four subsections (Listening, Reading, Writing, Speaking) influenced the

model's ENG4U grade prediction significantly (B/β = -0.02/-.01, $t(324)$ = -0.11, p = .911; B/β = 0.26/.09, $t(324)$ = 1.42, p = .158; B/β = 0.47/.14, $t(324)$ = 2.33, p = .020; and B/β = 0.13/.04, $t(324)$ = 0.62, p = .533, respectively; see Table 13). Also, after accounting for the other TOEFL iBT subsections' influence, none of the subsections were related to ENG4U grade (partial correlations: r (324) = -.01, p = .911; r (324) = .08, p = .158; r (324) = .13, p = .020; and r (324) = .04, p = .533, respectively).

Table 14

Multiple Regression Values of TOEFL iBT Subsection Scores, of Participants Meeting the

'Minimum Discretionary Range' Requirement, on ENG4U Grade.

	Unstandardized B	*Standardized β*	*t*	*p*	*Zero-order*	*Partial*
Intercept	59.91		9.29	.000		
LISTN	-.022	-.007	-.112	.911	.064	-.006
READ	.259	.086	1.42	.158	.131	.078
WRIT	.469	.142	2.33	.020	.181	.128
SPEAK	.127	.038	.624	.533	.106	.035

R^2 = .041

Significant at p < .01

Analysis of IELTS participants, who met the 'minimum discretionary range' requirement, for each IELTS subsection, Listening (M =7.09, SD =.781), Reading (M = 6.97, SD = .752), Writing (M = 6.36, SD = .435), and Speaking (M = 6.57, SD = .604) revealed a significant predictor model of ENG4U grade (M =81.11, SD = 7.76), F (4,1178) = 14.91, p< .001. The model accounted for 4.8% of the variance in ENG4U grade (adjusted R^2= .045). Within the model, IELTS Reading and Writing subsections added to the model's prediction significantly (Listening, B/β = 0.69/.07, $t(1178)$ = 2.13, p = .034; Reading, B/β = 1.55/.15, $t(1178)$ = 4.68, p< .001;Writing, B/β = 1.36/.08, $t(1178)$ = 2.58, p = .010; and Speaking, B/β = -.06/-.004, $t(1178)$ =

-.15, $p = .882$; see Table 14). This is supported by partial correlations of r (1178) = .06, $p = .034$;

r (1178) = .14, $p < .001$; r (1178) = .08, $p = 010$; and r (1178) = -.004, $p = .882$, respectively.

Although significant, the zero-order correlation suggests weak predictive ability of both Reading

and Writing subsection scores on ENG4U grade, r (1178) = .20, $p < .001$ and r (1178) = .12, $p =$

.010, respectively.

Table 15

Multiple Regression Values of IELTS Subsection Scores, of Participants Meeting the 'Minimum

Discretionary Range' Requirement, on ENG4U Grade.

	Unstandardized B	Standardized β	t	p	Zero-order	Partial
Intercept	57.16		14.61	.000		
LISTN	.689	.069	2.13	.034	.148	.062
READ	1.55	.150	4.68	.000	.195	.135
WRIT	1.36	.076	2.58	.010	.118	.075
SPEAK	-.058	-.004	-.148	.882	.062	-.004

$R^2 = .048$

Significant at $p < .01$

Chapter 5

5 Discussion

5.1 Academic Success

Based on a review of current literature, predictive validity studies of academic success are commonly used in determining the ability of English language proficiency tests to predict the academic success of applicants, and subsequent future students. The university in which this study was based published two sets of minimum English language proficiency test score requirements, the 'minimum score' and the 'minimum discretionary range', which applicants whose first language is not English and who have studied for fewer than three years in an Ontario high school must meet before being eligible for admission to undergraduate studies (University X, as in References, 2013, see Table 1). Current practice at this university dictates that applicants must first meet the 'minimum discretionary range' requirement before being reviewed for admission based on their academic admission average. This study reviewed the predictive validity of TOEFL iBT and IELTS test scores on the first-year GPA of OUAC 101 applicants to determine if TOEFL iBT or IELTS test scores are able to predict an applicant's academic success once enrolled in undergraduate studies. In order to review a varying range of requirements, both the 'minimum score' (a higher requirement), and the 'minimum discretionary range' (a lower requirement) were used to determine if TOEFL iBT and/or IELTS test scores meeting those requirements were able to predict academic success. Although the 'minimum discretionary range' subsample group also included those applicants in the 'minimum score' subsample group, both subsample groups were used in this exploratory study to determine if any patterns emerged which may stimulate future study. Both subsample groups were reviewed independently of the

other group and interpretations comparing those two groups were based on their projected use in practice.

The aim of this analysis was to determine if either range was optimal for use in practice to identify applicants with the potential to succeed academically once enrolled as students at this Ontario university. Overall the results of this study did not provide statistically significant support for TOEFL iBT or IELTS to predict the academic success of OUAC 101 applicants applying to this Ontario university and who have been required to submit proof of English language proficiency. Based on the small sample size of TOEFL iBT ($n = 37$) and IELTS ($n = 27$) 'minimum score' groups, no inferences were made as to the ability of those tests at the 'minimum score' requirement to predict academic success. Instead, the results of that analysis were provided to aid in the development of future research studies.

Although sample sizes were relatively small for all other groups, this study did not statistically support the use of TOEFL iBT or IELTS to be used when identifying OUAC 101 applicants whose first language is not English, have studied for fewer than three years Ontario high school and who have the potential to succeed academically at this Ontario University. Therefore the results of this study disprove the hypothesis that a strong significant relationship exists between each TOEFL iBT and IELTS, and academic success as measured by first-year GPA.

5.1.1   TOEFL iBT and Academic Success

Correlational analysis between the TOEFL iBT Overall test score and first-year GPA was conducted using all three subsample groups a) 'all participants', 2) those meeting the 'minimum score' requirement, and 3) those meeting the 'minimum discretionary range' requirement to

determine if the TOEFL iBT Overall score was able to predict the academic success of OUAC 101 applicants required to submit proof of English language proficiency. Among the 'all participant' group, no significant relationship with first-year GPA was found. Based on zero-order correlations and the results of multiple regression analysis the Overall TOEFL iBT score was not able to predict an applicant's academic success in their first year of undergraduate study at one Ontario University (i.e. no significant regression model was observed, nor were any correlational relationships found between TOEFL iBT subsections and first-year GPA for the 'all participants' group). It should be noted that the sample sizes for the TOEFL iBT Overall score correlation and multiple regression analyses with GPA were relatively small (n = 56). Therefore, although the results suggest that TOEFL iBT Overall score is not able to predict applicants' future academic success, this relationship may be limited by sample size.

Applicants must currently meet the 'minimum discretionary range' requirement in order to be considered eligible for admission based on the assumption that applicants meeting the 'minimum discretionary range' posses sufficient English language proficiency to succeed at this Ontario university, which is an English medium institution. Unlike the 'all participant' group which provided a more general relationship based on a wider range of data (as test scores both above and below the 'minimum discretionary range' requirement were included), analysis of the 'minimum discretionary range' requirement provided data specifically related to the current English language proficiency requirement used by this Ontario university. Correlational analysis of the TOEFL iBT participants meeting the 'minimum discretionary range' requirement identified a moderately significant correlation between the TOEFL iBT Overall score and first-year GPA. Those results appeared to initially suggest that the Overall TOEFL iBT score at the 'minimum discretionary range' was to be able to predict applicants' future academic success in

an undergraduate program at this Ontario University. As the sample size for this analysis was relatively small (n = 51), it is possible that with a larger sample size this moderately significant relationship may exhibit a stronger correlation coefficient.

The subsequent multiple regression model found none of the subsections to be significantly related to first-year GPA, indicating that although the Pearson correlation and multiple regression model were significant, TOEFL iBT score at the 'minimum discretionary range' requirement was in fact not able to predict the academic success of OUAC 101 applicants applying to this Ontario university. Specifically, the lack of a significant relationship between any of the TOEFL iBT subsection and first-year GPA in the regression model suggested that none of the TOEFL iBT subsections explained significant amounts of variance in first-year GPA, meaning the observed significant Pearson correlation and multiple regression model must be due to other confounding and unmeasured variables that are not represented in this study. Therefore, based on those results TOEFL iBT was not able to reliably predict the academic success of OUAC 101 applicants at the 'minimum discretionary range' requirement. Again it is important to note that with a larger sample size one or more of the TOEFL iBT subsections may be significantly related to GPA. Consequently, the use of the 'minimum discretionary range' requirement as a means for determining applicants with the potential to succeed academically once registered in an undergraduate program at this Ontario university is not, on its own, an adequate benchmark.

