

Fairness and Justice in High-stakes Language Testing:
Test-taker Perspectives on the IELTS Speaking Test for Immigration and Higher Education in
Canada

by

Paul P. Emme

BA, University of Victoria, 2010

A Thesis Submitted in Partial Fulfillment

Of the Requirements for the Degree of

MASTER OF ARTS

in the School of Languages, Linguistics and Cultures

© Paul Emme, 2025

University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

We acknowledge and respect the Lək^wəŋən (Songhees and X^wsepsəm/ Esquimalt) Peoples on whose territory the university stands, and the Lək^wəŋən and W̱SÁNEĆ Peoples whose historical relationships with the land continue to this day.

Supervisory Committee

Fairness and Justice in High-stakes Language Testing:

Test-taker Perspectives on the IELTS Speaking Test for Immigration and Higher Education in
Canada

by

Paul P. Emme

BA, University of Victoria, 2010

Supervisory Committee

Dr. Li-Shih Huang, Supervisor,

School of Languages, Linguistics, and Cultures

Dr. Ben Pin-Yun Wang, Outside Member

Department of Pacific and Asian Studies

Abstract

The International English Language Testing System (IELTS) is a high-stakes, standardized English proficiency test. It is used by governments and post-secondary institutions when deciding whether to accept or reject applicants for immigration and admission to academic programs. This study explores test-taker perspectives on the validity of the speaking portion of the IELTS test for these purposes, using Messick's (1989) concept of unified validity, framed by McNamara and Ryan (2011) as concerning the fairness and justice of the test. Seven participants who identified as having taken or needing to take a high-stakes standardized English test for admission into a post-secondary program and/or for immigration purposes provided perspectives on the fairness and justice of IELTS through questionnaire responses and semi-structured interviews. Subsequent coding of the interview transcripts resulted in 13 top-level codes and 133 child codes, for a total of 495 child code references across all seven transcripts. Child codes were divided into four overarching themes: *Factors that Impacted Performance*, *Relevance of the Test to Academic and Everyday Contexts*, *Costs and Benefits of the Test*, and *Considering Whether There Should Be a Test, What Kind of Test, and Who Should Take It*.

Quantitative data collected from the questionnaires and the qualitative data from coding of the semi-structured interviews produced a variety of findings relevant to the ongoing discussion about the validity of IELTS as it is currently used in Canada. These include the impact of emotional factors and test topic on test performance, the importance of language assessment in both immigration and academic admission, and the need for greater flexibility in how one demonstrates English language ability.

Table of Contents

Abstract	iii
List of Tables.....	vii
List of Figures	ix
List of Acronyms.....	x
Acknowledgements.....	xi
Dedication.....	xii
Chapter 1: Introduction	1
Chapter 2: Literature Review.....	4
2.1 Research Context	5
2.1.1 Immigration and Language Trends in Canada.....	5
2.1.2 High-stakes Tests and their Stakeholders.....	6
2.1.3 High-stakes Test Use in Canadian Post-Secondary Institutions	9
2.1.4 High-stakes Test Use in the Canadian Immigration System.....	10
2.2 An Introduction to IELTS	12
2.2.1 IELTS Band Descriptions and Meaning	12
2.2.2 The IELTS Speaking Test	14
2.2.3 The Origins and Development of IELTS	16
2.3 A Brief History of Test Validation.....	17
2.3.1 Test Validation Models: Criterion, Content, and Constructs	18
2.3.2 Aspects of Construct Validity.....	20
2.4 Unified Validity.....	22
2.4.1 The Facets of Unified Validity.....	23
2.4.2 Reliability in Unified Validity.....	26
2.4.3 Language Assessment Validation Post-Messick	27
2.5 Fairness and Justice.....	28
2.5.1 Testing Fairness through Construct Validity	29
2.5.2 Testing Fairness through Relevance/Utility	31
2.5.3 Testing Justice through Value Implications.....	32
2.5.4 Testing Justice through Consequential Validity.	34

2.6 Recent Studies of Test-taker Perceptions of Fairness and Justice in High-stakes Language Testing.....	36
2.6.1 Cheng and DeLuca (2011)	36
2.6.2 Hamid et al. (2019)	38
2.6.3 Hoang (2019)	39
2.6.4 Liu and Hamid (2024).....	40
Chapter 3: Methods.....	44
3.1 Positionality	44
3.1.1 Considerations of Power and Privilege.....	44
3.1.2 Relationship to IELTS.....	46
3.1.3 Research Paradigm.....	47
3.2 Participants.....	47
3.3 Data Collection Instruments	52
3.4 Data Collection Procedures.....	55
The data collections procedures are presented in three phases: recruitment, mock IELTS speaking test administration and questionnaire completion, and finally the follow-up interview.	55
3.4.1 Recruitment.....	55
3.4.2 Mock IELTS Speaking Test Administration and Questionnaire Completion	55
3.4.3 Follow-up Interview.....	57
3.5 Data Analysis	58
3.5.1 <i>Analysis of Questionnaire Data</i>	58
3.5.2 <i>Analysis of Interview Data</i>	59
Chapter 4: Results	63
4.1 Report of Questionnaire Results	63
4.1.1 Overall agreement rate of questionnaire questions	63
4.1.2 Individual participant responses to questionnaire questions.....	71
4.2 Report of Coding Results.....	78
4.2.1 Theme 1: Describing Factors that Impacted Performance.....	79
4.2.2 Theme 2: Appraising the Relevance of the Test to Academic and Everyday Contexts.....	92
4.2.3 Theme 3: Describing the Costs and Benefits of the Test	99

4.2.4 Theme 4: Considering whether there should be a test, what kind of test, and who should take it.....	103
4.3 Summary.....	113
Chapter 5: Discussion and Conclusions.....	115
5.1 The question of fairness.....	116
5.1.1 Do the participants view IELTS as fair from the perspective of construct validity? ..	116
5.1.2 Do the participants view IELTS as fair from the perspective of relevance/utility?	118
5.2 The question of justice.....	119
5.2.1 Do the participants view IELTS as from the perspective of social consequences?	119
5.2.2 Do the participants view IELTS as just from the perspective of values?	120
5.3. Implications.....	121
5.3.1 Theoretical Implications	122
5.3.2 Methodological Implications	123
5.3.3 Practical Implications.....	123
5.3.4 Reflective Implications	126
5.4 Limitations	127
5.5 Recommendations for Future Research.....	128
5.6 Conclusion	129
References.....	131

List of Tables

Table 1	<i>Common High-Stakes English Tests Used in Canada for Immigration, Academic Enrollment, and /or International Credential Recognition.....</i>	8
Table 2	<i>CRS Points Earned by Married and Unmarried PR Applicants Per IELTS Band Level (IRCC,2025e)</i>	11
Table 3	<i>Official IELTS Description of the Type of English User Described by IELTS Band Score.....</i>	13
Table 4	<i>Official IELTS Guidelines for Acceptability of Band Scores for Language Requirements of Academic and Training Courses (British Council, 2019, p.13)</i>	14
Table 5	<i>Official Description of IELTS Speaking Test Parts (British Council, 2019, p.13).....</i>	15
Table 6	<i>Facets of Validity.....</i>	25
Table 7	<i>Understanding Messick’s Validity Matrix.....</i>	26
Table 8	<i>Participants’ Demographic Information.....</i>	48
Table 9	<i>Participants’ Background with English-language Learning and Living in English-speaking Countries.....</i>	49
Table 10	<i>Participants’ Education Background and Current Studies.....</i>	49
Table 11	<i>Participants’ Work History and Current Employment Status.....</i>	50
Table 12	<i>Participants’ History with IELTS and Other Standardized English Tests.....</i>	51
Table 13	<i>Participants’ Future Plans for Taking IELTS.....</i>	51
Table 14	<i>Top-level Parent Codes With Number of Child Codes and Total Instances of Child Codes.....</i>	60
Table 15	<i>Agreement Rate for English for Academic Purposes Question Set.....</i>	64
Table 16	<i>Agreement Rate for Immigration Question Set.....</i>	66
Table 17	<i>Perceived Relevance of the IELTS Speaking Test to Speaking Skills Required by Academic Program.....</i>	71

Table 18	<i>Perceived Relevance of the IELTS Speaking Test to Speaking Skills Required for Daily Oral Communication in Canada.....</i>	72
Table 19	<i>Perceived Appropriacy and Necessity of IELTS to Participant Academic Programs.....</i>	73
Table 20	<i>Opinions Regarding Current Minimum IELTS Speaking Scores Required for Admission into Participants' Academic Programs.....</i>	73
Table 21	<i>Perceived Appropriacy and Necessity of IELTS to Obtaining Canadian Citizenship.....</i>	74
Table 22	<i>Opinions Regarding Current Minimum IELTS Speaking Scores Required for Obtaining Canadian Citizenship.....</i>	74
Table 23	<i>Perceived Appropriacy and Necessity of IELTS to Daily English Oral Communication in Canada.....</i>	75
Table 24	<i>Perceived Impact of Factors Other Than Speaking Skills on IELTS Speaking Test Score.....</i>	75
Table 25	<i>Impact of the Need to Reach the Minimum IELTS Speaking Test Score for Admission to Academic Program.....</i>	76
Table 26	<i>Impact of the Need to Reach the Minimum IELTS Speaking Test Score for Application for Canadian Citizenship.....</i>	76
Table 27	<i>Whether Everyone Should Demonstrate English Abilities for Admission to Participant Academic Programs and Whether Alternative Forms of Demonstration Should be Accepted.....</i>	77
Table 28	<i>Whether Everyone Should Demonstrate English Abilities when Obtaining Canadian Citizenship and Whether Alternative Forms of Demonstration Should be Accepted.....</i>	78
Table 29	<i>Total Codes and References Labelled in During Focused Coding.....</i>	79
Table 30	<i>Everyday Oral English Communication Acts Identified by Participants.....</i>	94
Table 31	<i>Academic Communication Needs Described by Participants.....</i>	97
Table 32	<i>Participant Descriptions of the Emotional Impact of Testing.....</i>	100
Table 33	<i>Proposed Alternative Ways of Proving English Abilities.....</i>	110

List of Figures

Appendix 1: Codebook	138
-----------------------------------	------------

List of Acronyms

CAEL	The Canadian Academic English Language assessment
CELBAN	The Canadian English Language Benchmark Assessment for Nurses
CELP	The Canadian English Language Proficiency Index Program
CELP-G LS	The Canadian English Language Proficiency Index Program General Listening and Speaking
CLB	Canadian Language Benchmarks
CLT	Critical Language Testing
CRS	Comprehensive Ranking System
EAP	English for Academic Purposes
ELL	English Language Learner
ELTS	English Language Testing Service
ESP	English for Specific Purposes
DET	The Duolingo English Test
IELTS	The International English Language Testing Service
IELTS GT	The International English Language Testing Service General Training
IRCC	Immigration Refugees and Citizenship Canada
MELA	The Michener English Language Assessment
LINC	Language Instruction for Newcomers to Canada
OET	The Occupational English Test
L1	First / native language
L2	Second or additional language
PI	Principal Investigator
PR	Permanent Residency
PTE	The Pearson Test of English
TOEFL (iBT)	The Test of English as a Foreign Language (internet-based test)

Acknowledgements

My thanks go out first to the participants in this study for sharing their time, their insights, and their stories. To my supervisor, Dr. Li-Shih Huang: your incredible patience, rigor, and generosity, both intellectually and with your time, made this thesis possible. Since I first took your undergraduate course as part of the Diploma of Applied Linguistics, you have been an inspiration. Thank you for setting the bar so high, and for the dedication and follow-through to help me strive toward it. Thank you to my instructors in the Department of Linguistics during my first year as a graduate student for sharing your passion and expertise: Dr. John Archibald, Dr. Leslie Saxon, Dr. Marion Caldecott, Dr. Hossein Nassaji, Dr. Sandra Kirkham.

Thank you to my small but deeply inspiring Applied Linguistics cohort, Monica Connors and Kate Garrison-Riker, for sharing this brief but very meaningful time, and to my extended cohort of Linguistics classmates. Thank you to the many colleagues whom I had the pleasure to collaborate with during various phases of this project: Haleah Mclellan, Fenna Herwig, Tegan Clarke, Mina Guan, among many others.

To Dr. Ben Wang: thank you for generously agreeing to be the outside reviewer for this thesis and for several semesters as my Chinese instructor. Thank you as well to Dr. Valia Spiliotopoulos for acting as the external examiner for my oral examination and to the examination chair, Dr. Daniela Damian.

Lastly, I would like to thank my friends and family for sustaining me throughout this lengthy, and sometimes winding process. To Zoë Sager: thank you for the years of friendship and commiseration. To Kyra Nabeta: thank you for teasing me relentlessly about finishing this. As promised, I anti-dedicate this to you! And to Mom, Dad, Monica, and Abbie for everything.

Dedication

Dedicated with gratitude to Jane and Gordon Paton, to Bob and Roxanna Emme, and to John Emme.

Chapter 1: Introduction

The rise of English as the global lingua franca of business, education, and diplomacy is attached to the parallel rise of wealthy, majority English-speaking countries as major destinations for international study and for immigration. Governments in these countries, along with post-secondary institutions and professional accreditation bodies, have increasingly relied on standardized English proficiency exams to help make decisions about who to admit and who to reject. Around the world, millions of test-takers sit a standardized English exam every year, resulting in millions of dollars paid annually in exam fees and preparatory courses. The standardized English exam industry is dominated by a fairly small number of key players. By its own account, the International English Language Testing System (IELTS) is the most widely taken and most widely accepted test in the industry (British Council, n.d.).

Exams of any type make a claim to knowledge and skills in that the results of the exam are interpreted by the test-user as providing information about the test-taker. Test-users such as governments and universities then make decisions based on the perceived meaning of a test's results. As observed by Foucault (1977) and reiterated by Shohamy (2001a), a test is an instrument of power. In the case of IELTS and other standardized English tests, the decisions being made based on the result of a test-taker's performance often has major consequences for said test-taker, including access to a post-secondary education, ability to have one's previous education recognized, and indeed the ability to remain in one's adopted country permanently or claim citizenship of that country.

Standardized English tests are tools of power, used by powerful organizations to make major decisions about the lives and livelihoods of individuals, which in turn has broader

consequences for society as a whole. It is not surprising then that there have been calls to critically assess the validity of the knowledge claims made by test-users, the defensibility of the decisions made based on those claims, and the epistemological and ethical connotations of both.

As a field, test validation is uniquely positioned to address the questions of whether a test should be used and if so, what kind of test. While researchers from numerous disciplines have been assessing the validity of tests for decades, the work of psychologist Messick (1989) and subsequent language assessment researchers like McNamara and Ryan (2011) offer a framework that is especially relevant to the critical assessment of high-stakes, standardized English exams. It is a framework of test validation that asks not only what information can be gained from a test, but also what are the consequences of giving the test, and what values are present in how a test is used and interpreted. In McNamara and Ryan's words, is the test both fair and just?