As the results of this analysis did not provide reliable support for the use of the current 'minimum discretionary range' requirements used in practice by this Ontario university, correlational and multiple regression analyses were conducted to determine if the higher 'minimum score' requirement allowed for applicants' future academic success to be predicted.

Based at least in part to the small sample size (n = 37), neither correlational nor multiple regression analysis of the TOEFL iBT participants meeting the 'minimum score' requirement were able to unearth any significant relationship between the TOEFL iBT Overall score or its subsections and academic success, as measured by first-year GPA. Therefore, the value of the 'minimum score' requirement as a single benchmark to be relied upon for judging future academic success is also highly suspect.

The results of this analysis did not provide reliable statistical support for the use of 'minimum discretionary range' TOEFL iBT requirement as currently set by University senate, when predicting the academic success and subsequently determining the ability of potential students' to succeed at this English-speaking university. It is important to note that alternative options to TOEFL iBT are limited. Therefore, although the continued use of those tests may be required to establish an objective benchmark, it is recommended that TOEFL iBT alone is not used as the sole determining factor of an applicant's ability to succeed once enrolled as an undergraduate student.

5.1.2 IELTS and Academic Success

Similar to the TOEFL iBT test score and academic success analysis, Pearson correlation and multiple regression analyses were conducted for three IELTS subsample groups 1) 'all participants', 2) those meeting the 'minimum score' requirement, and 3) those meeting the 'minimum discretionary range' requirement, to determine if IELTS test score was able to predict the academic success of OUAC 101 applicants applying to one Ontario university. None of the correlation or multiple regression analyses showed significant relations between the IELTS Overall or subsection scores and first-year GPA for either the 'all participant', suggesting that IELTS test score is also not able to significantly predict academic success for this range of test

scores. Although the results of the 'minimum score' subsample group were similar in suggesting that no significant relationship exists between IELTS test scores at this range and GPA, the sample size of this analysis was far too small to infer any such interpretation.

When the IELTS overall score meeting the 'minimum discretionary range' requirement was correlated with academic success, a weak significant zero-order correlation was identified. This correlation was not supported by multiple regression analysis, as neither the regression model overall nor the IELTS subsections in the model were significantly related to first-year GPA. Those results suggested that some other factors, not measured by IELTS or within this study, are responsible for the zero-order relationship which appeared between Overall IELTS score and academic success. Therefore, those results do not provide reliable support for the use of the IELTS 'minimum discretionary range' as the sole benchmark for identifying applicants with sufficient English language proficiency to success once enrolled as an undergraduate student at this Ontario university.

5.1.3 Predictability

A common assumption of English language proficiency tests is that they predict applicants' future academic success (Brooks & Swain, 2014; Simner, 1998). This assumption is the basis, at least in part, for admission procedures around the world. However, neither ETS nor the British Council claim that TOEFL iBT or IELTS are able to predict academic success. In fact, ETS clearly states that the intention of the TOEFL iBT is to a) "stimulate university communication", b) "help test takers determine their academic readiness" and c) "help institutions identify and select students with the English communication skills required to succeed" (Educational Testing Service, 2010). Although the developers of TOEFL iBT and IELTS claim their tests can determine if an applicant has the skills "required to succeed" in an

English-speaking academic environment, they do not claim to determine how well the applicant will succeed. When reviewing the predictive validity of TOEFL iBT or IELTS test scores, the use of correlation and multiple regression analyses, as linear analyses, sets out to roughly determine how well an applicant will perform academically, based on the test score received. Essentially, those predictive validity studies, this study included, aim to determine the linear relationship between those tests and academic success (i.e., determine if applicants with high TOEFL iBT or IELTS test scores are able to achieve high first-year GPA). Since ETS and the British council claim that their tests determine an applicants' ability to be successful, and not how successful those applicants may be, this study has taken an alternative approach and has viewed both test score and academic success as dichotomous variables in which applicants either meet a set requirement or do not.

Post hoc analysis was developed to determine if either TOEFL iBT or IELTS test scores at either minimum requirement were able to determine if an applicant would succeed in their first year of undergraduate education at this Ontario University, and therefore validate what TOEFL iBT and IELTS claim to measure. To do so, the 'minimum score' and the 'minimum discretionary range' requirements were each set as benchmarks that applicants must reach in order to be viewed as meeting the English language proficiency requirement. Any participant meeting the set English language proficiency requirement was placed into one of the two dichotomous test score groups, and those participants not meeting the requirement was placed in the other group. Participants' were divided into two dichotomous groups based on academic success as well. The academic success groups were divided based on participants' ability to meet their first-year GPA progression requirement of 55%. Therefore participants either met their first-year progression requirement (GPA > 55%) or did not (GPA < 55%). To validate the statement that TOEFL iBT and IELTS are able to determine if applicants are able to succeed, and

if progression within a program is viewed as succeeding academically then all applicants meeting the English language proficiency requirement should also have a GPA above 55%. A tetrachoric correlation was then conducted to determine the relationship between those two dichotomous variables. This categorical analysis of applicants' ability to succeed or not may be beneficial as other factors such as motivation should be limited as all participants should be motivated to at least meet their minimum progression requirements.

The intention of this analysis was to determine if those applicants who met either the 'minimum score' or 'minimum discretionary range' requirements were also able to meet the first-year progression requirement. Of both TOEFL iBT and IELTS analysis a significant relationship was only revealed between TOEFL iBT test score and academic success when the 'minimum discretionary range' requirement was used. Therefore, only TOEFL iBT when using the 'minimum discretionary range' requirement as a benchmark is able to statistically predict which applicants will likely succeed academically and be able to meet their first-year progression requirements once enrolled in undergraduate study at this Ontario university.

5.1.4   Differences in Mean GPA

To determine if any subsample groups contained participants with specifically higher mean first-year GPAs, an independent sample t-test was used to compare the mean first-year GPA of five comparison groups: 1) all TOEFL iBT with all IELTS participants, 2) TOEFL iBT participants meeting the 'minimum score' requirement with IELTS participants meeting the 'minimum score' requirement, 3) TOEFL iBT participants meeting the 'minimum discretionary range' requirement with IELTS participants meeting the 'minimum discretionary range' requirement, 4) TOEFL iBT participants meeting the 'minimum score' requirement with TOEFL iBT participants meeting the 'minimum discretionary range' requirement, and 5) IELTS

participants meeting the 'minimum score' requirement with IELTS participants meeting the 'minimum discretionary range' requirement (Table 9).

This analysis did not reveal any significant differences in the mean first-year GPA of the five comparison groups with the exception of the TOEFL iBT 'minimum score' and the TOEFL iBT 'minimum discretionary range' groups. Although a significant difference was found between the mean GPA of those two TOEFL iBT samples, those results alone do not support the 'minimum score' requirement as being better able to determine applicants with the potential to succeed once admitted to this Ontario university. Essentially, those results only suggested that applicants meeting the 'minimum discretionary range' only (excluding any students who would meet both requirements), do not achieve the same mean first-year GPA as those applicants meeting the required 'minimum score'. Since the intended use of TOEFL iBT by the admission department within this University is to identify applicants with sufficient English language proficiency to potentially succeed once enrolled, it is more important as to the applicants' ability to succeed than the level to which they do so.

5.1.5   Academic Success and OUAC 101 Applications

As part of its Strategic Action Plan the Ontario university in this study has set out to increase the number of international students and faculty (University X, as in References, 2014). This goal stems from a desire to internationalize the university and provide a deeper level of education than would be provided by a homogenous student population. In order to do so, it is critical that this Ontario university is able to admit international applicants who are highly likely to succeed. Although this study used applicants currently studying the Ontario curriculum, some of those applicants were studying on a study permit and were considered international students.

Therefore, it is important that this Ontario university understands how changing admission requirements would impact the number of eligible OUAC 101 applicants.

Analysis of the TOEFL iBT and IELTS subsample groups sample size revealed information that must be taken into consideration by senior administrators when determining the minimum English language proficiency requirements. Although 50.46% of participants who submitted TOEFL iBT met the 'minimum discretionary range' requirement on their first attempt, only 37.27% of those participants would have met the 'minimum score' requirement on their first attempt. Therefore, if the admission requirement was increased to the 'minimum score' requirement, rather than the 'minimum discretionary range' requirement, there may be an approximately 13.19% decrease in eligible applicants. In turn, that would almost invariably impact other university goals and targets such as first year class size. Similarly, approximately 55.10% of participants who submitted IELTS test scores met the 'minimum discretionary range' requirement on their first attempt. Of the participants submitting an IELTS test score, only a surprising 13.23% of applicants would have met the 'minimum score' requirement on their first attempt. Drawing on the sample numbers in this study, this would then result in a 41.87% decrease in eligible applicants if the IELTS minimum English language proficiency requirement was raised to the 'minimum score' requirement.

Those results alone do not support the increase of the minimum English language proficiency requirement from an institutional administrative perspective. To attract and educate the best minds from around the world, the university used within this study must ensure that applicants who have the potential to succeed academically are not denied or discouraged from applying. The results of the other analyses within this study do not provide reliable support for using the 'minimum score' requirement of TOEFL iBT or IELTS over the 'minimum discretionary range' requirement. Therefore, increasing the minimum English language

proficiency requirement to the 'minimum score' requirement could potentially disqualify or discourage many potentially successful applicants. As it can be argued that the high minimum requirements set by some universities have led applicants to make the unethical decision to submit fraudulent test scores, it can subsequently be argued that increasing the minimum requirement without sufficient evidence may propagate this issue.

5.2   Prior English Ability

The identification of fraudulent TOEFL iBT and IELTS test scores among Ontario Universities has the university used in this study, as an institution that prides itself on continuous improvement, questioning the current policies and procedures for reviewing English language proficiency tests during the admission of OUAC 101 applicants. To adapt to this emergent issue, this Ontario university regularly attempts to find new methods for identifying fraudulent test scores. Unfortunately, due to the limited amount of academic literature explaining English language proficiency test fraud, it is very difficult for institutions to adapt their policies and procedures. In addition, universities are merely told if the test score under review by ETS or the British Council is acceptable or if it has been cancelled. In the event that a test score is cancelled, no explanation is provided, again limiting the amount of information universities have available when reflecting on their current processes.

This study aimed to determine if there was a relationship between TOEFL iBT and/or IELTS and prior English ability, as measured by participants' ENG4U grade by comparing three subsample groups of participants who submitted either TOEFL iBT or IELTS with prior English ability 1) all participants, 2) participants meeting the 'minimum score' requirement, and 3) participants meeting the 'minimum discretionary range' requirement. Working under the assumption that fraudulent test scores are not representative of the applicant's true ability, it can

103

also be assumed that deviations from the hypothesized relationship between TOEF iBT or IELTS and ENG4U grade, may suggest that the test was not written by the same person who submitted the ENG4U grade as part of their application. This assumption can be made for strong significant correlations which infer strong predictive linear relationships in which test scores are able to predict academic success. Therefore this study worked under the assumption that deviations from a strong significant linear relationship between those significant TOEFL iBT or IELTS subsections and ENG4U grade could be used in practice to suggest that the test scores are not representative of the applicants projected ability, and may also suggest the existence of potentially fraudulent test scores.

The results of this study revealed that the TOEFL iBT Writing subsection and all IELTS subsection are weakly able to be predicted by ENG4U grades. As a weak correlation implies variance within the relationship, this weak correlation, although significant, would expect to see inconsistencies in the predictive relationship. Therefore, the weak correlation between the TOEFL iBT Writing subsection and all IELTS subsections with ENG4U grade is not a reliable relationship to use in practice, disproving the hypothesis that ENG4U grade may be used as a marker for identifying potentially fraudulent TOEFL iBT or IELTS test scores. Although the existence of a significant relationship was in line with the Research Question 2 hypothesis, only a weak relationship was present. Similarly, unlike this hypothesis, the Reading and Writing subsections were not the only significant TOEFL iBT or IELTS subsections.

5.2.1 TOEFL iBT and Prior English Ability

To determine the ability of ENG4U grade, as a measure of prior English ability, to be used as a marker for identifying potentially fraudulent TOEFL iBT test scores, it was important to first understand the relationship between those two variables. Pearson correlations determined

that the Overall TOEFL iBT score was able to predict ENG4U grade for all subsample groups. Although this relationship was moderately significant for the 'all participants' group, the 'minimum score' and 'minimum discretionary range' groups were only weakly correlated. The relatively weak strength of those relationships did not provide sufficient evidence for the use of ENG4U grade as a marker for identifying fraudulent Overall TOEFL iBT test scores.

To better understand those relationships, multiple regression analysis was conducted to determine the influence of each TOEFL iBT subsection on ENG4U grade. The 'all participants' and 'minimum discretionary range' groups both presented multiple regression models that were significant. In order to use those findings in practice to identifying a method for detecting potentially fraudulent TOEFL iBT tests, it is important to understand which subsections of TOEFL iBT are contributing to this relationship and the strength of this predictive relationship. Amongst the 'all participant' group, only the Writing subsection was found to weakly contribute to the relationship between TOEFL iBT and ENG4U grade. This finding is in line with the expectation that writing is the largest method of assessment used as part of the ENG4U curriculum (Ontario Ministry of Education, 2007). It would however have been expected that a stronger relationship was exhibited between the Writing subsection and ENG4U grade. As the sole contributing subsection, the Writing subsection accounted for only 12.6% of the variance in ENG4U grade, suggesting that many other non-TOEFL iBT factors were influencing this relationship. Since none of the TOEFL iBT subsections of the 'minimum discretionary range' group were found to be significantly correlated to ENG4U grade at the designated significance values of $p = .01$, those results suggested that the significant relationship exhibited by the multiple regression model and the correlation with the Overall test score must be attributed to some other underlying factors not included in this study.

The multiple regression model of the 'minimum score' group was not found to be significant at the designated significance value of $p = .01$, suggesting that ENG4U is not a sufficient marker for identifying fraud test scores of applicants in this higher TOEFL iBT score group. The inability to predict prior English ability in the 'minimum score' group may be partly attributed to fraudulent test scores that were not identified and removed from this study. Theoretically, it would be expected that fraudulent test scores would be above the minimum English language proficiency requirement, suggesting that the 'minimum score' group would potentially have the largest amount of fraud amongst all subsample groups.

Conversely the 'all participant' group, which contained scores below the minimum English language proficiency requirement, would be expected to be more representative of the applicant pool applying to university as all test scores, regardless of their ability to meet the minimum requirement, must be reviewed for potential fraud. For this reason, the 'all participant' group should be viewed as a better sample for identifying a method for detecting TOEFL iBT fraud in practice. As the TOEFL iBT Writing subsection was found to have a weak significant related with ENG4U grade, it is expected that variance within this relation exists weakening any predictive ability. Therefore deviations from the linear relationship between those variables are expected and cannot be used in practice when identifying potentially fraudulent test scores.

5.2.2   IELTS and Prior English Ability

Similar to the above TOEFL iBT analysis, the relationship between IELTS test scores and ENG4U grade were investigated to determine if ENG4U grade, as a measure of prior English ability, can be used as a marker for identifying potentially fraudulent IELTS test scores. This analysis found that the Overall IELTS test score of each of the three subsample groups 1) all

participants, 2) those meeting the 'minimum score' and 3) those meeting the 'minimum discretionary range', was able to predict ENG4U grade.

To better understand which subsections of the IELTS test are contributing to those relationships significantly, multiple regression analysis was conducted. Although multiple regression models were significant for all three subsample groups, the subsections contributing to those models varied. When the relationship amongst the IELTS subsection scores which met the 'minimum score' and ENG4U grade were analyzed, only the Writing subsection was found to be significantly correlated with ENG4U grade. As this relationship only account for 7.8% of the variation in ENG4U grade, there are many external factors which are influencing this relationship. The significant relationship between the Writing subsection and ENG4U grade is to be expected based on the types of assessments used to calculate the ENG4U grade (Ontario Ministry of Education, 2007). As with the TOEFL iBT analysis, it should be noted that the 'minimum score' subsample group was the group expected to be the most influenced by any undetected fraudulent test scores. This expectation was grounded in the assumption that fraudulent test scores are higher than the minimum English language proficiency requirement.