In recent years, discourse surrounding the use of standardized English testing has made some room for an expansive, or in Messick's terminology unified, understanding of test validity. In particular, greater emphasis has been placed on including test-taker perspectives, experiences, and opinions when discussing the fairness and justice of a test. This has in part addressed Shohamy's (2001b) call to democratize language testing.

There are, however, gaps in the research which need to be addressed. The testing landscape has continued to change in recent decades. As the communication requirements of studying, working, and living in Canada continue to evolve, it is only fitting that critical analysis of the fairness and justice of IELTS be continuously assessed as well. Previous test-taker centred studies range considerably in terms of geographical context, test-taker population, and aspects of the test being focused on. For this reason, there is not only a need for analysis of IELTS in the

Canadian context, but also a need for analysis that focuses on the specific modalities of the IELTS test (i.e., reading, writing, listening, and speaking). The present study focuses on the speaking test, as research into test-taker perspectives on this modality of the test are limited, particularly within the Canadian context.

In Chapter 2 of this thesis, I present the context of language testing in Canada as well as some essential details of the IELTS test and how it is used in this context. This is followed by a brief history of test validation, moving from early foundational concepts to Messick's (1989) unified validity, to a description of McNamara and Ryan's (2011) reframing of unified validity in language testing as the question of fairness and justice. I then describe a number of recent studies that have addressed questions of fairness and justice (although not always explicitly using these terms) with a particular focus on centering test-taker perspectives.

In Chapter 3, I describe the methods used in the present study. This includes my positionality as a researcher, data collection instruments, data collection procedures, and data analysis procedures. This is followed in Chapter 4 with a presentation of the findings of both the questionnaires and qualitative coding of the follow-up interview transcripts. Finally, in Chapter 5, I discuss these findings in light of the research questions, followed by a description of the theoretical, methodological, practical, and reflective implications of the study as well as limitations and areas for further research.

Chapter 2: Literature Review

This literature review provides some necessary context and background for the present study as well as my theoretical framework and an overview of several relevant recent studies.

In section 2.1, I describe the current demographic and language policy trends which establishes the motivations for the study and provides information deemed important in understanding the situation faced by high-stakes language testing stakeholders.

In section 2.2, I give a brief history and overview of the International English Language Testing Service (IELTS) test, with special attention given to the speaking section of the test as the main focus of this study.

In section 2.3, I provide some historical context for the present study, situating it within the field of test validation. In this section, I explore how the strengths and weaknesses of criterion validity, content validity, and construct validity laid the foundation for Messick's (1989) unified validity, which is itself a major theoretical framework for this study.

In section 2.4, I present Messick's unified validity theory in more detail and its connection to modern language assessment. Special attention will be given to the ways in which relevance/utility, social consequences, and value judgements are included with construct validity to create a theory of test validation with the potential the potential to answer the calls "democratize" (Shohamy, 2001b) and "humanize" (Hamid & Hoang, 2018) language testing.

In section 2.5, I present McNamara and Ryan's (2011) definitions of *fairness* and *justice* and their relationship to unified validity. In this section, I explore how McNamara and Ryan's concepts of fairness and justice can be used as a lens through which the experience and expertise

of test-takers can be centred in language testing validation. I describe current test-taker centred validation research and place these studies within the larger conversation regarding the fairness and justice of high-stakes English testing in general, and the IELTS test in particular.

2.1 Research Context

Addressing the question of fairness and justice in language testing can only be done with an understanding of the context of said testing and a discussion of who the stakeholders are within this context. As the present study focuses on high-stakes, standardized English testing in the Canadian context, I present an overview current trends, stakeholders, and uses of high-stakes tests in Canada, as well as an overview of the broader social context in which these tests are being used.

2.1.1 Immigration and Language Trends in Canada

According to the 2021 Census, 23% of Canada's total population were classified as immigrants, a term which includes both permanent residents and naturalized citizens (Statistics Canada, 2023). This number is projected to reach between 24.5% and 30% by 2036 (Statistics Canada, 2017a). While the number of immigrants who speak English or French as a first or native language (L1) has dropped significantly, owing to changing trends in the countries of origin of immigrants to Canada, a large majority of immigrants to Canada report knowledge of either English or French as a second or additional language (L2) (Statistics Canada, 2017b).

Additionally, half of newcomers to Canada have a bachelor's degree or higher (Statistics Canada, 2017b). This presents a portrait of Canada's immigrant population that is not only large and growing, but also more highly educated and multilingual than the population at large. Moreover, internationally trained professionals are filling key positions in regulated fields such as

healthcare. As of 2023, 27% of doctors are internationally trained (Canadian Institute for Health Information, 2023) as are 10% of registered nurses (Crea-Arsenio et al., 2023).

The trends observed in the Canadian population at large are also reflected in the student population at Canadian academic institutions. Statistics Canada (2024) reports that of the 2.2 million students enrolled in Canadian universities and colleges during the 2022/2023 school year, 468,087 of these were international students, roughly 21.3% of total enrollments. The vast majority of these students came from countries deemed non-English speaking for the purposes of enrollment, with India, China, Nigeria, and the Philippines being the most common for international students in Canada (Canadian Bureau for International Education, 2023).

In Canada, whether an individual is attempting to acquire permanent residency (PR) or citizenship, gain admission to university or college as an L2 English speaker, or obtain recognition for international credentials, there is a high likelihood that they will be required to prove proficiency in either English or French. In practice, this means that a significant number of immigrants, international students, and internationally educated professionals will be required to take at least one high-stakes language test.

2.1.2 High-stakes Tests and their Stakeholders

Nagy (2000) identified three roles of assessment: gatekeeping, accountability, and instructional diagnosis. Standardized tests, when used as proof of English proficiency for immigration, academic enrollment, or international credential recognition perform a gatekeeping role as “assessment determines who is granted a privilege such as admission or graduation” (p. 2).

Gatekeeping assessments frequently come in the form of high-stakes tests, defined by McNamara (2000) as “tests which provide information on the basis of which significant decision

are made about candidates” (p. 133). Within this definition, two primary stakeholder groups emerge: test-takers (i.e., individuals whose performance on the test is assessed) and test-users (i.e., individuals or organizations who make decisions based on this assessment). An expanded view of the stakeholders of high-stakes testing can also include test developers, test owners, educators involved in test preparation, managers and administrators working in regulated professions, and, in the most expansive view, communities and societies which will be affected by decisions made by test-users.

When describing high-stakes language tests, it is important to consider whether the test is designed for general English, English for Academic Purposes (EAP), or English for Specific Purposes (ESP). Most standardized EAP tests claim to assess the language abilities needed in order to begin academic studies in an English-speaking environment. Most major EAP tests try to minimize the role of topic-specific knowledge and instead claim to test academic communication more broadly. Standardized ESP tests often *do* assume a level of domain-specific knowledge, but what is assessed is the ability to use this knowledge to perform language-based tasks. It is worth noting that many regulated professions require both a language test and knowledge test as part of their foreign credential recognition program. The most commonly accepted English tests in for immigration, academic enrollment, and credential recognition in Canada are given in Table 1.

Table 1*Common High-Stakes English Tests Used in Canada for Immigration, Academic Enrollment, and /or International Credential Recognition*

Test Name by Acronym	Test Description
CAEL	The Canadian Academic English Language assessment is a Canadian-made EAP test. It was purchased by Paragon Testing Enterprises in 2015, who continue to own and operate the test (Paragon Testing Enterprises, 2025a).
CELBAN	The Canadian English Language Benchmark Assessment for Nurses is a nursing ESP test developed by the Centre for Canadian Language Benchmarks and the Touchstone Institute (CELBAN, 2024)
CELPPI	The Canadian English Language Proficiency Index Program is a general English proficiency test developed for a Canadian context. It is owned by Paragon Testing Enterprises (Paragon Testing Enterprises, 2025b).
DET	The Duolingo English Test was developed by Duolingo and released in 2021. Its marketing emphasizes its use for university placement and as such may be considered an EAP test (Duolingo, 2023)
IELTS	The International English Language Testing Service is currently the most widely used standardized English test. It can be taken either as a general English test, IELTS General Training (IELTS GT), or as a test of English for academic purposes, IELTS Academic. It is sometimes referred to as a multi-purpose test as it promotes itself for use in general English, EAP, and ESP. It is owned by the British Council, IDP Australia, and Cambridge Assessment of English (British Council, 2019).
MELA	The Michener English Language Assessment is a Canadian ESP test marketed towards healthcare professionals. It is owned and operated by the Michener Institute (The Michener Institute, n.d.).
OET	The Occupational English Test is an English for specific purposes test owned and operated by Cambridge Assessment English and Box Hill Institute. It is an ESP test which markets itself as being developed for 12 healthcare professions including medical doctors, nurses, pharmacists, and so on (OET, 2025).
PTE	The Pearson Test of English is a computer-based test owned and operated by Pearson PLC. PTE versions currently available are Academic, Core (a general English test), and several UK visa tests (Pearson PLC, n.d.)

Table 1 (continued)

TOEFL (iBT)	The Test of English as a Foreign Language (internet-based test) is arguably the main competitor of IELTS in its use as a globally available, multi-purpose English proficiency test. It is owned and operated by the English Testing Service (ETS, 2024).
-------------	---

2.1.3 High-stakes Test Use in Canadian Post-Secondary Institutions

Students applying for admission to a primarily English-medium Canadian university or college are typically required to demonstrate English-language proficiency when their previous education was completed in a non-English speaking country. While most major Canadian universities accept several of the tests listed in Table 1, IELTS markets itself as the most popular English test globally for education (British Council, n.d.).

While the minimum accepted score (also known as the cut score) for admission will depend on the institution and program, the most widely accepted minimum on the IELTS Academic test is an average of six point five out of nine in all four domains and no individual domain below a six (MacDonald, 2019). In most cases, applicants must present scores from a test taken within the last two years. This period of institutional acceptability is sometimes referred to as the test's *shelf life* (Hamid et al., 2019). Alternatively, many institutions offer EAP preparatory courses which allow students to enroll with a lower score and transition into their desired program upon completion of additional language training. However, the additional time and cost associated with these programs has led to Johnson and Tweedie (2021) to label student desire to “IELTS-out” (i.e., achieve the minimum IELTS score and bypass additional coursework) a wide enough phenomenon to deserve its own term.

2.1.4 High-stakes Test Use in the Canadian Immigration System

Immigrating to Canada is a multi-step process. It will typically involve initial acceptance into Canada via a temporary visa or attempt to claim refugee status. For those wishing to remain in Canada indefinitely, an individual will need to apply for and receive PR status. Gaining PR status is a necessary prerequisite for applying to become a naturalized Canadian citizen (Immigration Refugees and Citizenship Canada [IRCC], 2025a). Both PR and Citizenship applications can be costly and time-consuming, and both have language proficiency requirements which IELTS General Training (IELTS GT) can be used to fulfill. However, there are meaningful differences between PR and Citizenship applications both in how IELTS is used and which alternatives to IELTS are available to applicants.

Permanent Residency. For individuals residing in Canada with a temporary visa or refugee status, achieving PR status can offer a meaningful sense of stability and security. However, gaining PR status is competitive. While there are multiple routes to PR status, for most applicants, the process involves providing information such as age, education, work experience in Canada, and English or French language ability. Applicants are awarded points as part of the Comprehensive Ranking System (CRS) based on these and other factors. Multiple times per month, “top-ranking candidates” are invited to apply for PR status using CRS points as a deciding factor (IRCC, 2025b). For many pathways, a language test with a minimum score is required to apply, but higher scores will result in more CRS points. Currently, CELPIP General, IELTS GT, and PTE Core are the accepted as proof of English-language abilities (IRCC, 2025c). IRCC (2025d) lists the points awarded per testing modality (i.e., reading, writing, listening, and speaking) using the Canadian Language Benchmarks (CLB) (Centre for Canadian Language Benchmarks, n.d.). Table 2 shows CRS points awarded per IELTS band score using government-

endorsed IELTS equivalents (IRCC, 2025e). Note that PR applicants are awarded different amounts of points depending on whether or not they have a spouse or common-law partner.

Table 2

CRS Points Earned by Married and Unmarried PR Applicants Per IELTS Band Level (IRCC, 2025e)

IELTS GT Testing Domain	IELTS Score by Equivalent CLB Level						
IELTS GT Reading	3.5	4.0	5.0	6.0	6.5	7.0	8
IELTS GT Writing	4.0	5.0	5.5	6.0	6.5	7.0	7.5
IELTS GT Listening	4.5	5.0	5.5	6.0	7.5	8.0	8.5
IELTS GT Speaking	4.0	5.0	5.5	6.0	6.5	7.0	7.5
PR Applicant Marital Status	CRS Points Earned						
With a spouse or common-law partner (Maximum 128 points)	0	6	8	16	22	29	32
Without a spouse or common-law partner (Maximum 136 points)	0	6	9	17	23	31	34

Note. CLB = Canadian Language Benchmarks; CRS Comprehensive Ranking System; IELTS GT = International English Language Testing System – General Training; PR = Permanent Residency.

It is worth noting that whichever test the applicant provides, it must have been taken within the last two years and cannot be substituted for different versions of the test (i.e., an IELTS Academic test cannot be used in place of an IELTS GT).

Citizenship. Permanent residents who have been present in Canada for three of the last five years may apply for Canadian citizenship (IRCC, 2025f). Applicants who are between the ages of 18 to 54 must complete a citizenship test and prove proficiency in English or French. Proof can include attendance at an English or French language secondary or post-secondary institution (IRCC, 2025g). Alternatively, applicants can complete a recognized language course such as the federally funded Language Instruction for Newcomers to Canada (LINC), achieving a CLB Level 4 or higher in speaking and listening (IRCC, 2025h). Finally, applicants can also fulfill the language requirement for citizenship by achieving a minimum speaking and listening

score, equivalent to CLB Level 4, in either the IELTS GT or the CELPIP General Test (or CELPIP-G LS which is only the listening and speaking components).

Somewhat unusually for standardized-testing requirements, as of the time of writing, there are no restrictions on how recently the IELTS or CELPIP scores were acquired, meaning that a score achieved more than two years ago could still be used to apply for citizenship, although as stated earlier, academic versions of these tests are not accepted.

2.2 An Introduction to IELTS

IELTS is an English language proficiency test jointly managed by the British Council, IDP Australia, and Cambridge Assessment of English. It promotes itself as “the world’s most popular English language proficiency test for higher education and global migration” (British Council, n.d.). It is comprised of four sections: reading (60 minutes), writing (60 minutes), speaking (11-14 minutes), and listening (30 minutes). As mentioned earlier in Table 2, there are two versions, or “types” of the test available: IELTS Academic and IELTS GT. The listening and speaking sections are identical in these two types; the reading and writing sections in IELTS Academic purport to be reflective of the kinds of reading and writing skills needed for study at an English-speaking university, while the IELTS GT attempts to measure “English language proficiency in a practical, everyday context. The tasks and texts reflect both workplace and social situations” (IELTS, 2019, p. 3).