Similar to the 'minimum score' subsample group, the IELTS 'minimum discretionary range' subsample group also identified a relationship between ENG4U grade and the Writing subsection. In addition, this analysis found a relationship between ENG4U grade and the Reading subsection. The relationship between these two subsection scores and ENG4U grade only account for 4.8% of the variation seen in ENG4U grade suggesting that many external factors not included in this study are impacting this relationship. Again, based on the type of assessments used to calculate ENG4U grade, the relationship between those variables was not unexpected (Ontario Ministry of Education, 2007).

As mentioned with the TOEFL iBT and ENG4U grade analysis, the 'all participant' group presents the most useable data when developing a method for identifying English language proficiency test fraud as this sample is the most representative of the OUAC 101 applicant pool for which this method could be used. Specifically, the 'all participant' group represented all applicants regardless of their ability to meet the minimum English language proficiency requirement. Amongst the 'all participant' group all subsections were found to significantly contribute to the relationship between IELTS and ENG4U grade. However, those contributing subsections only account for 15.2% of the variance in ENG4U grade, suggesting that other factors not identified in this study are also contributing to this relationship. As all IELTS subsection scores are only weakly able to predict ENG4U grade, the linear relationship between those variables is not sufficient to reliably permit all IELTS subsections to be used to flag potentially fraudulent IELTS test scores. Therefore, deviation from the hypothesized relationship may not suggest potentially fraudulent IELTS test scores.

Chapter 6

6   Conclusion

6.1   Review of Study

To stay competitive in the global market of the 21st century and to add both cultural and intellectual diversity to campuses, Canadian public universities need to attract the best and brightest students from around the world. Although attracting diverse talent in the form of international students helps broaden knowledge, cultural experiences, and intellectual capacities, senior administrators must also reflect on their responsibility to Parliament and other stakeholders who help fund public universities within Canada. At the same time, senior administrators need to be cognisant of external influences, including the media, as well as the internal demands made by students, staff and faculty. In order to balance these competing tensions, those universities must develop financially wise risk management strategies to ensure that all applicants have sufficient English language proficiency to succeed in an institution where English is the dominant language of instruction. Currently, English language proficiency tests such as TOEFL iBT and IELTS are common requirements of the admission process when making decisions regarding the English language proficiency of applicants whose first language is not English as they provide an objective benchmark around which senior administrators can base their admission procedures.

Research supporting the use of TOEFL iBT and IELTS to predict academic success varies in its results to the extent that some studies present data which support a significant relationship between test score and academic success, while other studies do not (Al-Musawi& Al-Ansari, 1999; Arcuino, 2013; Cho & Bridgeman, 2012; Cotton & Conrow, 1998; Dooey & Oliver, 2002; Feast, 2002; Fu, 2012; Ng, 2007; Simner, 1998; Wait & Gressel, 2009; Woodrow, 2006; Yen & Kuzma, 2009). Although neglected amongst current literature, OUAC 101 applicants who may

currently be studying the Ontario curriculum in an English-speaking high school may be required to submit proof of English proficiency. Specifically in this case, OUAC 101 applicants whose first language is not English and who have studied for fewer than three years in an Ontario high school were required to submit proof of English language proficiency as part of their undergraduate application. This study aimed to review the use of TOEFL iBT and IELTS during the admission process for OUAC 101 applicants applying for undergraduate education at one Ontario university.

Senior administrators are challenged with the task of identifying the appropriate benchmark value at which they deem applicants' English language proficiency to be sufficient in order to succeed academically. Currently, OUAC 101 applicants whose first language is not English are required to meet the 'minimum discretionary range' requirement for the specific test score submitted in order to be identified as having sufficient English language proficiency to succeed in this English-speaking academic institution. Although it may be suggested that increasing the minimum English language proficiency requirements may subsequently identify applicants more likely to succeed academically once enrolled as students, this debate is ongoing. During the undertaking of this research this exploratory predictive validity study correlated TOEFL iBT and IELTS test scores meeting two different English language proficiency requirement ranges ('minimum score' and 'minimum discretionary range') with academic success, as measured by first-year GPA, in order to determine if either requirement range was able to identify potentially successful applicants. Both the TOEFL iBT and IELTS 'minimum discretionary range' subsample groups did not reveal any significant relationships with academic success, therefore suggesting that test scores above the 'minimum discretionary range' are not able to reliably predict first-year GPA alone. As made evident by the results of this analysis,

many other factors not considered within this study contribute to the variance exhibited in first-year GPA.

As the predictive relationship tested during this correlation analysis was linear, a significant relationship would have implied that the higher tests scores were able to predict higher first-year GPAs. As this relationship was not found, and higher test scores are not indicative of higher first-year GPAs, those results do not support the increase of the minimum English language proficiency requirement from the 'minimum discretionary range' to the 'minimum score' requirement. Although the results of the 'minimum score' subsample group were provided, the low sample size of this subsample groups did not allow for any statistically reliable results. Therefore interpretations of such data are not presented. Data for the 'minimum score' sample group were provided within this exploratory study to aid in the development of future research studies.

To understand the change in the average first-year GPA of participants, a comparative independent sample t-test analysis of the mean GPA amongst five different TOEFL iBT and IELTS participant groups was conducted. Among those groups included comparisons between the mean GPA of the 'minimum score' group and the 'minimum discretionary range' group for each TOEFL iBT and IELTS. Amongst the five comparison groups, a significant difference was only found between participants meeting the TOEFL iBT 'minimum discretionary range' only (excluding any applicants meeting both ranges) and TOEFL iBT 'minimum score' subsample groups. Although those results suggest that a significant difference exists amongst the mean first-year GPA of those groups, those results alone are not sufficient to support the increasing of the minimum English language proficiency requirement to the 'minimum score'. In fact, the mean first-year GPA of the 'minimum discretionary range' group was 70.14% which is still well

above the 55% first-year progression requirement that students are required to achieve in order to progress within their program. Further analysis of the descriptive statistics showed a large decrease in the number of applicants eligible for admission based on meeting the 'minimum score' on their first attempt, again providing no support for increasing the minimum English language proficiency requirement.

When evaluating English language proficiency tests, like all other documents submitted for admission purposes, universities must show due diligence in order to identify, as far as reasonably possible, any fraudulent documentation. Although it may be difficult to typically think of fraud in most education settings, no university is immune from such actions. Data supporting the existence of English language proficiency test fraud is virtually nonexistent among current peer reviewed literature; however the existence of such fraud is present within the media and the post-secondary education community (eg. "The China Conundrum", 2011; Guhr, & Furtado, 2013). Therefore it is important that institutions take it upon themselves to conduct studies to help strengthen their relevant policies and procedures. This study was intended to help influence the process which governs the review of TOEFL iBT and IELTS test scores at this Ontario university. Working under the assumption that fraudulent tests score are not representative of applicants' true abilities, and that a deviation from the expected predictable relationship between TOEFL iBT or IELTS and ENG4U grade is also not representative of the applicants' true abilities, this study examined the relationship between TOEFL iBT and IELTS with ENG4U grade to determine if a predictable linear relationship existed. Amongst the 'all participant' group which was selected as the most representative of the population for which this method will be employed, the TOEFL iBT Writing subsection and all IELTS subsections were linearly related to ENG4U grade, suggesting that a predictable relationship exists. Although the

intention of this study was to identify ENG4U grade as a marker for comparison with TOEFL

iBT or IELTS to help flag potentially fraudulent test scores, the weak relationships, although

significant, brings into question the use of this method in practice. Results of this study could be

further understood by qualitative analysis that investigates the views of students currently

studying at university.

6.2   Practical Consideration and Recommendations

6.2.1   TOEFL iBT and IELTS for Predicting Academic Success

University admission offices typically operate under the assumption that English language

proficiency tests act as highly reliable predictors of future academic success and applicants'

"real-life" academic ability once they are admitted to university (Brooks & Swain, 2014; Simner,

1998). In order for an English language proficiency test score, such as TOEFL iBT or IELTS, to

predict applicants' academic success, there must be a linear relationship amongst those two

variables. This study reviewed the predictive validity of each TOEFL iBT and IELTS of three

subsample groups 1) 'all participants', 2) those meeting the 'minimum score' requirement, and 3)

those meeting the 'minimum discretionary range' requirement with academic success, as

measured by first year GPA, however no significant relationship was found for any of those

groups. Those results do not provide support for the ability of TOEFL iBT or IELTS alone to

reliably predict the academic success of OUAC 101 applicants who have been requested to

submit proof of English language proficiency. Essentially, the inability TOEFL iBT and IELTS

test scores to reliably predict academic success at either of the minimum English language

proficiency requirement ranges in this study, brings into question any sole reliance of those tests

to identify applicants with the English language skills require to succeed at this Ontario

university. As previously mentioned other limitations, such as sample size, may have impacted

the results of this study and the inability of this study to identify significant relationships between test scores and academic success.