2.2.1 IELTS Band Descriptions and Meaning

Strictly speaking, one does not pass or fail an IELTS test. Test results are provided on a scale of 1-9 in each of the four sections of the test. These are called “bands” in

IELTS terminology, and test-takers receive their results in the form of four band scores, one for each of reading, writing, speaking, and listening. The four band scores received by a test-taker result in an overall average which many institutions will also consider when make score-based decision. Table 3 shows the score meanings for test-takers receiving each of the nine bands according to IELTS (2019). Note that scores ending in “.5” are also possible as well as a “0” score for registered test-takers who did not submit a score for the test or a section of the test.

Table 3

Official IELTS Description of the Type of English User Described by IELTS Band Score

Band score	Type of “English user”	Description
9	Expert user	Has fully operational command of the language: appropriate, accurate and fluent with complete understanding.
8	Very good user	Has fully operational command of the language with only occasional unsystematic inaccuracies and inappropriacies. Misunderstandings may occur in unfamiliar situations. Handles complex, detailed argumentation well.
7	Good user	Has fully operational command of the language, although with occasional inaccuracies, inappropriacies and misunderstandings in some situations. Generally handles complex language well and understands detailed reasoning.
6	Competent user	Has generally effective command of the language despite some inaccuracies, inappropriacies and misunderstandings. Can use and understand fairly complex language, particularly in own field.
5	Modest user	Has partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field.
4	Limited user	Basic competence is limited to familiar situations. Has frequent problems in understanding and expression. Is not able to use complex language.
3	Extremely limited user	Conveys and understands only general meaning in very familiar situations. Frequent breakdowns in communication occur.
2	Intermittent user	No real communication is possible except for the most basic information using isolated words or short formulae in familiar situations and to meet immediate needs. Has great difficulty understanding spoken and written English.

Table 3 (continued)

1	Not user	Essentially has no ability to use the language beyond possibly a few isolated words.
0	Did not attempt the test	No assessable information provided.

As mentioned earlier, decisions made based on band scores will vary by institution and use. While *how* these standards are set is an area of discussion within test validation studies (e.g., MacDonald, 2019), IELTS does provide its own general guidelines. Table 4 shows the official IELTS guidelines for the acceptability of band scores for the language requirements of academic and training courses.

Table 4

Official IELTS Guidelines for Acceptability of Band Scores for Language Requirements of Academic and Training Courses (British Council, 2019, p. 13)

Band score	Linguistically demanding academic courses	Linguistically less demanding academic courses	Linguistically demanding training courses	Linguistically less demanding training courses
7.5-9	Acceptable	Acceptable	Acceptable	Acceptable
7	Probably acceptable	Acceptable	Acceptable	Acceptable
6.5	English study needed	Probably acceptable	Acceptable	Acceptable
6	English study needed	English study needed	Probably acceptable	Acceptable
5.5	English study needed	English study needed	English study needed	Probably acceptable

2.2.2 The IELTS Speaking Test

The four modalities of the IELTS test, reading, writing, speaking, and listening, each present their own unique challenges to test-takers. Detailed breakdowns of the tasks included in the four sections of the IELTS test are available to test-takers, as are study aids, practice tests, and general

information about marking, although the specific rubric used by raters in the speaking and writing sections are kept confidential. As the present study is focused on the speaking test, a general outline of the tasks, procedures, and published marking guides will be helpful. Table 5 shows the official description of the three parts of speaking test.

Table 5

Official Description of IELTS Speaking Test Parts (British Council, 2019, p. 13)

Part number and length	Official part description
Part 1 (4-5 minutes)	Test takers answer general questions about themselves and a range of familiar topics, such as their home, family, work, studies and interests.
Part 2 (3-4 minutes)	Test takers are given a card which asks them to talk about a particular topic. They have one minute to prepare before speaking for up to two minutes. The examiner may then ask one or two questions on the same topic to finish this part of the test.
Part 3 (4-5 minutes)	Test takers are asked further questions connected to the topic in Part 2. These questions give the test taker an opportunity to discuss more abstract issues and ideas.

Test-takers are marked on fluency, grammatical accuracy, lexical resource, and pronunciation, but receive a single band score for the speaking section of the test based on an average of these four marks. Specific information about the test topics and, as previously mentioned, the exact rubric used for grading, are not shared with the public. However, IELTS has published information outlining the four criteria that test-takers are graded on and the meaning of each band score as it relates to these criteria (see IELTS, n.d.).

2.2.3 The Origins and Development of IELTS

The current IELTS test reflects developing attitudes and trends in language teaching and testing. Understanding the history of IELTS and the aims of its development is useful when analyzing the test critically. IELTS was developed in the 1980s by the British Council and the University of Cambridge Local Examinations Syndicate (now Cambridge Assessment of English). These organizations collaborated on an earlier test called the English Language Testing Service (ELTS), in many ways a predecessor to the IELTS. ELTS, whose primary function was to assess the English proficiency of international students receiving awards to study in the United Kingdom, was heavily influenced by the growing importance of the communicative approach to language teaching and the increasing popularity of English for Specific Purposes (Read, 2022). Based on the work of Canale and Swain (1980), the communicative approach prioritizes communicative competence over explicit linguistic knowledge and places real-world language use as the ideal to be striven for in language instruction and assessment. This influence can be seen in the later IELTS test as well. The role of English for Specific Purposes (ESP) however, proved logistically difficult for the long-term viability of ELTS, which had five domain specific versions (representing major fields of study for test-takers) and one general version.

The combination of logistical issues along with some evidence for superior validity in the general components of the test was the catalyst for the development of a more multi-purpose test of Academic English. IDP Australia became the third partner in the ELTS Revision Project and the first IELTS exams were administered in 1989. Weir and O'Sullivan (2017) summarize both the original aims as well as the subsequent

development and promotion of IELTS in this way: “The history of IELTS after 1989 is really all about how the original ELTS was turned into a valid, psychometrically respectable, high-stakes test that could be administered globally on an industrial scale” (p. 220).

While the global administration and industrial scale elements of IELTS history have undeniably come to pass, the validity and psychometrical respectably of IELTS have been an ongoing discussion. This is at least in part owing to the complexity of validating a test like IELTS multiplied by the ever-increasing range of contexts, uses, and stakeholder groups of the test. While no comprehensive review or revamp of IELTS has been made public by the test’s owners (Read, 2022), a number of studies (both independently undertaken and officially funded by the test-developers) have been conducted, exploring questions of the test’s validity. The bulk of these have been focused on stakeholders who took the test for university admission. A smaller but growing body of work has emerged focusing on stakeholders who took the test for immigration purposes or professional credential recognition. However, stakeholders do not always neatly fall into one category. It is not unusual, for example, for a skilled immigrant to take the test as part of the immigration proses while also working towards official recognition of credentials (e.g., Frost, 2017). The growing number and variety of IELTS-focused validation studies are situated within a larger history of the field of test validation. It is this history in which the present study is also situated.

2.3 A Brief History of Test Validation

Questions of validity are of keen interest to a number of IELTS stakeholders. While the term validity can be defined in various ways, Messick’s (1989) definition will serve for

the present study: “Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment” (p.5, emphasis in original).

For the owners of IELTS, the *perceived* validity of the test (also known as face-validity) is paramount for its continuing marketability. For test-users, the confidence with which they make decisions based on IELTS band scores will be directly impacted by the test’s validity. Of course, it is test-takers themselves who arguably have the most at stake when it comes to the quality and validity of a test. How test-takers view the test in this regard, and how they have already been impacted by the use of high-stakes test is central to the questions of this study. The following section will explore both the theory and practice of test validation and the role of test-takers in this process.

2.3.1 Test Validation Models: Criterion, Content, and Constructs

The theoretical foundations of the present study place it within the larger, interdisciplinary field of test validation. The history of twentieth century test validation shows a trend not of dismissing older forms of test validation, but rather adding to them to create or more complex, multifaceted understanding of whether a test is “valid” or not. The major development presaging the current study is the movement from criterion validity, to content-based validity, to construct validity.

Criterion validity was among the preferred approaches to test validation in psychology and education in the first half of the twentieth century. Criterion validity is asserted through a correspondence between the test being validated and a separate, measurable criteria. Or, in the words of Kane (2001): “how well scores estimated or predicated the criterion scores” (p. 319).

While useful, and still used today, criterion validity does have shortcomings that necessitated critique. As test validation often seeks to establish a correlation between the test and the “real” variable the test is trying to measure, researchers found they either had to use criteria that were themselves removed from the “real” variable or they had to use variables that could not be operationalized as true criteria in the established sense. For example, criterion validity can be used to predict TOEFL (iBT) scores based on IELTS scores but claiming that sufficient English proficiency for a given real-world context is predicted based on IELTS score requires something else.

Content-based validity models were proposed for instances where no criteria existed or was possible. Rather than using external criteria to validate a test, the content-based validity model relies on “a review of the test content by subject-matter experts” (Angoff, 1988, p. 22, cited in Kane, 2013, p. 449) This model asserts that a test is valid insofar as its contents are representative of the domain of the test, as supported by subject-matter experts. Content validity solved some of the problems remaining from the criterion model but was itself criticized for being overly subjective and for subject-matter experts’ tendency to prefer assessment methods and standards that were already in use. Like the criterion-based model, the content-based model has some utility, but could not be relied on as the sole argument for a test’s validity.

Construct validity emerged as a response to the perceived inadequacies of criterion and content based models. However, construct validity is not a rejection of earlier models, but rather a reframing of the focus of validation which can (and should) be investigated using multiple tools, often including criterion and/or content based forms of validation. A construct is “an attribute, proficiency, ability, or skill that happens in the human brain and is defined by established theories” (Brown, 2000, p. 9). Test results are interpreted as indicating the extent to

which the test-taker possesses the construct being tested. Or, in the words of Cronbach and Meehl (1955), two important figures in the development of construct theory: “Here the trait or quality underlying the test is of central importance, rather than either the test behavior or the scores on the criteria” (pg. 176). The challenge with giving traits or qualities a central place in validation is that they cannot be observed directly; they can however be inferred from observed behaviour, as in a test. Construct validation in testing is therefore an investigation into interpretations made from test performance. While criterion and content validity are often expressed mathematically, when investigating construct validity, it has also become standard practice to take an argument-based approach, that is, using evidence and logical inferences to support or reject interpretations of test results.

2.3.2 Aspects of Construct Validity

The development of construct validity may have addressed some of the philosophical issues left unanswered by earlier models. However, construct validity does provide unique challenges to researchers working within test validation. As stated earlier, mental constructs are by their nature not directly observable, and any evidence used to make inferences about mental constructs must necessarily be based on an incomplete picture of the test-taker as a whole. Messick (1996) proposed six aspects of validity, five of which dealt directly with construct validity, which are useful when considering the problem of how to approach testing construct validity. These aspects are not meant to be seen as mutually exclusive, making it possible to collect multiple sources of data and addressing multiple aspects when considering construct validity. The sixth aspect of validity, *consequences*, will be addressed in the next section. Messick’s five aspects of *construct* validity are as follows:

Content aspect of construct validity refers to “determining the knowledge, skills and other attributes to be revealed by the assessment tasks” (Messick, 1996, p. 249). Essentially, what constructs are necessary and sufficient for the purposes laid out by the test-user?

Substantive aspect of construct validity refers to the “theoretical rationales for observed performance regularities and item correlations” (Messick, 1996, p. 240). It describes whether constructs are appropriately sampled from the domain and whether the target constructs are actually engaged or whether test performance could be better explained by a different construct entirely. An example that is much discussed in standardized testing is test-wiseness, in which constructs related to taking tests are arguably responsible for performance rather than the actual constructs being tested.

Structural aspect of construct validity refers to “the fidelity of the score scales to the structure of the construct domain at issue” (Messick, 1996, p. 248). In other words, do the test’s tasks and rating scales measure the constructs effectively? Constructs may be very well selected for the target domain, but poorly designed tasks and rating scales will impede the test-user from making meaningful inferences from scores or other measures of performance on the test.

Generalizability aspect of construct validity refers to “the extent to which score properties and interpretations generalize to and across population groups, settings and tasks” (Messick, 1996, p. 248). Messick connects generalizability to score meaning. When test-users interpret a score, they are in effect making a generalization based on a limited piece of information. A score interpretation that is highly successful for making generalizations is an argument in favour of the construct validity of the test.

External aspect of construct validity refers to “the extent to which the assessment scores’ relationships with other measures and nonassessment behaviours reflect the expected high, low, and interactive relations implicit in the theory of the construct being assessed” (Messick 1996, p. 251). This is an aspect of construct validity in which criterion-based validation is still very useful. Studies that compare IELTS scores to university performance are investigating construct validity through the external aspect. Although usually not considered sufficient by itself, this aspect can strengthen the case for overall construct validity.

2.4 Unified Validity

While construct validity proved to be a major theoretical breakthrough in test validation, it is now widely viewed as one aspect of a more expansive view of validity. The work of Messick in the 1980s and early 1990s was a major inspiration for this shift in perspective. Messick’s own definition of validity, quoted earlier, is helpful in understanding why construct validity alone cannot be viewed as synonymous with validity. In particular, Messick’s (1989) inclusion of the “*adequacy and appropriateness of [both] inferences and actions*” (p.5, emphasis in original) in his definition is significant.

Earlier definitions of test validation might be summarized as asking: is the test *telling* us what we want to *know*? But, by including both inferences and actions in his definition of validity, Messick reframed validation in a way that could be rephrased as: is the test *doing* what we want it to *do*? This framing includes not only the earlier focus on construct validity, but it now also challenges the test-user to question the consequences of the test and the ethical motivations and ramifications of giving the test. If test-users (or individuals living in a society in which a test is used) find the information from the test highly useful, but find the *way* the test is being used to be ineffective, ethically dubious, or actively harmful, the test may not be doing what we want it

to do. Of course, who counts as “we” in this framing and whose needs and desires are considered is an issue that needs to be addressed. However, an expanded understanding of validity has the potential to place this issue within validation research rather than outside of it.

2.4.1 The Facets of Unified Validity

Messick’s (1989) unified validity is his attempt to combine the facets of an expanded vision of validity into a single theoretical framework. In this framework, Messick proposes two bases for claims of test validity: the evidential basis (justifying claims based on empirical evidence) and the consequential basis (justifying claims based on intended and unintended consequences). Messick argues that evidential basis and consequential basis are relevant when investigating both how a test is interpreted and how a test is used.

The evidential basis of test interpretation, Messick argues, is essentially equivalent to construct validity. It is “an integration of any evidence that bears on the interpretation or meaning of test scores” (Messick, 1992, p. 1491). A test can be said to have a high level of construct validity if inferences made based on the score are supported by evidence that the score is a result of the test-taker having whatever knowledge or skill corresponds to the construct being tested and not construct-irrelevant factors such as test familiarity or rote memorization. For example, if a university infers that a high score on an English test implies the test taker has the linguistic skills necessary to succeed in an English-medium program, a test for construct validity might involve discounting other factors such as those mentioned previously, which would undermine the strength of the inference.

The evidential basis of test use includes construct validity with the addition of relevance/utility for the applied use of the test (Messick, 1992). A test can be shown to enable

highly accurate inferences about a certain construct, but this does not necessarily imply that this information is relevant to the purpose for which it is being administered. Previous media attention given to the subject of non-academic uses of IELTS has focused on this facet of test validity (e.g., Collins, 2015). The criticism here is not that IELTS is not successful at enabling inferences of English proficiency within a certain domain of use, what Messick (1992) calls the *construct universe*, but that these inferences are not relevant or useful in the actual domain for which the test is being implemented.

The consequential basis of test interpretation is “the appraisal of value implications of score meaning” (Messick, 1992, p.1494). This is a significant facet of test validity as qualities considered *desirable* or *undesirable* are highly contingent on frequently opaque social values. For example, in a productive domain like speaking, a preference for directness or clarity reflects values that need not be assumed to be universal. To address this, researchers must seek to uncover how the inferences made from the results of a test are influenced by the values inherent in how the test is scored or graded.

The consequential basis of test use is “the appraisal of both potential and actual social consequences of applied testing” (Messick, 1992, p. 1494). In establishing a consequential basis for using a test, the test-user must consider negative social consequences that may result from the test being used rather than a different method of assessment or no assessment at all. Critics of using academic standardized tests outside of academic contexts challenge not only construct validity, relevance/utility, and value implications, but also point to the effects these tests have on the examinees as individuals (Collins, 2015), as members of often vulnerable or marginalized groups (McNamara & Ryan, 2011), or on society as a whole (Tapper, 2017).

Taken together, the evidential and consequential bases of test interpretation and use can be shown as in Table 6, Messick’s Facets of Validity matrix. The argument in favour of unified validity is that it acknowledges the functional worth of a test needs to take into account both the predictive power of the information the test provides and the consequences of the application of the test.

Table 6

Facets of Validity

	<i>Test Interpretation</i>	<i>Test Use</i>
Evidential Basis	Construct Validity (CV)	CV + Relevance/Utility (R/U)
Consequential Basis	CV + Value Implications (VI)	CV + R/U + VI + Social Consequences

(Messick, 1989, p. 10)

Unified validity was proposed as part of an ongoing theoretical discussion regarding the philosophical problems arising from tests in general, touching on both ethics and epistemology. As such, Messick’s original language tends to be phrased in the abstract. This has led subsequent researchers to reinterpret and rephrase the facets of validity to suit their particular needs. Working within the field of language assessment, McNamara and Roever (2006) rephrased the matrix as can be seen in Table 7.

Table 7

Understanding Messick's Validity Matrix.

	<i>What test scores are assumed to mean</i>	<i>When tests are actually used</i>
Using evidence in support of claims: test fairness	What reasoning and empirical evidence support the claims we wish to make about candidates based on their test performance?	Are these interpretations meaningful useful and fair in particular contexts?
The overt social context of testing	What social and cultural values and assumptions underlie test constructs and hence the sense we make of scores?	What happens in our education systems and the larger social context when we use tests?

(McNamara & Roever, 2006, p.14)

2.4.2 Reliability in Unified Validity

Messick's unified validity is clearly an expansion of what has traditionally been labelled validity; its enduring influence in language testing validation speaks to its power and utility. However, one aspect of language testing research whose place within unified validity remains something of an open question is reliability. Researchers such as Moss (1994) and Carlsen (2007) have argued that much of what has traditionally been referred to as *reliability* in testing research could be framed as part of Messick's description of construct validity. Although reliability, which Carlsen (2007) defines as "the degree of stability or consistence of a measurement instrument" (p. 1), is sometimes considered a thing wholly separate from validity, issues such as inter-rater reliability (the extent to which two raters give similar scores for similar performances) and rater-internal reliability (the extent to which one rater gives two similar scores for similar performances) clearly have an impact on generalizability.

Separating validity from reliability is challenging because it frequently demands that we make distinctions between construct selection, task design, rating scale, prescribed test procedures, and actual test implementation that are not always easy to separate. Carlsen herself suggests that whether one chooses to think of reliability and validity as separate or unitary may best be determined by the focus of the individual researcher. Because of the aforementioned complexity, and given Messick's (1995) own statement that "all factors that influence the suitability of scores as a basis for inferences and actions could... be classified together as sources of variance" (p. 744), I will consider issues pertaining to reliability alongside construct validity as related aspects test fairness.

2.4.3 Language Assessment Validation Post-Messick

While Messick's proposed unified validity model has been highly influential in the field of language assessment validation, the question of how to apply the theory has resulted in a number of competing (although not necessarily mutually exclusive) approaches. These tend to take an argument-based approach to establish construct validity but are expanded to include both test meaning and use as advocated by Messick. These include the Test Usefulness Framework (Bachman & Palmer, 1996), the Fairness Framework (Kunnan, 2004), Assessment Use Argument (Bachman & Palmer, 2010) and the Fairness Argument (Xi, 2010). These approaches allow researchers to address a test as a whole, using the evidential and consequential bases of interpretation and use, thereby making a validity argument which is unified.

While argument-based approaches to test validation have been put to use, they tend to be highly systematized and are not appropriate for every question related to test validation. A more open-ended framing, still rooted in established theory, can garner useful information in its own right, and potentially provide support as part of future validity arguments as well. McNamara and

Ryan's (2011) proposal of expressing unified validity in terms of *fairness* and *justice* provides a balance theoretical grounding and flexible applicability.

2.5 Fairness and Justice

For McNamara and Ryan (2011), fairness refers to “the extent to which the test quality, especially its psychometric quality, ensures procedural equality for individual and subgroups of test-takers and the adequacy of the representation of the construct in test materials and procedures” (p. 163). In Messick's terminology, fairness is the evidential basis of test interpretation and use. When questioning the fairness of a test, we are asking: does the test effectively assess the desired construct and is this representative of the real life situations in which the construct is deemed important? For a test to be deemed *fair*, there should be strong evidence of construct validity and strong evidence that the test is relevant and useful for the context in which it is used.

Justice, then, corresponds to the consequential basis of test interpretation and use. It is concerned not with the technical quality of the test but rather “the values implicit in test constructs, and the social uses to which language tests may be put” (McNamara & Ryan, 2011, p. 161). For a test to be deemed *just*, it should result in beneficial consequences as a result of the values implicit in what scores are interpreted to mean and there should be beneficial consequences as a result of using the test.

As can be seen, establishing fairness and justice in language testing requires a combination of gathering and analyzing information about the test and target language domain as well as documenting and discussing the impact and sociopolitical context of the test in the real

world. The following sections explore the ways in which researchers in the field have attempted to “test the test” in terms of fairness and justice.

2.5.1 Testing Fairness through Construct Validity

Testing fairness through construct validation requires an articulation of the construct being tested (and a justification for its inclusion), a plausible argument that the test as interpreted can make claims about the construct, and established tools for addressing these claims. It is also necessary to investigate threats to construct validity, principally construct-irrelevant variance and construct under-representation. McNamara (2000) defines construct-irrelevant variance as the “difference in the performance of the candidates on a test which are unrelated to the ability or skill being measured” (p. 132) and construct underrepresentation as the oversimplification of the construct. Construct validation research employs a wide range of methodologies including quantitative, qualitative, and mixed methods. The role that stakeholder perceptions play in construct validation varies depending on the aspect of construct validity being examined.

Domain experts such as administrators, educators, and working professionals within the test domain have frequently played an important role when addressing what Messick called the *content aspect* of construct validity. For example, Manias and McNamara (2016), using thematic analysis of medical and nursing clinical educators’ evaluation of domain-specific role-plays, found not only considerable overlap with constructs deemed important in the domain and what was currently assessed in the Occupational English Test (OET), an ESP test designed for various medical professions. However, the authors also found areas where the test may be suffering from construct underrepresentation, for example, the importance of English as a Lingua Franca (ELF) interactions in a clinical setting.

Domain experts can also provide useful information when validating score meaning, an important part of the *generalizability aspect* of content validity. A cut score (also written as cut-score), the minimum allowable score with which a test taker has “passed” the test, is one of the strongest and most direct statements of test interpretation. Domain experts can provide useful information in this aspect of construct validity as well. O’Neill et al. (2007) is an example of the use of the Analytic Judgement method in establishing a cut score through domain experts’ perceptions of the language needed in their fields. The Analytic Judgement method “involves determining distinctions between performance levels and then computing the averages of scores near performance-level boundaries to determine final cut scores” (Manias & McNamara, 2016, p. 236).

A similar method was used in Murray et al. (2014) in which school principals were asked to make yes/no decisions about whether they would offer employment to an IELTS test-taker based on their performance on the speaking and writing test. Murray et al.’s study was additionally useful as subsequent coding of principals’ justifications for their decisions provided information about which constructs they deemed necessary to work in a school setting. While Murray et al. (2014) is an example of stakeholder perceptions being consulted for cut score appropriacy and the content aspect of construct validity, there continues to be a need for greater consultation not only with test-users such as school principals, but also test-takers as in Hoang’s (2019) study with test-takers in the context of the Australian immigration system.

Another major question when addressing the fairness of a test is the prevalence of construct-irrelevant variance. Construct-irrelevant variance is among the more widely studied aspects of construct validity in high-stakes language testing, and with good reason. Any test item which claims to assess a construct but whose results are meaningfully affected by something

other than the knowledge, skill, or ability represented by the claim has introduced construct-irrelevant variance. Of special importance for global tests like IELTS, where the number of tests and testing locations is so varied, is reducing construct-irrelevant variance by ensuring procedural fairness, which can include factors such as testing environment, behaviour of test-centre administrators, and behaviour of examiners in the case of the speaking test. Test-takers offer a uniquely authentic perspective on these issues as can be seen in test-taker centred studies by Cheng and DeLuca (2011) and Hoang (2019).

2.5.2 Testing Fairness through Relevance/Utility

While construct validity refers to a test's ability to assess a mental attribute via an interpretation of test performance, relevance/utility refers to the question of whether a given construct *should* be assessed for the stated purpose. Assessing relevance/utility requires test-developers or test-users to articulate a justification for a construct's inclusion within the test given the stated purpose. In the words of Fulcher and Davidson (2009) "unless intended score meaning is explicitly and carefully linked to test design, it is extremely difficult to demonstrate the link between the users' interpretation of the score and the decisions that they take on the basis of the score" (p. 125).

Relevance/utility is an important facet of the fairness of tests generally, but when tests are applied outside of their original stated use, the risk of a loss of coherence in the argument for the use of the test grows. Fulcher and Davidson (2009) refer to the process of applying a previously developed test to a new context as a *change retrofit*. A change retrofit does not necessarily mean a loss of relevance/utility, but it would require evidence that the constructs tested are sufficiently representative of the language needs of the new domain to justify inferences about the test-takers and allow test-users to make decisions based on those inferences. For their part, Fulcher and Davidson (2009) argue that "it is unlikely that any such change retrofits could be conducted

without making serious alterations to the construct framework” (p. 136). That being said, change retrofits are a fact of high-stakes testing, including the IELTS, so seeking out evidence for or against claims of the relevance / utility of the test’s use for various purposes will be important in establishing the fairness of using the test in these ways.

Relevance/utility can be addressed in a number of ways. Discourse analysis allows researchers to identify language used in the target domain and compare their findings with the test (Sedgwick & Garner, 2017) . Questionnaires, focus groups, and interviews are also popular ways of investigating relevance/utility. Previous studies include those focused on test-users like administrators within professional bodies (Merrifield, 2016), domain experts (Murray et al., 2014), and test-takers in several contexts. Test-taker perceptions of relevance/utility have been collected for IELTS for university admission (A. W. Ata, 2018; Liu & Hamid, 2024; Rajendram et al., 2019), IELTS for immigration (Hoang, 2019; Merrylees, 2003), and IELTS for professional credential recognition (Gribble et al., 2016; Read & Wette, 2009; Rumsey et al., 2016). There is a strong argument for including test-taker perceptions of the relevance/utility of a test when addressing test fairness. This is not only because the stakes for test-takers are so high, but also, as Cheng and DeLuca (2011) have pointed out, because test-takers are frequently domain experts themselves and have valuable insights into the language needs of the domain for which they are being tested.

2.5.3 Testing Justice through Value Implications

McNamara (2013) summarizes the inquiry into value implications of tests with the question “whose values are represented in the test, and whose interests are advanced through its existence and operation?” (p. 1). Understanding the values implicit within the test necessitates exploring the power relationships that exist between stakeholders. This was forcefully articulated in

Shohamy's (2001a) *The Power of Tests*. Building on ideas put forward by Foucault (1977), Shohamy conceives of tests as tools by which the owners and users of tests enforce their agendas. While it is conceivable that these agendas could be morally defensible, Shohamy warns that intentions behind test design and use can be both overt and covert. Investigating the values imbedded within the test requires examining both the stated aims of test users and unstated (perhaps even unconscious) aims that the test implies. In the words of Frost (2017): "tests offer a low cost and efficient means of manipulating the flow of people into a country, as well as a means of defining and enforcing national language ideologies and norms" (p. 46).

The partially covert nature of the values implicit in language testing presents a challenge for researchers interested in this facet of validity. For academics and test developers, Shohamy's Critical Language Testing (CLT) provides a useful framework for investigation. Broadly defined, CLT "assumes that the act of testing is not neutral. Rather, it is both a product and an agent of cultural, social, political, educational and ideological agendas that shape the lives of individual participant, teachers and learners" (Shohamy, 1998, p. 332).

Calls to "democratize" (Shohamy, 2001b) or "humanize" (Hamid & Hoang, 2018) high-stakes testing are another vantage by which to approach the problem of values in testing. These calls are relevant to the values of the test insofar as they treat test-takers as active participants in all facets of the test. Test-taker values can become relevant points of comparison for those values already observed in the test. Investigating how these values align and how they do not, with consideration of the role of dominant power structures within these values, could provide a useful guiding tool when addressing this facet of test validity.

A growing body of research, especially in the form of interviews and case studies, has attempted to give test-taker perceptions a greater voice in the conversations about the value

implications imbedded in the IELTS test. Some examples include research into test-taker perceptions on the acceptability of world Englishes in IELTS (Hamid, 2014) and the extent to which the native-speaker-as-ideal is reflected in speaking test scores (Frost, 2017). The role of IELTS as a money-making enterprise, and how this may impact scores, has also received some scrutiny from test-takers (Frost, 2017; Hamid et al., 2019; Hoang, 2019; Rumsey et al., 2016). While the body of work has grown in recent years, there remains a need for test validation research to give voice to the values of test-takers and to test-taker perceptions of how these values are reflected (or not reflected) in the IELTS test and in its use.

2.5.4 Testing Justice through Consequential Validity.

The consequences of test use can be examined through various lenses. Washback, “the effect of the test on the teaching and learning leading up to it,” (McNamara, 2000, p. 138) is an element of consequential validity that has received a considerable amount of scholarly attention (see Allen, 2016; Nguyen, 2023). Meanwhile, effects on society more broadly are often referred to as test impact (McNamara & Roever, 2006). Early appraisals of consequential validity have tended to focus on the potential benefits of language testing. For example, Bachman (2004) summarized it in this way:

we use language tests to help us identify second or foreign language learners in schools, to place students into language programs, to screen potential immigrants and to select employees. Language tests thus have the potential for helping us collect useful information that will benefit a wide variety of individuals. (p. 3)

Similarly, Kunnan (2004) codified a view of consequential validity as part of his fairness framework, in what he terms The Principle of Beneficence. Sub-principle 1 states “A test ought

to promote good in society by providing test-score information and social impacts that are beneficial to society” (p. 34).