English language proficiency tests score such as TOEFL iBT and IELTS continue to be used by most public universities in Canada after many years of scrutiny and debate by the educational community (eg. "Curtin University, 2011). Although this study did not provide support for the use of those tests during the admission of OUAC 101 applicants, one question that the findings prompt is, "what is the alternative?". Based on a review of the current requirements of many Ontario universities as well as those aboard, it appears that universities are no longer willing to take a 'back seat' to the issue of English language proficiency testing. In fact, some universities are beginning to offer their own English language proficiency tests in addition to university-owned and run English as a Second Language (ESL) programs. For example, York University offers the York English Language Test (YELT) (York University, 2016). The YELT is a two part test consisting of a written portion and an oral portion during which reading, writing, comprehension, and spoken English are evaluated at a university level. It is also intended to determine whether applicants require additional ESL courses as part of their university program. Similar to TOEFL iBT and IELTS, YELT test results are valid for a period of two years.

Another example in which a university has decided to implement their own English language test is the Diagnostic English Language Assessment (DELA) at the University of Melbourne in Australia (University of Melbourne, 2016a). At present, the University of Melbourne is the top university in Australia and ranks 33 in the world (Times Higher Education, 2016). In order to ensure their students' success, regardless of their English language proficiency test scores, The University of Melbourne has created a test to assess the strengths and weaknesses of their students in academic reading, writing, and listening. This post-entry assessment cannot

be failed, in the sense that students may be required to rewrite the test, and does not impact students' admission to university. Instead, this test is intended to help students identify their areas of improvement and help direct them to the services provided by the institution. Not only does this diagnostic method of assessment benefit the student population, but it puts use to the many services already available to help students succeed.

Recommending that the Ontario university used within this study suspend their use of TOEFL iBT and IELTS goes beyond the scope of this study. It is also acknowledged that universities may be using TOEFL iBT or IELTS tests as a method of risk management, although they are aware that neither test is able to reliably predict applicant who will success. As no support for the use of those tests during the admission of OUAC 101 applicants was identified here, it is recommended that this university develops its own diagnostic English language proficiency assessment to be used in combination with currently accepted standardized English language proficiency tests such as TOEFL iBT and IELTS. As the development of a sound new test to predict applicants' academic success is very difficult and expensive to design and implement, due to the many factors other than language ability which influence academic success, it is recommended that this university approach English language assessment similar to the University of Melbourne by developing a student service focused diagnostic assessment.

Beginning July 2015 the university in this study launched a new English language program intended to help applicants with slightly low TOEFL iBT or IELTS test scores improve their English language skills during an intensive summer program. Applicants who submitted test scores slightly below the 'minimum discretionary range' requirement but who were academically admissible, were given a conditional offer of admission to an undergraduate program pending the successful completion of this intensive English language program. This

program comes at an excellent time as this study did not identify support for the sole use of TOEFL iBT or IELTS test scores during the admission process of OUAC 101 applicants. As it was seen that neither test score alone was able to predict the academic success of OUAC 101 applications at the current minimum requirement ranges, continuing to use such requirements may potentially disqualify applicants who are able to succeed during an undergraduate degree at this Ontario university. This intensive summer English language program allowed applicants below the 'minimum discretionary range' requirement, who may have been disqualified from direct admission to an undergraduate program based on their TOEFL iBT or IELTS test score alone, the opportunity to improve their skills and prepare them for their future education. As this program is still in the early stages, it is recommended that future studies be done to ensure that applicants who successfully complete this program are subsequently able to succeed in their undergraduate program.

Having found no significant correlation between TOEFL iBT or IELTS and academic success does not mean that applicants meeting the 'minimum discretionary range' are not able to succeed academically during their undergraduate education. In fact, the mean GPA of participants meeting the TOEFL iBT or IELTS 'minimum discretionary range' was 75.5% or 74.56% respectively. For this reason, creating an institution-specific diagnostic test which allows students to identify their weaknesses and seek out the services provided to them may help ensure that even those students whose English is not sufficient to succeed academically in an English-speaking institution have the services available to help them subsequently succeed.

6.2.1.1 Another View of Academic Success

The intended use of English language proficiency tests such as TOEFL iBT and IELTS during the admission process is to identify applicants who have sufficient English language

proficiency to, in theory, succeed once enrolled in an English-speaking university. In order to study this relationship, many researchers use predictive validity studies, which determine the ability of TOEFL iBT or IELTS to predict academic success along a linear continuum. In the event that a significant relationship exists, it can be expected that a high TOEFL iBT or IELTS test score can predict a high first-year GPA, and subsequently predict an applicants' ability to succeed.

If admission processes require the identification of applicants who will succeed academically, does it matter at what level those applicants will succeed? For example, if success in first-year is measured by students' ability to progress to their second year of study, by meeting their first year progression requirements, then it can be argued that any applicant who can meet this requirement can essentially succeed. For this reason, this study examined both test score (TOEFL iBT or IELTS) and academic success as dichotomous variables in which applicants either pass a given threshold or do not. This tetrachoric correlation found that only TOEFL iBT test scores above the 'minimum discretionary range' were found to be significantly related to participants' ability to succeed academically in their first year of undergraduate education. Those results suggested that although TOEFL iBT test scores meeting the 'minimum discretionary range' were not able to predict the academic success of OUAC 101 applicants, they were able to identify students who will succeed by meeting their first-year progression requirements. Additionally, those results support the use of the TOEFL iBT 'minimum discretionary range' requirement during the admission of OUAC 101 applicants in order to identify applicants who possess sufficient English language proficiency skills to succeed academically.

Again the recommendation holds that this university should develop a diagnostic assessment that may allow applicants the ability to identify their strengths and weaknesses with

regards to the English language.   During the development of this assessment, it may be beneficial to consider the time at which students will be required to complete this diagnostic. For example, although it is recommended that all students take this test prior to enrolling in their first-year courses, it may be beneficial for some students to utilize this test at multiple junctions throughout their education. For those students who achieve only slightly above the first-year progression requirement, it may be beneficial to reassess their English language skills and to reintroduce them to the services provided by the university.

6.2.2   Increasing the Minimum English Language Proficiency Requirements

The minimum English language proficiency test scores required for admission vary by university, since there is ongoing debate as to the optimal minimum test score at which applicants are able to succeed in an English-speaking university. During this debate, the idea of increasing the minimum requirements was a common suggestion. Currently the university used within this study uses their lower English language proficiency test score requirement when evaluating applications for OUAC 101 applicants whose first language is not English and who have studied for fewer than three years in an Ontario high school. This study reviewed the use of TOEFL iBT and IELTS at both the 'minimum score' and the 'minimum discretionary range' requirement in order to determine if there is any evidence for increasing the minimum English language proficiency requirement from the 'minimum discretionary range' to the 'minimum score' requirement. As previously mentioned, this study found no predictive validity of TOEFL iBT or IELTS on the academic success of the 'all participants' or the 'minimum discretionary range' subsample groups. Again although correlation and multiple regression results for the TOEFL iBT and IELTS 'minimum score' groups with GPA were presented to assist with the design of

future research studies, the sample size of those groups was too small to obtain any reliable results. Therefore the following analysis neglects the 'minimum score' subsample group.