The issue with such framings is that *what* constitutes a benefit and *who* is included in society requires serious scrutiny. Scholars like Shohamy (2001b) have argued that, without rigorous investigation into the power structures served by testing, benefit will very likely be conferred to those already in power, and that marginalized voices will continue to be ignored when considering the consequences of testing.

Recent research in high-stakes language testing has tried to address this issue by centering test-takers in impact studies. As with value implications, test-taker centred research on consequential validity has mostly been conducted through interviews and case studies. These studies serve to humanize the consequences of high-stakes testing as they are experienced by test-takers themselves. Some of these the consequences include: causing test-takers to change education or career paths (Ahern, 2009; Frost, 2017), delaying skilled professionals from entering the career for which they were trained (Frost, 2017), delaying immigrants’ acquiring permanent residency or citizenship status (Ahern, 2009; Frost, 2017), many of these resulting in the financial burden of retaking the test multiple times. High-stakes testing also has emotional consequences for test-takers which should not be ignored (Cheng & DeLuca, 2011; Frost, 2017; Hoang & Hamid, 2017; Yucel & Iwashita, 2016). This is not to say that test-takers report exclusively negative consequences. Many test-takers report that preparing for the test helped them improve their English (Hoang, 2019) and that the successful completion of the test affected their emotional state and feelings of self-efficacy positively (Hoang, 2019; Yucel & Iwashita, 2016). However, discussions of consequential validity must weigh the positive and negative

consequences to both individual stakeholders and society at large in order to make a coherent and ethically defensible argument for test interpretation and use.

2.6 Recent Studies of Test-taker Perceptions of Fairness and Justice in High-stakes

Language Testing

I have so far established the importance of IELTS as a tool of government, academic, and professional institutions, described McNamara and Ryan's (2011) framing of fairness and justice as a conceptual framework for addressing the validity of IELTS, and argued for a test-taker centred approach to investigating the fairness and justice of IELTS. I now turn to some of recent studies of test-taker perceptions of fairness and justice in high-stakes language testing. While not all of the studies use McNamara and Ryan's language, the studies described share a common approach to validity that is unified and test-taker centred.

2.6.1 Cheng and DeLuca (2011)

I begin with Cheng and DeLuca (2011), who gave 59 English language learner (ELL) university students a prompt to write about their experience with a real-life language testing event. The participants were asked to give "a reflection upon [their] positive, neutral, or negative experience with the language test" and "address an issue of test validity and test use" (p. 108). Using deductive, qualitative coding, Cheng and DeLuca identified eight overarching themes: "(a) test administration and testing conditions, (b) timing, (c) test structure and content, (d) scoring effects, (e) preparation and test-taking strategies, (f) test purpose, (g) psychological factors, and (h) external factors and test consequences" (pp. 109-110).

The most commonly identified theme was psychological factors "typically characterized by stress, nervousness, and anxiety related to test performance and test consequences" (Cheng &

DeLuca, 2011, p.114). Psychological factors were also connected environmental conditions, negative feedback from instructors or raters, and issues related to timing (i.e., the impact of the length of the test, or not having enough time to complete a test or task). The authors go on to argue that the co-occurrence of this theme with preparation and test-taking strategies suggests that greater test preparation may mitigate feelings of anxiety.

Another theme with a high occurrence was test structure and content. The authors delineated this theme into four aspects: “(a) authenticity of tasks, (b) choice in constructed response item, (c) psychometric properties and perceived limitations of fixed-response items, and (d) understanding score criteria and item instructions” (Cheng & DeLuca, 2011, p. 111). Tests were perceived to be more valid when the tasks were more authentic and more choice was provided. The latter point reflects earlier research on test choice conducted by Jennings et al. (1999).

Cheng and DeLuca’s (2011) study asked participants to choose a testing experience to describe, rather than focusing on a specific test as in the present study. However, Cheng and DeLuca’s work is important to IELTS research for methodological reasons. The authors’ deductive qualitative approach has clear advantages for incorporating test-taker perspectives in validation research. The relative open-endedness of the prompt allowed the participants to focus on whichever aspects of test validity they deemed significant. While the themes arising from the study may not allow commentary on the fairness and justice of IELTS specifically, they do point to potential areas for further investigation.

2.6.2 Hamid et al. (2019)

Hamid et al.'s (2019) study on test-taker perspectives of global English tests (focusing mainly on IELTS) is a major contribution to the current discussion of fairness and justice. An earlier study conducted by Hamid (2014) produced a 18,500 word corpus of written comments, provided by 343 mostly ELL participants in Australia. Hamid et al. (2019) then used qualitative coding and content analysis to explore themes of fairness and justice¹. The study is significant both for its comparatively large sample size, but also because it represents test-takers for academic purposes (70%), immigration purposes (13.7%), and a mix of both (5.8%) (Hamid, 2014, p. 266).

The picture of test-taker perspectives on the fairness and justice of IELTS found in Hamid et al. (2019) is a complex. Participants were more likely to view IELTS overall positively. However, the comments on specific aspects of the tests fairness and justice were more critical. For example, on the question of fairness, a high number of participants “asserted that their language ability was not reflected in their test scores, thus problematizing the assumption that good English = good scores” (Hamid et al., 2019, p. 11). Questions of the representativeness of scores to actual English ability were frequently connected to a perception that scores were highly variable. The authors offer the example of participants taking the IELTS test multiple times in a fairly short amount of time and receiving scores of up to a full band’s difference in each sitting.

Comments on the question of justice reported in Hamid et al.'s (2019) data were especially pointed on the topic of IELTS as a money-making venture. Multiple participants described IELTS as being primarily a source of income for test-developers and test-providers. The issue was compounded for some due to the need to take the test multiple times. A major

¹ Hamid et al. (2019) use the terms fairness, justice, and *validity*, which they delineate from fairness. For the sake of consistency, I will consider validity in this sense to be an aspect of fairness.

reason that participants reported needing to take the test multiple times was to achieve the necessary score for immigration or acceptance into a desired academic program. However, some participants also discussed the 2-year shelf life of the test and the frequent necessity to sit the test multiple times for different purposes (i.e., because the participant applied to a post-secondary institution and then later applied for PR in Australia).

Hamid et al. (2019) also discuss the justice of requiring IELTS for L1 English speakers. The Australian system requires L1 English speakers to provide test results when applying to immigrate, but Australian post-secondary institutions do not typically require L1 English-speaking international students to likewise present test scores. On the latter topic, Hamid et al. quote an L1 English-speaking immigrant from Ireland, who described the need to take the IELTS test as “madness, and a waste of money” (p. 14). On the topic of IELTS for academic purposes, Hamid et al. note that several of their participants questioned whether it was just to require L2 speakers to achieve test scores which L1 speakers may not be able to achieve.

Hamid et al. (2019) is an important precursor to the current study as it provides a wide range of perspectives for IELTS used in the Australian context, which is comparable if not identical to the situation in Canada. The larger sample size of Hamid et al.’s study strengthens its ability to support claims about how widespread the opinions it reports are. However, using a corpus of questionnaire comments as a dataset means that there is still a need for the more in-depth data which can be provided by other sources.

2.6.3 Hoang (2019)

Hoang (2019) offers a useful alternative approach to the wider ranging work of Hamid et al. (2019). Hoang (2019) collected questionnaire data from 39 participants who had taken either

IELTS or TOEFL for the purpose of immigration to Australia. The author then conducted semi-structured interviews with six of these participants and coded the resulting transcripts.

As in earlier studies, more of the participants in Hoang's (2019) study agreed that the test they took was an effective measurement of their English abilities than disagreed. Likewise, a majority of respondents expressed the belief that the reliance on IELTS and TOEFL within Australia's skilled migration program was appropriate and that the current cut score for the program was also appropriate. Also noteworthy was that three quarters of participants reported experiencing positive washback from taking the test, but a majority reported finding the cost of the test significant.

As in earlier studies, psychological/emotional impact was identified as an important theme. A majority of questionnaire respondents reported that feeling stressed or nervous while being tested interfered with their performance.

After psychological/emotional impact, questionnaire respondents identified test content/topic as the second most likely cause of interference with test performance and scoring. All participants interviewed for the study responded that they "would have performed better if the topics had been related to their field or study or work" (Hoang, 2019, p. 141). Hoang recounts that many participants framed receiving a familiar topic as a kind of "luck", which she connects to an ongoing discussion of whether topic variability in non-field specific test like IELTS threatens of the validity of these tests (see Jennings et al., 1999).

2.6.4 Liu and Hamid (2024)

The final and most recent study published at time of writing was conducted by Liu and Hamid (2024). As in Hoang (2019), Liu and Hamid made use of coded, semi-structured interviews to

study test-taker perspectives of the validity IELTS. In this case, the focus is on IELTS for higher education rather than immigration. The authors interviewed 17 ELL international students, already studying at Australian post-secondary institutions.

A major contribution made by Liu and Hamid (2024) is in offering a glimpse into how test-takers compare their experience with IELTS to their language needs during subsequent post-secondary studies. One such point of comparison between IELTS and real communication in higher education was an apparent disconnect in the importance of accuracy versus communicative competence. The participants reported feeling that the focus on accuracy in IELTS (particularly spelling and grammar) did not reflect their real-world experiences. This was in part because IELTS limits tools that would usually be available to the participants such as spell checkers. Perhaps even more significantly, many participants described a real-world need for communicative flexibility and adaptability in their interactions with others, with a high emphasis on ELF communication which was not reflected in the IELTS test. Recall that the underrepresentation of ELF communication in academic settings mirrors earlier findings made by Manias and McNamara (2016) when comparing medical professionals' perceptions of their workplace communication needs and the OET.

Another theme emerging from Liu and Hamid (2024) was the importance of topic coverage. The general trend emerging from the interviews was that the participants viewed the topic covered during their IELTS test as a significant factor in how well they performed. This impact could be positive or negative, with younger test-takers being at a potential disadvantage, having less of a range of experience to call upon during the test.

The preceding literature review has sought to establish a few key facts as they pertain to the use of IELTS in Canada. It is deeply integrated into Canadian immigration, higher education, and professional accreditation systems. Because IELTS is a tool used by powerful institutions with meaningful stakes and consequences for stakeholders, its continued interpretation and use warrants ongoing, critical appraisal. Unified validity provided an important shift in that it considers the consequences and values of a test alongside its construct validity and relevance/utility. The decades following the description of the unified validity framework have seen increasing calls for testers and test validation researchers to address the fairness and justice of high-stakes English language testing in its various uses (Deygers, 2019). However, to make test interpretation and use truly democratic, the questions of fairness and justice must centre the voices and experiences of test-takers. This has led related calls for greater engagement with test-taker perspectives both on the research and policy levels (Jin, 2023). While work in this area is growing, so too is the need for ongoing research.

As of the time of writing, the Canadian context is underrepresented in recent test-taker centred research into the validity of IELTS. Given the importance of Canada as a destination country for international studies and immigration and given the importance of IELTS to the process of studying in or immigrating to Canada, this represents a meaningful gap in the research.

Moreover, previous studies have focused on testing in general as in Cheng and DeLuca (2011) or on IELTS as a whole as in Hoang (2019), Hamid et al. (2019), and Liu and Hamid (2024). While there are advantages to a holistic approach, this does mean that issues unique to only one of speaking, reading, writing, and listening may go underexplored. For this reason, studies which focus on one specific part of the test are represent a meaningful gap in the

research. The speaking portion of the IELTS test (along with listening) has the added issue of being identical in both IELTS GT and IELTS Academic. The choice to have a single speaking test for all IELTS test-takers would seem a clear area of concern for questions of fairness and justice.

With these considerations in mind, the research questions I investigated were as follows:

1. Do test-takers view the IELTS speaking test as *fair* (as conceived as both construct validity and relevance / utility)?
2. Do test-takers view the IELTS speaking test as *just* (as conceived as both value implications and social consequences)?

Chapter 3: Methods

In this chapter, I begin by outlining my positionality as a researcher. A description of the research methods follows, in which I describe the setting, participants, data collection instruments, and procedures. Finally, I describe the methods used to analyze the data collected.

3.1 Positionality

The following discussion of my positionality lays out the elements of my identity which I believe are relevant to this study as well, my relationship with IELTS, and the research paradigm underpinning my research.

3.1.1 Considerations of Power and Privilege

As argued in Rose et al. (2020) and elsewhere, the subjectivity of the researcher cannot ever truly be removed from the study they are conducting. This is especially true given the relevance of power imbalances to the present study. As the author of this study, my identities as a white, male, university educated, native (L1) English speaker, who is a settler/non-immigrant Canadian, are all significant to my positionality. The aforementioned identities imply a level of privilege within the Canadian context. IELTS takers, by the nature of needing to take the test, will almost certainly be at least one of: a speaker of English as an additional language, an immigrant or newcomer to Canada, and/or seeking to begin a university education. As such, my relative position results in a dynamic which researchers like Kubota (2016), among others, refer to as *studying down*. As the name suggests, studying down (as opposed to studying across or studying up) implies a research context in which the researcher holds a privileged status compared to participants. This is particularly important to note when the loci of the imbalances (for example,

country of birth) are relevant to the study. Studying down presents both ethical and methodological issues which the researcher needs to address.

Conducting research that does not exploit, coerce, or otherwise harm participants is, of course, paramount. However, ensuring this does not take place, and taking steps to address how one's privilege may be affecting the research being done requires conscious effort. I will now outline two ways which I hope have positively impacted my positionality during the present study: ongoing reflexivity and collaboration with peers.

I have attempted to conduct my research reflexively with the help of wider reading on power imbalances related to language, education, testing, and immigration, as well as seeking out talks and reading groups to expand my own understanding of these topics. Personal journaling was used as a way of cementing these ideas and connecting them to the current study, and also as a way of challenging my own identities as a researcher and instructor with a position of privilege.

The collaborative aspect of the current study was at first simply a product of the fact that this study of fairness and justice of IELTS is part of a larger ongoing project being conducted at the University of Victoria. While often a slower, more deliberative process, it has resulted in set of questionnaires and interview questions which are the product of hours of discussions among myself, my supervisor who is the Principal Investigator (PI) on the multi-year project, and numerous research assistant collaborators of varying backgrounds. While no one can or should claim to have overcome the biases and blind spots that privilege produces, it is my hope that the collaborative, consensus-driven nature of the research implementation process used in this study has helped me to see some of these biases and blind spots where they have affected my own work as a researcher.

Additionally, the PI's main study is being conducted with the reciprocity principle as a significant guiding philosophy. It is my aim to conduct my own research similarly guided by this principle of giving back to the participants and stakeholders.