To determine if applicants meeting the 'minimum score' performed significantly better academically in comparison to those applicants only meeting the lower end of the 'minimum discretionary range', comparisons among five TOEFL iBT and IELTS subsample groups were conducted which found no significant differences in the mean first-year GPA, or the academic average of first-year students, across those groups, with the exception of the TOEFL iBT 'minimum score' and 'minimum discretionary range' group. This independent sample t-test analysis excluded any of the 'minimum score' participants from the 'minimum discretionary range' group in order to minimize any confounding results. In practice however, using the 'minimum discretionary range' would allow applicants in both ranges to be eligible for admission based on their English language proficiency requirements and would not exclude applicants in the 'minimum score' group. Assuming that one purpose for using TOEFL iBT or IELTS during the admission process is to identify all applicants with sufficient English language proficiency, then it should be recommended that universities set their minimum English language proficiency requirements to the lowest possible score at which successful applicants are still selected. Using the lowest possible score at which an applicants' is still able to demonstrate sufficient English language proficiency should be used as to not disqualify any potentially successful applicants. Although the TOEFL iBT 'minimum score' group appeared to have a higher mean first-year GPA than the TOEFL iBT 'minimum discretionary range group, those participants meeting the 'minimum discretionary range' were still able to achieve above the first-year progression requirement of 55% with a mean group GPA of 70.14%. Therefore it can be argued that the

results of this study do not support the use of the 'minimum score' requirement over the 'minimum discretionary range' requirement when identifying potentially successful applicants.

Implications of increasing the minimum test score requirement extend beyond its use to identify applicants most likely to succeed. In order for universities to increase the diversity of their student population, universities must continue to attract the best and brightest student minds from around the world. Increasing the admission requirements above the minimum optimal test score required for admission may deter applicants with sufficient English language proficiency from applying. In addition, using the higher 'minimum score' requirement may decrease the number of eligible applicants for admission, resulting in universities not meeting their projected class, or international enrollment targets. A review of the sample sizes for each group in this study suggested that increasing the minimum requirement from the 'minimum discretionary range' to the 'minimum score' requirement would result in a decrease of approximately 13.19% of eligible applicants submitting TOEFL iBT and a decrease of approximately 41.87% of eligible applicants submitting IELTS. It can be argued that increasing the minimum English language proficiency test score requirement to the essentially arbitrary 'minimum score' requirement may, as previously noted, disqualify many potentially successful candidates from admission. For this reason it is suggested that should TOEFL iBT and IELTS continue to be used during the admission of OUAC 101 applicants, that the test score not be increased. Again in order to ensure that those applicants selected for admission based on TOEFL iBT or IELTS tests score, even those at the low end of the 'minimum discretionary range' are able to succeed, a student success focused diagnostic assessment is recommended.

6.2.3 ENG4U Grade as a Marker for Identifying Potentially Fraudulent Test Scores

Although English language proficiency test fraud is not well understood or well documented, its existence is concerning for university admission offices. In response to fraudulent test scores submitted to Ontario universities, this study aimed to provide one Ontario university with a method for identifying potentially fraudulent test scores. In doing so, ENG4U grade, which is required by all OUAC 101 applicants to this Ontario university, was compared with TOEFL iBT and IELTS test scores to determine if there was a correlation between these variables. As discussed earlier, all test scores regardless of their ability to meet the 'minimum discretionary range' requirement must be reviewed for potential fraud. Therefore, the 'all participant' group provided data most representative of the OUAC 101 applicant population at this Ontario university. This study identified the TOEFL iBT Writing subsection and all IELTS subsections as having a weak significant correlation with ENG4U grade, suggesting that a weak linear relationship existed between those significant subsections and ENG4U grade. Although a strong linear relationship would have allowed TOEFL iBT or IELTS subsections to predict ENG4U grade (i.e. high subsection test scores corresponding to high ENG4U grade), the weak relationship found amongst these variables does not. As a weak linear relationship suggests variance within the predictive relationship, deviations from the predicted relationship are expected and cannot be used to identify test scores which are not representative of the applicants' abilities. Those results were supported by weak R^2 values which indicated that many factors, other than the TOEFL iBT or IELTS subsection, were contributing to the already weak correlation. Therefore the results of this study do not support the use of ENG4U grade as a marker for identifying potentially fraudulent test scores in practice.

Developing a method for identifying fraudulent test scores is no easy task as data supporting the existence of such fraud is virtually nonexistent. Institutions would, of course, benefit from the proprietary knowledge of the English language proficiency test developers in order to protect themselves against the unethical act of fraud. Although testing companies such as ETS and the British Council do not release data which may support the existence of fraud, it can be argued that studies aiming to understand fraud would benefit those stakeholders as well. For examples, if testing centers were able to employ a method for detecting fraudulent test scores before being released to test takers and institutions, then a level of trust may develop between test centers and test users that is currently lacking. Similarly, the ability of test centers to be proactive and identify cases of fraud including impersonation may help deter those unethical events from taking place.

6.3   Suggestions for Future Research

The results of this study, although not supported by strong significant correlations, contribute to the body of knowledge surrounding English language proficiency test scores for use during the admission of OUAC 101 applicants to undergraduate education. From those results several implications were made for the use of TOEFL iBT and IELTS in practice. In addition to developing our current understanding of TOEFL iBT and IELTS, many questions arose which may build upon this exploratory study and assist in the development of future studies.

6.3.1   Identifying the Optimal Minimum Requirement

Zareva (2005) discussed a typical TOEFL PBT score of 550 which was accepted as the minimum TOEFL PBT Overall score for an undergraduate student to be proficiency in English. Currently, TOEFL iBT and IELTS test score minimum requirements vary among universities as no set requirement exists which is accepted as the minimum score for an undergraduate student to

be proficient in English. It is suggested that future analyses be conducted to isolate an optimal minimum TOEFL iBT and IELTS test score. To conduct such a study it is suggested that universities select students who fall into the following categories 1) insufficient English proficiency to succeed, 2) minimal English proficiency, but able to succeed, and 3) sufficient English proficiency to succeed. Following the categorization of students by the university, each student would write both TOEFL iBT and IELTS tests. Following the completion of those tests, test scores could be analyzed to determine if any patters emerge between student English ability, as determined by the universities categorization, and the student's test score. If this suggested analysis is conducted there is potential for those results to inform the minimum English language proficiency test score required for applicants to succeed as defined by the university as having sufficient English proficiency skills.

6.3.2   Identifying the Lowest Minimum Requirement

The current study found no predictive validity of TOEFL iBT or IELTS on academic success, as defined by first-year GPA. Even though TOEFL iBT at the 'minimum discretionary range' was able to predict an applicant's ability to meet their first-year progression requirements, this correlation was weak and may be influenced by many factors not included in this study. Therefore, admitting applicants based on the current 'minimum discretionary range' requirements may potentially disqualify applicants who are able to succeed during undergraduate education at this Ontario university. It is suggested that future research be conducted to identify if a test score exists that is able to predict academic success. Should such a score exist, it is recommended that the lowest possible version of this score be identified in order to eliminate the disqualification of potentially qualified applicants. In addition, it can be argued that English language proficiency test fraud may be due, at least in part, to the high test scores required by many Tier 1 universities.

Finding the lowest acceptable test score for which applicants are still able to succeed may, in theory, reduce the stress placed on applicants which may subsequently results in less applicants committing English language proficiency test score fraud. High test scores do not alone sufficiently explain the current amount of English language proficiency test score fraud. It should be noted that there is likely no one possible perfect test for identifying applicants with the English language proficiency required to succeed academically once enrolled as multiple factors influence an applicant's ability to succeed. It can be argued that some universities continue to use TOEFL iBT or IELTS in light of their inability to provide useful data as a risk management strategy which allows for the objective review of those applications requiring English language proficiency proof. As alternative methods to English language proficiency tests are limited, identifying the best way to use currently available test scores may prove beneficial to universities.

6.3.3   Controlling for Additional Factors

In order to identify the minimum requirement that is most successfully able to predict academic success, it is suggested that many other factors be considered including but not limited to specific subsection scores, or the nature of learning in a program an applicant has applied to. Many Ontario universities have minimum English language proficiency requirements which vary by subsection similar to that of the TOEFL iBT 'minimum score' requirement (88 Overall with a score of 22 in Speaking, Reading, and Writing, and a 20 in listening) (University X, as in References, 2015). Although the results of this study provided no support for the relationship between either TOEFL iBT or IELTS and academic success, it is suggested that the notion of different subsection requirements be taken into consideration when developing future TOEFL iBT and IELTS predictive validity studies.

It can also be argued that different levels of English language proficiency may be required for applicants to be successful in different programs. Due to the sample size of this study, program analysis was not able to be completed, however it is suggested that analysis be conducted to determine what the specific subsection minimum requirements should be, and to determine if those requirements vary by program. By way of an example, Ryerson University currently publishes minimum English language proficiency requirements that vary depending on the program. Although implementation of different requirements per program brings many practical challenges, the information gained from this future study may help admission department develop new policies and procedures.