3.1.2 Relationship to IELTS

My personal experience with IELTS education and research has led me to the conclusion that there are very few teachers, researchers, or students involved with IELTS who do not have an opinion about it. It is important, then, in the interest of transparency, to discuss my own background and relationship with the test. I have worked as an English-language instructor, both in Canada and China, from 2012 up to the time of this study. During this time, I have had numerous students and other friends and acquaintances who were required to take the test, both for immigration purposes and for university admission. Additionally, I have worked as an IELTS speaking examiner on two separate occasions. This final point warrants the greatest reflexive scrutiny. While my IELTS examiner experience has been an asset to conducting this study, it could also introduce an element of bias or even conflict of interest which must be addressed.

My familiarity with the test's procedures and grading scheme was relevant at various points while collaborating on the mock tests, the earlier versions of the questionnaire, and the interview questions for the participants involved in this study. However, at no point was confidential information used or made available. The mock tests were based on unofficial practice materials available online and official, publicly available descriptions of the speaking test format. Likewise, questions pertaining to the grading of the test were written using the wording of the official, publicly available descriptions of the grading schemes, not the confidential grading rubrics used by examiners during the test. Finally, some participants were given written feedback on their performance on the mock test when requested by the participant.

The feedback given was of the general sort a teacher or tutor would provide, not a score or estimated score. Participants were explicitly told they would not receive a score.

3.1.3 Research Paradigm

Reflexivity requires a clear description of the researcher's ontological, epistemological, and axiological assumptions (Pretorius, 2024). The centrality of test-taker perspectives to this study implies a belief that claims about the fairness and justice of the IELTS speaking test can be answered by individuals through subjective experience. I have therefore adopted a primarily constructivist approach while conducting this study. However, IELTS test-takers are both individuals and members of a larger group of stakeholders. Focusing solely on individual experiences without considering broader trends would paint an incomplete picture. For this reason, the philosophical underpinnings of the present study are best understood as pragmatist, bringing both constructivist and positivist elements to bear as required for addressing a real-world issue in language testing and policy. In practice, this means that I will require both qualitative and quantitative data to address my research questions in their full complexity.

3.2 Participants

Participants were recruited online via social media posts. The present study focuses on seven participants from a larger ongoing study. These participants were selected as they were the first seven participants to complete the interview portion of the study.

As can be seen in Table 8, the participants were four females and three males, all between the ages of 20 and 39. Each of the seven participants identified a different country of origin, with Portuguese, Hindi, Arabic, Urdu, Indonesian, and Cantonese listed as an L1 of one participant each, and English listed as an L1 by three of the participants. Note that the participants were

asked to self report their first language. Those who listed more than one first language were not asked to identify which of these they felt more comfortable using.

Table 8

Participants' Demographic Information

P#	Age	Gender	Country of Origin	First Language(s)	Other languages spoken
P1	30-41	Female	Brazil	Portugues	English, Spanish, and French
P2	20-24	Male	India	English, Hindi	Haryanvi
P3	20-24	Female	Malaysia	English	Some Mandarin
P4	30-34	Male	Tunisia	Arabic	French
P5	30-34	Female	Pakistan	Urdu	English, Punjabi
P6	35-39	Male	Indonesia	Indonesian	English, Portuguese
P7	20-24	Female	Honk Kong	Cantonese, English	Mandarin

Note. P# = Participant Number

The participants' history with English as well as some basic background on their experience living in Canada (and other Majority English-speaking countries) is presented in Table 9. The participants reported having begun their English studies between the ages of one and 11, with a total number of years spent learning English (formally or informally) ranging from eight to 19 years. The amount of time spent living in Canada ranged from five months to seven years. Additionally, P2 reported having lived in the USA previously, and P6 reported having lived in both England and New Zealand before coming to Canada.

Table 9*Participants' Background with English-language Learning and Living in English-speaking Countries*

P#	Age started learning English	Number of years learning English	Other English-speaking countries resided in	Length of time in Canada
P1	11	19	None	22 months
P2	3	21	USA	5 years
P3	1	22	N/A	4 years
P4	10	8	0	5 months
P5	8	18	None	5 years
P6	8	15	England, New Zealand	6 years
P7	2	11	None [visited U.S and U.K.]	about 7 years

Note. P# = Participant Number

The participants' education and work backgrounds are listed Table 10. All seven participants reported having completed at least one university degree. P1, P4, and P7 reported that they were currently students as well, pursuing further education.

Table 10*Participants' Education Background and Current Studies*

P#	Highest level of education completed?	Are you currently a university student?	Are you studying now?
P1	University: History	Yes – Discipline: History	Yes, I am currently studying a subject in my primary field of work/interest.
P2	University: Comp Sci; Minor in Psych	No	No, I am not studying.
P3	University: Psych	No	No, I am not studying.
P4	Post-graduate education: Law	Yes – Discipline: HR	Yes, I am studying another subject in order to be allowed to work in Canada.
P5	University: Science	No	No, I am not studying.
P6	University: Kinesiology	No	No, I am not studying.
P7	University: Comms; Gender, Sexuality and Women's Studies minor	No	Yes, I am currently studying a subject in my primary field of work/interest.

Note. Comp Sci = Computer Science; P# = Participant Number; Psych = Psychology; HR = Human Resources; Comms = Communications.

Table 11 shows the participants' work history and current employment status. Five of the seven participants reported that they were working at least part time in their primary field of work (i.e.,

the field that they had trained or studied for). P4 reported that he was not currently working and was a student at the time of the interview, and P5 was a stay-at-home primary caregiver.

Table 11

Participants' Work History and Current Employment Status

P#	Field(s) of work before moving to Canada	Field of work after moving to Canada	Current situation
P1	Education	Education	Working part-time in my primary field of work.
P2	Engineering	Engineering	Working full-time in my primary field of work.
P3	Not applicable / I don't know	Healthcare	Working part-time in my primary field of work.
P4	Law	N/A	I am not working.
P5	Education	Stay-at-home parent, caregiver, etc.	Stay-at-home parent, caregiver, etc.
P6	Education	Education	Working one full-time job outside my primary field of work, and working part-time in my primary field of work.
P7	Not applicable / I don't know	Government, Healthcare	Working full-time in my primary field of work.

The participants' previous experience with IELTS and other high-stakes standardized English tests is presented in Table 12. Five of the seven participants had taken an IELTS test before participating in the study, with P2, P3, and P7 reporting having taken the academic version of the test, and P5 and P6 reporting having taken the general version. P1 and P4 had not taken an official IELTS test before participating, but P1 reported to be considering IELTS in the future for a permanent residency application. P4 discussed having considered taking the IELTS test before deciding to take the Duolingo English Test for university admission.

Table 12*Participants' History with IELTS and Other Standardized English Tests*

P#	Tests taken	IELTS test(s) taken	Number of IELTS taken	Date of last IELTS	Highest IELTS Speaking score	Reason for taking IELTS
P1	TOEFL	N/A	0	N/A	N/A	N/A
P2	IELTS, TOEFL	Academic	1	Dec-18	7.5	For College / University entrance
P3	IELTS, Cambridge Exams	Academic	1	2018	I don't remember	For College / University entrance, For a visa application
P4	DET	N/A	0	N/A	N/A	N/A
P5	IELTS	General	1	2014	I don't remember	For a PR application
P6	IELTS	General	2	22-Jul-23	8 or above	For a PR application
P7	IELTS, TOEFL, CELPIP	Academic	1	2014	8 or above	For College / University entrance

Note. DET = Duolingo English Test; IELTS = International English Language Testing System; N/A = Not applicable; P# = Participant Number; TOEFL = Test of English as a Foreign Language.

Finally, Table 13 shows the participants' future plans regarding the IELTS and reason for taking the IELTS in the future. P1 and P2 were the only participants who indicated that they were planning on taking an IELTS exam in the future. Both P1 and P2 reported that this reason for this was a PR application.

Table 13*Participants' Future Plans for Taking IELTS*

P#	Do you plan on taking the IELTS in the future?	Why do you plan on taking the IELTS?
P1	Yes	For a PR application
P2	Yes	For a PR application
P3	No	N/A
P4	No	N/A
P5	No	N/A
P6	No	N/A
P7	No	N/A

Note. N/A = Not applicable; P# = Participant Number.

3.3 Data Collection Instruments

The three primary data collection instruments were mock IELTS speaking tests, questionnaires, and interviews.

(a) Mock IELTS Speaking Tests. To account for the variability of participant test-taking experiences, the administration of questionnaire was preceded by a mock IELTS speaking exam. These were created by a research team, including myself and my supervisor. Eight mock exams were selected using publicly available test preparation materials, with some editing to ensure consistency in terms of number of questions and wording. To ensure a variety of topics and multiple tests to randomly draw from: two mock tests were broadly focused on academic study, two on daily life in Canada focus, two relating to work or professions, and two which were focused on topics not clearly relevant to any of these.

(b) Questionnaires. After participants completed the mock speaking test, they were given access to an online questionnaire via SurveyMonkey, an institutionally supported platform. Participants could answer directly after the mock exam or later as per their preference, and a research assistant was available to assist the participants as needed.

The questionnaires were developed collaboratively with the research team. The questionnaires used a combination of Likert, multiple-choice, checklist, and open-ended questions. The quantitative data taken from the questionnaires to collect a base data set. The questionnaires' primary use was to help guide and personalize the questions asked in the in-depth follow-up interviews.

Three initial versions of the questionnaires were created for participants who have taken or intend to take IELTS for various purposes: university admission, immigration or citizenship,

professional credential recognition. The professional credential recognition questionnaire was not ultimately used as no participants reported taking IELTS for this reason. Subsequent combined versions were created by another research team led by the PI to allow participants for whom more than one category is applicable to answer from multiple perspectives within the same questionnaire. All versions of the questionnaire had 15 items related to general information such as participants' L1, education, work history and so on. All versions of the questionnaire also had nine items related to participants' previous experience with IELTS and other English language testing, reasons for taking the test, and future plans to take a high-stakes English test. The final section of the test asked participants to offer opinions about the IELTS and relevant language testing policies. The number of items included was based on the participant's reported reason for taking the test: 16 items for the university admission questionnaire, 22 items for the immigration questionnaire, and 36 items for a combined university admission and immigration questionnaire. Questionnaires were refined and piloted by the PI and a research assistant prior to implementation. To view questionnaires, readers are directed to contact the PI.

(c) Interviews. The primary source of data for this study are the transcripts of semi-structured interviews conducted shortly after the completion of the mock test and questionnaire. These interviews, conducted either in person or online according to participant preference, lasted approximately one hour. While a "script" of interview questions was used to guide the interview, each set of interview questions was developed or modified in advance in response to each participants' questionnaire responses. The questions pertaining to the seven participants included in this study were produced collaboratively by a research assistant, the PI, and myself. As an example of a question tailored to address a specific questionnaire response, we asked P4 to expand on a response he gave in the questionnaire:

Interviewer: ...why do you think that the score is only SOMEWHAT accurate of an indicator of your English-speaking skills?

P4: Because there is some luck on it. If you don't get the right question or you don't like you don't feel comfortable, sometimes even for me, if... , like with... the same level with my English, that won't change, if <you ask me one question> I can feel comfortable, present like thousands of arguments or talk about something that like I like or I don't know (???). <And you can ask me questions that I don't find any interesting things to say so it will reflect on my on my marks later.

As another example, we asked P1 to elaborate on why she chose to take a TOEFL test rather than the IELTS test:

Interviewer: ...why did you choose the TOEFL test over another standardized test, like IELTS?

P1: Okay. Uh, why... I think it was cheaper back in time than IELTS, for example. Also, I think at that time I thought that IELTS will require me, like, British English, instead of only like English, so I was, like, Okay, I will choose TOEFL because, yeah, I want to go to Canada and I think it'll be like the same English. Yeah, kind of the same English-North American English.

In total, the mean number of questions asked to each participant was 46.

3.4 Data Collection Procedures

The data collections procedures are presented in three phases: recruitment, mock IELTS speaking test administration and questionnaire completion, and finally the follow-up interview.

3.4.1 Recruitment

With ethics approval having been granted, recruitment began in the summer of 2023.

Recruitment was conducted through online advertisements and social media posts on social networks primarily geared towards newcomers to Canada, international workers, and international student communities. The study was advertised as an opportunity to contribute to research related to standardized testing as well as an opportunity to practice taking a mock IELTS speaking exam.

Respondents were asked to provide their preference for participation via an online or in-person mock tests as well as their reason(s) for taking the IELTS test (i.e., university admission, immigration, and/or professional credential recognition). This information would determine which version of the questionnaire they would receive and which version of the mock speaking test would be given to them. Participants were also given the option to receive feedback on their mock test performance (upon request) which was given in the form of written feedback.

3.4.2 Mock IELTS Speaking Test Administration and Questionnaire Completion

Shortly before the scheduled mock exam session, a research assistant would decide via coin toss whether the participant would be given a “relevant” (i.e., relevant to their stated reason for taking the test) or “general” version of the speaking test. As two general mock tests were compiled as well as two relevant mock tests for each of the three reasons for taking the test, which version the

participant was given (e.g., General Test 1 or General Test 2) was staggered from participant to participant.

After the research assistant determined which mock test would be given, the participant was invited into the mock test meeting space, either online via a zoom chat or in the Applied Linguistics Lab at the university campus. In both cases, the mock test was attended by the participant, a research assistant, and the mock test “examiner” (in this case, myself). Before proceeding with the mock test, the research assistant reminded the participant that their participation was voluntary and that they could end the mock test at any time before requesting permission to record the mock test.

As with a real exam, the mock exams were divided into three parts: part one, short questions on familiar topics, part two, and part three, a more open conversation. In part one, the examiner begins by reading a short introduction of the test and begins asking the participant part one questions in order. The examiner must adhere to the questions as written but may ask “why?” or “why not?” as a follow-up question.

In part two, the participant is given the part two prompt and three additional suggestions of what they should talk about when addressing the prompt. When the mock exam was conducted in person, the prompt was written on a piece of paper. When the mock exam was conducted via Zoom, the prompt and suggestions were sent in the text box. The participant was given one minute to write and then two minutes to speak on the topic. The examiner remained silent for the full two minutes unless the participant ran out of things to say before the full two minutes had elapsed. In this case, the examiner could prompt the participant further by saying “tell me more about that”, encouraging the participant to elaborate on what they just said or “tell

me more about...” indicating one of the given suggestions. After two minutes, the examiner could ask an optional rounding off question and proceed to part three.

In part three, the examiner had two sets of three thematically related questions to ask the participant. Like part one, these questions must be asked in order. However, unlike part one, the examiner was encouraged to rephrase the questions and ask impromptu follow-up questions as timing allowed.

In all three parts, the examiner must maintain the test’s procedure and order of questions. Ideally all questions would be asked, but the time limit of five minutes for part one, four minutes for part two (including one minute preparation and two minutes to speak), and five minutes for part three must be adhered to. The examiner firmly but politely interrupted the participant in order to maintain the test timing procedures.

When the mock test was finished, the recording was turned off. The participant was then instructed on how to access the questionnaire, which they were encouraged to complete at a time of their choosing. The research assistant remained available to answer questions about questionnaire items if the participant needed clarification. The questionnaire took approximately thirty minutes to complete.