6.4   Limitations

It can be argued that this study was limited, in part, by the use of quantitative measures alone, and may have benefitted from a mixed method design. Although quantitative studies are essential for understanding the theoretical relationship amongst variables, academic success is a complex variable which cannot be explained by one factor or a set of factor that can easily be controlled. If English language proficiency were the deciding factor in one's ability to succeed academically then it can be argued that native speakers of English would never fail. Instead, many factors together help to drive an applicant to succeed academically. Therefore this study would have benefitted from a qualitative component in which a sample of participants were interview to gain personal insight and shine some light on the results presented in this study.

Another notable limitation of this study and many others like it is the temptation to place confidence in the existing view that English language proficiency tests have the ability to predict applicants' academic success. In practice however, universities use those tests to identify applicants whom are likely to succeed, and not the level to which those applicants are able to

succeed. Rather than falling in line with the mass majority of predictive validity studies, it may be beneficial for researchers to view English language proficiency tests as a means to determine minimum level of English language proficiency. This categorical approach positions the use of English language proficiency tests in line with how those tests are used and are intended to be used.

6.5   Closing Remarks

A university is a dynamic environment in which knowledge is created and transferred amongst its faculty and students. In order to succeed in today's global world, universities must attract many of the best and brightest minds from around the world who bring with them, and/or contribute to, diversity, cultural acceptance, and new perspectives; all of which help foster a rich intellectual culture. In recruiting applicants from around the world, universities must set admission criteria that allow them to select the best candidates to contribute to an intellectual culture while also ensuring that the selected applicants possess the skills required to succeed in an English-speaking academic institution. To help select applicants with the potential to success, universities use English language proficiency tests such as TOEFL iBT and IELTS to make admission decisions regarding English language ability.

As with many high stakes tests, issues of fraud are often a concern among English language proficiency tests. The extent of research pertaining to English language proficiency test fraud is limited, which is unhelpful to universities, such as the Ontario university used within this study that have strict and high standards regarding the admission of undergraduate applicants. This study was not able to identify a reliable method for identifying TOEFL iBT and IELTS test fraud, however it is the hope of this study that those results will initiate further conversation about test fraud and subsequently encourage other researchers to conduct further exploratory

126

studies. It will not be until we, as an academic community, are able to establish a solid body of knowledge around test fraud that universities will be able to fully develop methods for entirely identifying the elusive issue that is test fraud. As no evidence is available to suggest that cases of fraud will decrease in the near future, it is suggested that a network of Ontario universities be developed to track the increasingly sophisticated efforts made by students to submit fraudulent test scores. This joint network may foster collaborative research across institutions and the identification of methods for identifying, or reducing English language proficiency test fraud while simultaneously sharing the financial burden of this project amongst all 21 publically-funded universities in Ontario.

Although this study was not able to identify the ability of TOEFL iBT or IELTS to predict the academic success of OUAC 101 applicants whose first language is not English and who have studied for fewer than three years in an Ontario high school, the need for resources which aid the admission process is recognized. As alternative methods for identifying applicants with sufficient English language proficiency skills to succeed are currently limited, it is acknowledged that universities may currently use such standardized tests as a measure or risk management. The lack of support, found within this study, for the use of TOEFL iBT or IELTS suggested that universities should consider the development and implementation of other options for assessing the English language proficiency of their applicants. To adapt to the concerning reliability of TOEFL iBT and IELTS, some universities are beginning to develop their own tests, or create programs to help students' whose first language is not English succeed once they are admitted. This study recommends the development of a diagnostic assessment to be used following the admission of applicants. The aim of this assessment should be to help new students identify their strengths and weaknesses with regards to the English language and help orient them

with the existing services offer by the university. It is expected that the future will bring with it many new innovations for assisting students' whose first language is not English which expands far beyond the use of and what is captured through TOEFL iBT or IELTS.

References

Abunawas, M. E. (2014). *A meta-analytic investigation of the predictive validity of the test of English as a foreign language (TOEFL) scores on GPA* (Doctoral dissertation).Retrieved from ProQuest.

Al-Musawi, N. M., & Al-Ansari, S. H. (1999). Test of English as a Foreign Language and First Certificate of English tests as predictors of academic success for undergraduate students at the University of Bahrain. *System 27*, 389–399.

Arcuino, C. L. T. (2013). *The relationship between the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS) scores and academic success of international Master's students.*(Doctoral dissertation).Retrieved from ProQuest.

Australian Associated Press. (2011, August 19). *Tough sentences over English visa scam.* The Australian. Retrieved from: http://www.theaustralian.com.au/news/nation/tough-sentences-over-engllish-visa-scam/story-e6frg6nf-1226118090364.

Bartlett, T., & Fischer, K. (2011, November 3). *The China Conundrum.* The New York Times. Retrieved from: http://www.nytimes.com/2011/11/06/education/edlife/the-china-conundrum.html?_r=0.

Bell, S. (2013, December 26). *Chinese women gets two years for scam that has imposters write Canadian English-language proficiency tests.* National Post. Retrieved from: http://news.nationalpost.com/news/canada/chinese-woman-gets-two-years-for-scam-that-had-imposters-write-canadian-english-language-proficiency-tests.

Brennan, R. L. (1983). *Elements of generalizability theory.* Iowa City, IA: American College Testing Program.

Brooks, L., & Swain, M. (2014). Contextualizing performances: Comparing performance during

      TOEFL iBT and real-life academic speaking activities. *Language Assessment Quarterly*,

      11, 353−373.

Cambridge ESOL (2003). *IELTS handbook*. Cambridge: Cambridge University press.

Cambridge ESOL (2006). IELTS test performance data 2004. *Research Notes 23,* 13−15.

Canadian Bureau for International Education (CBIE). (2009). Canada First: The 2009 Survey of

      International Students. Ottawa, ON. Retrieved from:

      http://files.eric.ed.gov/fulltext/ED549797.pdf.

Canadian Government. Department of Foreign Affairs and International Trade. (2012).

      Economic impact of international education in Canada − An update. Vancouver, BC.

      Retrieved from:

      http://www.international.gc.ca/education/assets/pdfs/economic_impact_en.pdf.

Celestine, C., & Ming, C. S. (1999). The effect of background disciplines on IELTS scores.

      IELTS Research Reports 2, 37−61.

Chapelle, C. A., Jamieson, J., & Enright, M. K. (Eds.) (2008). *Building a validity argument for*

      the Test of English as a Foreign Language.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT scores to academic performance:

      Some evidence from American universities. *Language Testing 29*(3), 421−442.

Citizenship and Immigration Canada. (2016). *Study in Canada: Visas, work and immigration for*

      international students. Government of Canada. Retrieved from:

      http://www.cic.gc.ca/english/resources/publications/study.asp#apply.

Cleophas, T. J., & Zwinderman, A. H. (2010). *SPSS for starters* (1st ed.). Dordrecht: Springer

      Verlag. doi:10.1007/978-90-481-9519-0.

Cohen, L. , Manion, L., and Morrison, K. (2007). Research methods in education 6th edition.

London: Routledge

Cotton, F., & Conrow, F. (1998). An investigation of the predictive validity of IELTS amongst a

group of international students at the University of Tasmania. *IELTS Research Reports 1*,

72–115.

Cummins, J. (1981). The role of primary language in promoting educational success for language

minority students. In California State Department of Education (Ed.), *Schooling &*

language minority students: A theoretical framework (pp. 3–49). Los Angeles, CA:

Evaluation, dissemination & assessment center, California State University, Los Angeles.

Darnell, D. K. (1968). *The development of an English language proficiency test of foreign*

students, using a Clozentropy procedure. (Office of Education Project No. 7-H-010).

Washington, DC: U.S. Department of Health, Education & Welfare.

DeLuca, C., Cheng, L., Fox, J., Doe, C., & Li, M. (2013). Putting testing researchers to the test:

An exploratory study on the TOEFL iBT. System, 41(3), 663-676.

doi:10.1016/j.system.2013.07.010.

Dooey, P., & Oliver, R. (2002). An investigation into the predictive validity of the IELTS test as

an indicator of future academic success. *Prospect 17(1),* 36–53.