3.4.3 Follow-up Interview

As part of the questionnaire, participants were asked to indicate their willingness to participate in a follow-up interview. Interviews were scheduled with participants who expressed an interest after completing the questionnaire. These interviews were semi-structured and were conducted within two weeks of completion of the questionnaire. During the interview, follow-up questions were given to the research assistant interviewer via private message in order to reduce disruption

of the flow of the conversation. The interview was conducted in person on two occasions. In these cases, I opted to interject directly.

In all cases, the research assistant interviewer began each interview by obtaining ongoing, informed consent, informing the participant that their participation was voluntary and confidential and that they could end the interview at any time and without any explanation.

3.5 Data Analysis

Two sources of data were collected for the present study: the questionnaires and interview transcripts. The methods of analysis for these data sources are described in the following sections.

3.5.1 Analysis of Questionnaire Data

Data gathered from the questionnaire was analyzed by assigning a numerical value to each Likert scale question. “Strongly disagree” was assigned a value of one. “Disagree” was assigned a value of two. “Somewhat disagree” was assigned a value of three. “Somewhat agree” was assigned a value of four. “Agree” was assigned a value of five. “Strongly agree” was assigned a value of six. Responses of “not sure” or “neither agree nor disagree” were not assigned a value. The mean of response values (excluding responses of “not sure” or “neither agree nor disagree”) was calculated in Excel to find the agreement rate for each question, as was the standard deviation (*SD*) of these responses. A mean value above 3.5 indicates some level of agreement, with a mean value of 5.0 or greater suggesting strong agreement. A mean value below 3.5 indicates some level of disagreement, with a mean value of 2.0 or lower suggesting strong disagreement. A mean value of 3.5 represents neutrality, neither agree nor disagree.

3.5.2 Analysis of Interview Data

Interviews were transcribed by a research assistant and checked for accuracy. Data analysis was conducted primarily through qualitative, thematic coding of the seven interview transcripts. Coding was conducted electronically using NVivo 14, beginning in December 2023 and ending in August 2024. Codes were created, refined, and organized over multiple cycles of coding, using the procedures described in Saldaña (2013). As themes developed, a codebook was created.

First Cycle (Initial) Coding. In the first cycle of coding, I took an Initial Coding (also referred to as Open Coding) approach. This more open-ended, descriptive approach was chosen as it gave me the flexibility to describe and label data as it appears in the transcript without forcing the data to fit with my language or theoretical assumptions. A deductive coding approach allows for the possibility of unexpected themes and patterns to develop. Moreover, it can help mitigate the biases and expectations that the researcher brings to the process of coding.

Second Cycle (Focused) Coding. I began second cycle coding in the spring of 2024. For this cycle, I used a Focused Coding approach. While the priority of the Initial Coding done in the first cycle was a descriptive openness, the Focused Coding of the second cycle aimed to group codes into meaningful categories and subcategories. This was done through multiple passes of all seven transcripts, allowing for codes to be combined or separated as appropriate.

Codebook Creation. After multiple rounds of second cycle Focused Coding, I was able to use the codes and categories which emerged to create a codebook for this study. During the codebook creation process, I produced finalized code names and descriptive definitions for each code, which are presented in Appendix 1. During the codebook creation process, I selected individual instances for each code to serve as an example. The individual instances were selected

with the aim of providing both a clear example of the code and a balance of each of the seven participants. All code definitions and examples were reviewed by my supervisor. Several rounds of these revisions resulted significant refinement to the final codebook. For example, the choice to differentiate whether a participant was explicitly describing the mock test or a real testing experience was the result of the revision process.

Finally, several codes and categories were removed from the main codebook as they were deemed outside of the focus of the present study. For example, several codes were created for participants describing tests other than IELTS. While this data may be valuable for future studies, the choice was made to only include these codes for analysis when the participant did mention IELTS, typically as a point of comparison. After all rounds of Focused Coding and codebook revision were complete, the finalized codebook contained 133 distinct codes, each coded under one of 13 top-level parent codes. Table 14 shows the top-level parent codes names and abbreviations along with number of child codes.

Table 14

Top-level Parent Codes with Number of Child Codes and Total Instances of Child Codes

Top-level code abbreviation	Top-level code name	# of child codes	# of child code references
AFFECTED PERF	Expresses Belief That IELTS Performance or Score Was Affected by Factors Other Than English Language Ability.	33	82
ALTERNATIVES	Proposes an Alternative Way of Proving English Abilities	16	30
CHANGES 2 IELTS	Proposes Changes to IELTS Itself.	2	3
COMM SKILLS	Describes Communication Skills Required (or Not Required) for an Academic, Daily Life, or Workplace Setting	7	30

Table 14 (continued)

ENG LVL	Discusses the English Abilities of Native English Speakers OR Appraises English Abilities or Changes in English Abilities Over Time	2	22
IMPACT	Describes the Impact of IELTS	15	41
MIN SCORE	Offers Opinion about Minimum IELTS Score Accepted by an Institution (i.e., the Government, Universities, etc.)	3	9
MOTIVE	Gives Motive for Taking IELTS	4	9
MULTI	Discusses Taking IELTS Multiple Times (either an Experience OR an Opinion about the Necessity of Taking the Test Multiple Times)	1	5
OPINION	Expresses an Opinion about IELTS	34	216
PREP	Describes Preparing for IELTS	2	14
REASON	Gives Reasons for Choosing or not Choosing IELTS	4	7
STRAT	Describes Test-Taking Strategies Used (or Didn't Use) for IELTS	12	27

Note. IELTS = International English Language Testing System

Intra-Coder Reliability. Intra-Coder Reliability was established via blind recoding of three of the seven transcripts. That is, one week after the final pass and codebook creation phase, I recoded three randomly selected transcripts using the previously established codes from the codebook. Participant 3 (P3), Participant 5 (P5), and Participant 7 (P7) were selected and recoded. The total agreement was calculated at 85% agreement for P3's transcript, 86% for P5's, and 84% for P7's, resulting in an average of 85% agreement. In the cases where discrepancies caused me to reconsider my initial coding, I made of note of this and updated the original code. As an example, the following was excerpts from P3 was coded as *Give Prepared Presentation* during the reliability check but not in the original: "For academic purposes, I mean they- if they want to, they could give you a topic in advance and make you prepare for a presentation because that gives you a chance, a chance to prepare" (P3). After reviewing this coding decision, a note was made and the code was added to the original transcript to reflect this decision.

This phase served as a final check to address any major discrepancies between the earlier Focused Coding and this later check for reliability.

This chapter laid out my positionality as a researcher, the setting, participants, data collection instruments, procedures, and methods used to analyze the data. In the following chapter, I report the findings of this data analysis.

Chapter 4: Results

In this chapter, I present the findings of the questionnaires and qualitative coding. In, Section 4.1, I present the participant opinions about and experience with IELTS as collected through the post-mock speaking test questionnaire. In Section 4.2, I present an overview of the coding data followed by a detailed description of the four themes which emerged as particularly relevant to the research questions of this study, i.e., to what extent do test-takers view the IELTS speaking test as fair and just. In Section 4.3, I offer a summary of the findings as they relate to the research questions of this study.

4.1 Report of Questionnaire Results

Questionnaire data were collected from the participants in order to help guide the subsequent semi-structured interviews and to complement the qualitative data collected from said interviews. Questions were given to the participants based on their stated reason or reasons for taking a high-stakes English test. Six of the seven participants (all except P5) indicated that they had taken or would take a test for admission to an academic program and so completed the English for Academic Purposes (EAP) question set. P5 and P7 indicated that they had taken or would take a test for immigration purposes and so completed the immigration question set. Because P7 indicated both academic program admission and immigration as reasons for taking a test, she was given a version of the questionnaire with both the EAP and immigration question sets.

4.1.1 Overall agreement rate of questionnaire questions

Table 15 shows the agreement rate and *SD* for the EAP questions set. This is followed by Table 16, showing the agreement rate and *SD* for the immigration question set.

Table 15*Agreement Rate for English for Academic Purposes Question Set*

Questionnaire Question	Total Disagree Value	Total Agree Value	Not Sure / Neither	Mean Value	SD
Based on the simulated IELTS Speaking test I just took, part 1 of the test is relevant to the English speaking skills required by my academic program.	5 (n = 1)	15 (n = 3)	0 (n = 1)	4.0	1.6
Based on the simulated IELTS Speaking test I just took, part 2 of the test is relevant to the English speaking skills required by my academic program.	3 (n = 1)	23 (n = 5)	0 (n = 0)	4.3	0.9
Based on the simulated IELTS Speaking test I just took, part 3 of the test is relevant to the English speaking skills required by my academic program.	1 (n = 1)	27 (n = 5)	0 (n = 0)	4.7	1.7
The IELTS Speaking test (in its totality) is relevant to the English speaking skills required by my academic program.	2 (n = 1)	28 (n = 5)	0 (n = 0)	5.0	1.4
The IELTS Speaking test is the most appropriate test to measure the English speaking skills required by my academic program.	0 (n = 0)	20 (n = 4)	0 (n = 2)	5.0	0.7
The minimum IELTS Speaking test score for admission is an accurate indicator of the English speaking skills required by my academic program.	0 (n = 0)	13 (n = 3)	0 (n = 3)	4.3	0.5
Reaching a required minimum IELTS Speaking test score should be necessary for admission to my academic program.	3 (n = 1)	20 (n = 4)	0 (n = 1)	4.6	1.2
The minimum IELTS Speaking test score required for my academic program should be changed.	6 (n = 2)	8 (n = 2)	0 (n = 2)	3.5	0.5
The minimum IELTS Speaking test score required for my academic program should be lower.	6 (n = 2)	8 (n = 2)	0 (n = 2)	3.25	1.5

Table 15 (continued)

The minimum IELTS Speaking test score required for my academic program should be higher.	4 (<i>n</i> = 2)	9 (<i>n</i> = 4)	0 (<i>n</i> = 0)	4.3	1.7
Factors other than my English speaking skills may affect my IELTS Speaking test score.	0 (<i>n</i> = 0)	27 (<i>n</i> = 5)	0 (<i>n</i> = 1)	5.4	0.9
The need to reach the minimum IELTS Speaking test score for admission to my program delayed the start of my academic studies.	7 (<i>n</i> = 4)	10 (<i>n</i> = 2)	0 (<i>n</i> = 0)	2.8	1.7
The need to reach the minimum IELTS Speaking test score for admission to my program has made me consider not pursuing further academic studies.	3 (<i>n</i> = 3)	11 (<i>n</i> = 2)	0 (<i>n</i> = 1)	2.8	2.2
The need to reach the minimum IELTS Speaking test score for admission to my program has made me decide not to pursue further academic studies.	3 (<i>n</i> = 5)	0 (<i>n</i> = 0)	0 (<i>n</i> = 1)	1.4	0.5
Demonstrating English speaking skills should be required of anyone (including speakers of English as a first language) applying for admission to my program from outside of Canada.	9 (<i>n</i> = 1)	9 (<i>n</i> = 4)	0 (<i>n</i> = 1)	4.6	1.2
I should be able to demonstrate my English speaking skills to pursue academic studies in Canada in a way other than taking the IELTS Speaking test or other standardized speaking tests (e.g., by completing courses, interviews, etc.).	3 (<i>n</i> = 1)	28 (<i>n</i> = 5)	0 (<i>n</i> = 0)	5.2	1.1

Note. IELTS = International English Language Testing System; Not Sure / Neither = total number of participants responding “not sure” or “neither disagree nor agree”; SD = standard deviation; Total Agree Value = total Likert scale value of “somewhat agree”, “agree”, or “strongly agree” responses; Total Disagree Value = total Likert scale value of “somewhat disagree”, “disagree”, or “strongly disagree” responses; Likert scale values: “strongly disagree” = 1; “disagree” = 2; “somewhat disagree” = 3; “somewhat agree” = 4; “agree” = 5; “strongly agree” = 6.

Table 16*Agreement Rate for Immigration Question Set*

Questionnaire Question	Total Disagree	Total Agree	Not Sure / Neither	Mean	SD
Based on the simulated IELTS Speaking test I just took, part 1 of the test is relevant to the English speaking skills required for my daily oral English communication in Canada.	1 (<i>n</i> = 1)	5 (<i>n</i> = 1)	0 (<i>n</i> = 0)	3.0	2.0
Based on the simulated IELTS Speaking test I just took, part 2 of the test is relevant to the English speaking skills required for my daily oral English communication in Canada.	0 (<i>n</i> = 0)	9 (<i>n</i> = 2)	0 (<i>n</i> = 0)	5.5	0.5
Based on the simulated IELTS Speaking test I just took, part 3 of the test is relevant to the English speaking skills required for my daily oral English communication in Canada.	0 (<i>n</i> = 0)	10 (<i>n</i> = 2)	0 (<i>n</i> = 0)	5.0	1.0
The IELTS Speaking test (in its totality) is relevant to the speaking skills required for my daily oral English communication in Canada.	0 (<i>n</i> = 0)	12 (<i>n</i> = 2)	0 (<i>n</i> = 0)	6.0	0
The IELTS Speaking test is the most appropriate test to measure the speaking skills required for my daily oral English communication in Canada.	0 (<i>n</i> = 0)	10 (<i>n</i> = 2)	0 (<i>n</i> = 0)	5.0	0
The minimum IELTS Speaking test score for obtaining citizenship is an accurate indicator of the speaking skills required for my daily oral English communication in Canada.	1 (<i>n</i> = 1)	5 (<i>n</i> = 1)	0 (<i>n</i> = 0)	3.0	2.0
Based on the simulated IELTS Speaking test I just took, part 1 of the test (short questions on familiar topics) is relevant to the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	0 (<i>n</i> = 0)	0 (<i>n</i> = 2)	N/A	N/A
Based on the simulated IELTS Speaking test I just took, part 2 of the test is relevant to the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	4 (<i>n</i> = 1)	0 (<i>n</i> = 1)	4.0	N/A

Table 16 (continued)

Based on the simulated IELTS Speaking test I just took, part 3 of the test is relevant to the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	5 (<i>n</i> = 1)	0 (<i>n</i> = 1)	5.0	N/A
The IELTS Speaking test (in its totality) is relevant to the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	5 (<i>n</i> = 1)	0 (<i>n</i> = 1)	5.0	N/A
The IELTS Speaking test is the most appropriate test to measure the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	4 (<i>n</i> = 1)	1 (<i>n</i> = 1)	4.0	N/A
The minimum IELTS Speaking test score for obtaining citizenship is an accurate indicator of the English speaking skills required for obtaining citizenship.	0 (<i>n</i> = 0)	4 (<i>n</i> = 1)	0 (<i>n</i> = 1)	4.0	N/A
Reaching a required minimum IELTS Speaking test score should be necessary for me to obtain citizenship.	0 (<i>n</i> = 0)	5 (<i>n</i> = 1)	0 (<i>n</i> = 1)	5.0	N/A
The minimum IELTS Speaking test score required for obtaining citizenship should be changed.	0 (<i>n</i> = 0)	5 (<i>n</i> = 1)	0 (<i>n</i> = 1)	5.0	N/A
The minimum IELTS Speaking test score required for obtaining citizenship should be lower.	1 (<i>n</i> = 1)	5 (<i>n</i> = 1)	0 (<i>n</i> = 0)	3.0	2.0
The minimum IELTS Speaking test score required for obtaining citizenship should be higher.	2 (<i>n</i> = 1)	0 (<i>n</i> = 0)	0 (<i>n</i> = 1)	2.0	N/A
Factors other than my English speaking skills may affect my IELTS Speaking test score.	0 (<i>n</i> = 0)	5 (<i>n</i> = 1)	0 (<i>n</i> = 1)	5.0	N/A
The need to reach the minimum IELTS Speaking test score delayed my application for citizenship.	2 (<i>n</i> = 2)	0 (<i>n</i> = 0)	0 (<i>n</i> = 0)	2.0	0
The need to reach the minimum IELTS Speaking test score has made me consider not applying for citizenship.	3 (<i>n</i> = 2)	0 (<i>n</i> = 0)	0 (<i>n</i> = 0)	1.5	0.5
The need to reach the minimum IELTS Speaking test score has made me decide not to apply for citizenship.	3 (<i>n</i> = 2)	0 (<i>n</i> = 0)	0 (<i>n</i> = 0)	1.5	0.5
Demonstrating English speaking skills should be required of anyone (including speakers of English as a first language) applying for citizenship.	0 (<i>n</i> = 0)	6 (<i>n</i> = 1)	0 (<i>n</i> = 1)	6.0	N/A

Table 16 (continued)

I should be able to demonstrate my English speaking skills to obtain citizenship in a way other than taking the IELTS Speaking test or other standardized speaking tests	0 (n = 0)	10 (n = 2)	0 (n = 0)	5.0	0
--	--------------	---------------	--------------	-----	---

Note. IELTS = International English Language Testing System; N/A = not applicable; Not Sure / Neither = total number of participants responding “not sure” or “neither disagree nor agree”; SD = standard deviation; Total Agree Value = total Likert scale value of “somewhat agree”, “agree”, or “strongly agree” responses; Total Disagree Value = total Likert scale value of “somewhat disagree”, “disagree”, or “strongly disagree” responses; Likert scale values: “strongly disagree” = 1; “disagree” = 2; “somewhat disagree” = 3; “somewhat agree” = 4; “agree” = 5; “strongly agree” = 6.