Educational Testing Service (1994). *Guidelines for the use of TOEFL scores, 1994-1995 Edition*.

Retrieved from http://www.ets.org/Media/Research/pdf/TOEFL-SUM-9495.pdf.

Educational Testing Service (2006). The official guide to the new TOEFL iBT. New York:

McGraw-Hill.

Educational Testing Services (ETS). (2010). *TOEFL iBT research insight*. TOEFL iBT Research

Series 1, Volume 2.

Educational Testing Services (2011). *Reliability and comparability of TOEFL iBT scores*.

    TOEFL iBT Research Series 1, Volume 3.

Ezell, A. (2008) *Counterfeit diplomas and transcripts*. Washington, DC: American Association

    of Collegiate Registrars and Admissions Officers.

Feast, V. (2002). The impact of IELTS scores on performance at university. *International*

    Education Journal, 3(4), 70–85.

Fu, Y. (2012). The effectiveness of traditional admissions criteria in predicting college and

    graduate success for American and international students (Doctoral dissertation). ProQuest.

Graham, J. (1987). English language proficiency and the prediction of academic success. *TESOL*

    Quarterly, *21(3)*, 505–521.

Guhr, D. J., and Furtado, N. (2013). *Policy response to fraud and deception in international*

    education. Council of Ministers of Education, and Department of Foreign Affairs and

    International Trade Canada. Retrieved from: http://www.cbie-bcei.ca/wp-

    content/uploads/2014/11/icg-international-education-fraud-final-2013.pdf.

Hill, K., Storch, N., & Lynch, B. (1999). *A comparison of IELTS and TOEFL as predictors of*

    academic success. IELTS Research Reports, 2, 52–63.

Humphreys, P., Haugh, M., Fenton-Smith, B., Lobo, A., Michael, R., & Walkinshaw, I. (2012).

    Tracking international students; English proficiency over the first semester of

    undergraduate study. *IELTS Research Reports 1*, 1–41.

IELTS Canada. (2016). *What is IELTS?* Retrieved from:

    https://www.ieltscanada.ca/WhatIsIELTS.

Jamieson, J., Jones, S., Kirsch, I., Mosenthal, P., & Taylor, C. (2000). TOEFL® 2000

    framework: A working paper (TOEFL Monograph No. MS-16). Princeton, NJ: ETS.

Kerstjens, M., & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance. *IELTS Research Reports 3*, 85–108.

Koenig, A. M., & Devlin, E. (2012). Fighting domestic and international fraud in the admissions and registrar's offices. *College and University, 88*(1), 18.

Lane, B. (2011, March 30). *Curtin University may decide IELTS is too* risky. The Australian. Retrieved from: http://www.theaustralian.com.au/higher-education/curtin-university-may-decide-ielts-is-too-risky/story-e6frgcjx-1226030293003.

Light, R. L., Xu, M., & Mossop, J. (1987). English proficiency and academic performance of international students. *TESOL Quarterly*, *21*(2), 251–261.

Lo, J. (2002). The relationship between TOEFL scores and first-year GPA: A study of freshmen international students attending Texas A&M university-Kingsville from 1996–2001 (Doctoral Dissertation). ProQuest.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Manganello, M. (2011). *Correlation in the new TOEFL era: An investigation of the statistical relationship between iBT scores, placement test performance, and academic success of international students at Iowa State University* (Masters dissertation). ProQuest.

Moore, T., & Morton, J. (1999). Authenticity in the IELTS Academic module writing test: A comparative study of Task 2 items and university assignments. *IELTS Research Reports 2*, 74–116.

Moore, T., Morton, J., & Price, S. (2011). Construct validity in the IELTS Academic Reading test: A comparison of reading requirements in IELTS test items and in university study.

IELTS Research Reports 11, 1–89.

Nakatsuhara, F. (2011). The relationship between test-takers' listening proficiency and their performance on the IELTS Speaking Test. *IELTS Research Reports 12*, 1–50.

Ng, J. N. K. (2007). *Test of English as a Foreign Language (TOEFL): Good indicator for student success at community colleges?* (Doctoral dissertation). ProQuest.

O'Brien, J. (2014, March 11). *London police charge three people with using phony IDs to write English proficiency exams*. The London Free Press. Retrieved from: http://www.lfpress.com/2014/03/11/london-police-charge-three-people-with-using-phony-ids-to-write-english-proficiency-exams.

Ontario Ministry of Education. (2007). The Ontario curriculum grades 11-12: English. Retrieved from: https://www.edu.gov.on.ca/eng/curriculum/secondary/english1112currb.pdf.

Sawaki, Y., Stricker, L., & Orange, A. (2008). *Factor structure of the TOEFL Internet-based test: Exploration in a field trial sample*. ETS Research Report (RR-08-66). Princeton, NJ: Educational Testing Service.

Seedhouse, P., Harris, A., Naeb, R., & Ustunel, E. (2014). The relationship between speaking features and band descriptors: A mixed methods study. *IELTS Research Reports 2*, 1-30.

Simner, M. L. (1998). Use of the TOEFL as a standard for university admission: A position statement by the Canadian Psychological Association. *Journal of Psychological Assessment,* 4(3), 261–265.

Simner, M. L., & Mitchell, J. B. (2007). Validation of the TOEFL as a Canadian university admissions requirement. *Canadian Journal of School Psychology 22*(2), 182–190.

Stoynoff, S. (2009). Recent developments in language assessment and the case of four large-scale tests of ESOL ability. *Language Teaching*, *42*(1), 1–40.

Stricker, L. J., & Rock, D. A. (2008). Factor structure of TOEFL internet-based test across

subgroups. *TOEFL iBT Research Report*. Retrieved from

http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.2008.tb02152.x/epdf.

Telbis, N. M., Helgeson, L., & Kingsbury, C. (2014). International students' confidence and

academic success. *Journal of International Students, 4*(4), 330.

Time Higher Education. (2016). Work University Rankings 2015-2016. Retrieved from:

https://www.timeshighereducation.com/world-university-rankings/2016/world-

ranking#!/page/0/length/25/sort_by/rank_label/sort_order/asc/cols/rank_only.

University of Melbourne. (2016). Academic skills: Enabling students to realize their full

academic potential and achieve excellence. Retrieved from:

http://services.unimelb.edu.au/academicskills/english_as_a_second_language/DELA.

University X. (2013). *Academic Handbook*. London, ON.

University X. (2014). Strategic plan: Achieving excellence on the world stage. London, ON.

University X. (2015). Viewbook 2016. London, ON.

Wait, I. W., & Gressel, J. W. (2009). Relationship between TOEFL score and academic success

for international engineering students. *Journal of Engineering Education, 98:* 389–398.

Woodrow, L. (2006). Academic success of international postgraduate education students and the

role of English proficiency. *University of Sydney Papers in TESOL, 1,* 51–70.

Wu, K. (2011, July 14). *Ghost in the Machine*. China Daily. Retrieved from:

http://www.chinadaily.com.cn/hkedition/2011-07/14/content_12898594.htm.

Yen, D. A., & Kuzma, J. (2009). Higher IELTS score, higher academic performance? The

validity of IELTS in predicting the academic performance of Chinese students. *Worcester

Journal of Leaning and Teaching, 3,* 1–7.

York University. (2016). York English Language Test (YELT). Retrieved from:

      http://futurestudents.yorku.ca/requirements/yelt.

Zareva, A. (2005). What is new in the new TOEFL-iBT 2006 test format? *Electronic Journal of*

      Foreign Language Testing, *2*(2), 45–57.

Zhang, Y. (2008). *Repeater analyses for TOEFL iBT*. ETS Research Memorandum (RM-08-05).

      Princeton, NJ: Educational Testing Service.

Curriculum Vitae

Name:	Samia <u>Lahib</u>
Post-secondary Education and Degrees	Western University London, Ontario, Canada 2010 B.Sc
	Western University London, Ontario, Canada 2011 B.Ed
	Western University London, Ontario, Canada 2012 M.Ed Curriculum Studies
	Western University London, Ontario Canada 2016 Ed.D Educational Leadership
Related Work Experience:	Teacher and Curriculum Developer Canada eSchool Embrun, Ontario, Canada 2011 – 2012
	Student Central Assistant Western University London, Ontario, Canada 2012
	Admissions Department Western University London, Ontario, Canada 2012 – 2016