When considered as a whole, the questionnaire results reveal a number of noteworthy trends. These are presented thematically, beginning with areas of high agreement to areas with the lowest agreement.

Factors Impacting Performance. Five of the seven participants responded with agreement to the questionnaire statement that factors other than English ability may affect their score with the other two responding that they did not know. This statement was among the more widely agreed up in the questionnaire, receiving a mean agreement value of 5.4 on the EAP question set and 6.0 on the immigration question set, making this question among the most strongly agreed with in the questionnaire.

Alternatives. Another the question that scored high in agreement was on being able to demonstrate one’s English ability in a way other than taking an IELTS test. Six participants agreed and one disagreed. This question resulted in a mean agreement rate of 5.2 in the EAP question set and 5.0 in the immigration set.

Appropriacy. All six participants who responded to the EAP question set agreed that IELTS was the most appropriate test to measure the speaking requirements of their academic program, with a mean agreement rate of 5.0. Both participants who responded to the immigration question set agreed that IELTS was the most appropriate test to measure the speaking skills required for life in Canada, also with an agreement rate of 5.0.

Relevance to Communication Needs. All questionnaire items pertaining to relevance had a mean agreement rate above 3.5, suggesting general agreement with the item statements. The question of the relevance of IELTS in its totality to academic and daily English use being rated particularly high; the mean agreement for the former was 5.0 and the mean agreement for the latter was 6.0. Considering the relevance/utility of the test for academic purposes as three separate sections, all three parts were generally viewed as relevant, with part one receiving a mean agreement rate of 4.0, part two receiving 4.3, and part three receiving 4.7.

Necessity. On the question of academic admission, four participants agreed that achieving a minimum test score should be necessary versus one disagree, with a mean agreement rate of 4.6. On the question of citizenship, one participant agreed that the achieving a minimum score was necessary and the other responded that they did not know.

Scoring. All six participants who responded to the EAP question set agreed that the cut score for the IELTS speaking test was an accurate indicator of the speaking skills required by their academic programs, with a mean agreement rate of 4.3. When asked whether the score should be changed, two responded that they agreed it should, two disagreed, and two responded that they did not know and were not included when calculating the mean. The result was a mean agreement rate of 3.5, suggesting an overall trend which favouring neither agreement nor disagreement. The responses to lowering the cut score were similarly split with two agreements,

two disagreements, and two responding that they did not know. Again excluding the response of “do not know”, the mean agreement rate for this question was 3.25, suggesting an overall trend which slightly favours disagreement. Interestingly, there was higher agreement for the reverse statement, that the cut score should be higher. For this question, four respondents agreed and two disagreed resulting in a mean agreement rate of 4.3.

For the immigration question set, one respondent agreed that the minimum required IELTS speaking score was an accurate indicator of the speaking skills required to obtain citizenship. Likewise, the question of changing the cut score for citizenship resulted in one agreement and one participant responding that they did not know. Interestingly, the question of raising the cut score resulted in one participant agreeing and one responding that they did not know, but the reverse question of lowering the cut score resulted in one participant agreeing and one disagreeing.

Consequences. The questions with the highest overall disagreement rates related to personal consequences. On the EAP question set, four participants disagreed that they need to reach a minimum speaking test score delayed the start of their academic program and two agreed, resulting in a mean agreement rate of 2.8, suggesting overall disagreement. Three participants disagreed that the need for said minimum made them consider not pursuing academic studies, two agreed, and one responded that they were not sure. This also resulted in a mean agreement rate of 2.8, likewise suggesting overall disagreement. Five participants disagreed that the need for a speaking test score made them decide not to pursue academic studies and one responded that they did not know, resulting in a mean agreement rate of 1.4, the lowest agreement rate of any questionnaire question, suggesting very strong disagreement. Similarly, neither of the participants who took the immigration question set agreed with any of the

equivalent statements about the choice to pursue Canadian citizenship. For all three of these questions, both participants responded with disagreement.

4.1.2 Individual participant responses to questionnaire questions

The previous section showed the overall trends in agreement with the questionnaire items. However, *SD* calculations indicate a noteworthy level of variability in participant responses, warranting a closer inspection at an individual level. To give a clearer picture of participants' individual responses, this section presents responses with reference to each participants' assigned number. Questions are grouped by topic for ease of comparison. Table 17 shows the participants' responses to statements about the relevance of the IELTS speaking test to the speaking skills required in the participants' academic program. Three of six participants agreed that Part 1 of the speaking test was relevant to their program with two participants disagreeing and one neither disagreeing or agreeing. For Part 2 and Part 3, five of the six participants expressed at least some level of agreement that these parts of the test were relevant, with one participant disagreeing. These numbers are mirrored when appraising the relevance of the test as a whole.

Table 17

Perceived Relevance of the IELTS Speaking Test to Speaking Skills Required by Academic Program

P#	Part 1 of the test is relevant	Part 2 of the test is relevant	Part 3 of the test is relevant	The Speaking test (in its totality) is relevant
P1	Agree	Strongly agree	Strongly agree	Strongly agree
P2	Agree	Agree	Strongly agree	Strongly agree
P3	Neither	Somewhat agree	Agree	Agree
P4	Disagree	Somewhat agree	Agree	Agree
P6	Agree	Strongly agree	Agree	Agree
P7	Somewhat disagree	Somewhat disagree	Strongly disagree	Strongly disagree

Note. Neither = Neither disagree nor agree; IELTS = International English Language Testing System; P# = Participant number.

Table 18 shows participant responses to the equivalent questions from the immigration question set. For all three parts of the test and the speaking test as a whole, with one exception, the participants agreed that the IELTS speaking test was relevant to the skills required for their daily oral English communication in Canada.

Table 18

Perceived Relevance of the IELTS Speaking Test to Speaking Skills Required for Daily Oral Communication in Canada

P#	Part 1 of the test is relevant	Part 2 of the test is relevant	Part 3 of the test is relevant	The Speaking test (in its totality) is relevant
P5	Strongly disagree	Agree	Somewhat agree	Strongly agree
P7	Agree	Strongly agree	Strongly agree	Strongly agree

Note. IELTS = International English Language Testing System; P# = Participant number.

The next set of questions related to the appropriacy IELTS as a measurement of the speaking skills required by the participants' academic programs, the accuracy of the minimum speaking score as an indicator of having the required speaking skills, and the necessity of setting a minimum score. Table 19 shows that four of six participants agreed that the IELTS speaking test was the most appropriate test to measure the required speaking skills and three of six participants agreed that achieving the minimum speaking score as an accurate indicator of having the required speaking skills for their program. Four of six participants agree that achieving a required minimum speaking score should be necessary for admission to their program, with one participant disagreeing.

Table 19*Perceived Appropriacy and Necessity of IELTS to Participant Academic Programs*

P#	The IELTS Speaking test is the most appropriate test to measure the English speaking skills required by my academic program.	The minimum IELTS Speaking test score for admission is an accurate indicator of the English speaking skills required by my academic program.	Reaching a required minimum IELTS Speaking test score should be necessary for admission to my academic program.
P1	Agree	Neither	Neither
P2	Strongly agree	Agree	Somewhat agree
P3	Neither	Not sure	Somewhat disagree
P4	Agree	Somewhat agree	Somewhat agree
P6	Neither	Not sure	Strongly agree
P7	Somewhat agree	Somewhat agree	Strongly agree

Note. Neither = Neither disagree nor agree; IELTS = International English Language Testing System; P# = Participant number.

Table 20 shows participants' responses to the related questions of whether the previously discussed minimum score should be changed and if it should be made lower or higher. Three of four participants responded that they agreed it should be changed and one disagreed.

Table 20*Opinions Regarding Current Minimum IELTS Speaking Scores Required for Admission into Participants' Academic Programs*

P#	The min score should be changed	The min score should be lower	The min score should be higher
P1	Somewhat agree	Agree	Disagree
P2	Agree	Somewhat agree	Strongly agree
P3	Not sure	Note sure	Disagree
P4	Somewhat disagree	Somewhat disagree	Somewhat agree
P6	Not sure	Not sure	Somewhat agree
P7	Somewhat agree	Strongly disagree	Strongly agree

Note. IELTS = International English Language Testing System; P# = Participant number.

The participants who answered the immigration question set also addressed questions of the appropriacy of the test and minimum score requirements, in this case for meeting the language requirements when applying for Canadian citizenship. Table 21 shows that P5 was unsure whether IELTS was the most appropriate test to measure the speaking skills required for

obtaining citizenship and neither agreed nor disagreed that the minimum score was an accurate indicator of whether a test-taker had the speaking skills required to do so. P7 responded that she somewhat agreed with both of these statements. Both P5 and P7 agreed that reaching a required minimum IELTS speaking score should be necessary for obtaining citizenship.

Table 21

Perceived Appropriacy and Necessity of IELTS to Obtaining Canadian Citizenship

P#	The IELTS Speaking test is the most appropriate test to measure the English speaking skills required for obtaining citizenship	The minimum IELTS Speaking test score for obtaining citizenship is an accurate indicator of the English speaking skills required for obtaining citizenship.	Reaching a required minimum IELTS Speaking test score should be necessary for obtaining citizenship.
P5	Not sure	Neither	Somewhat agree
P7	Somewhat agree	Somewhat agree	Strongly agree

Note. IELTS = International English Language Testing System; Neither = Neither disagree nor agree; P# = Participant number.

Table 22 shows P5 and P7’s opinions regarding the current minimum score needed to obtain citizenship. Both participants agreed that the score should be changed but had opposite opinions about how it should be changed, with P5 agreeing that the minimum score should be made lower and P7 strongly agreeing that the score should be made higher.

Table 22

Opinions Regarding Current Minimum IELTS Speaking Scores Required for Obtaining Canadian Citizenship

P#	The minimum score should be changed	The minimum score should be lower	The minimum score should be higher
P5	Agree	Agree	Disagree
P7	Strongly agree	Strongly disagree	Strongly agree

Note. IELTS = International English Language Testing System; P# = Participant number.

In addition to the citizenship-related questions shown in Table 23, P5 and P7 also answered two questions related to the appropriacy and score accuracy of the IELTS speaking test as it related to their daily English oral communication. Their responses are shown in Table 7. P5

and P7 both responded that they agreed the IELTS speaking test was the most appropriate test to measure their oral English communication skills. However, while P7 agreed that the minimum score required to obtain citizenship was an accurate indicator of the speaking skills required for her communication needs in Canada, P7 strongly disagreed with this statement.

Table 23

Perceived Appropriacy and Necessity of IELTS to Daily English Oral Communication in Canada

P#	The IELTS Speaking test is the most appropriate test to measure my daily English oral communication in Canada.	The minimum IELTS Speaking test score for obtaining citizenship is an accurate indicator of the English speaking skills required for my daily English oral communication in Canada.
P5	Agree	Strongly disagree
P7	Agree	Agree

Note. IELTS = International English Language Testing System; P# = Participant number.

Table 24 shows responses to a question that was present in all forms of the questionnaire and so was answered by all seven participants. Six of seven participant expressed some level of agreement that factors other than their English speaking skills may affect their IELTS speaking test score, with P6 responding that he wasn't sure.

Table 24

Perceived Impact of Factors Other Than Speaking Skills on IELTS Speaking Test Score

P#	Factors other than my English speaking skills may affect my IELTS Speaking test score.
P1	Strongly agree
P2	Strongly agree
P3	Somewhat agree
P4	Somewhat agree
P5	Somewhat agree
P6	Not sure
P7	Agree

Note. IELTS = International English Language Testing System; P# = Participant number.

The next group of questions related to the impact that needing to achieve a minimum IELTS speaking test score had on the participants. Table 25 shows that four of the six participants who answered the EAP question set disagreed that needing to reach a minimum

score delayed the start of their academic studies. Two participants, P1 and P2, agreed that the minimum score requirements had caused them a delay and both of these participants also agreed that this requirement caused them to consider not pursuing further academic studies. Only P2 said that he strongly agreed that the minimum speaking test score requirement made him decide not to pursue further academic studies.

Table 25

Impact of the Need to Reach the Minimum IELTS Speaking Test Score for Admission to Academic Program

P#	Delayed the start of my academic studies.	Made me consider not pursuing further academic studies.	Made me decide not to pursue further academic studies.
P1	Somewhat agree	Strongly agree	Disagree
P2	Strongly agree	Agree	Strongly Agree
P3	Strongly disagree	Strongly disagree	Strongly disagree
P4	Somewhat disagree	Neither	Disagree
P6	Strongly disagree	Strongly disagree	Strongly disagree
P7	Disagree	Strongly disagree	Strongly disagree

Note. IELTS = International English Language Testing System; Neither = Neither disagree nor agree; P# = Participant number.

Table 26 shows P5 and P7's responses to an equivalent set of questions relating to applying for Canadian citizenship. Both participants disagreed that the need for a minimum IELTS speaking test score caused a delay in their citizenship application or caused them to consider not applying or decide to not apply for citizenship.

Table 26

Impact of the Need to Reach the Minimum IELTS Speaking Test Score for Application for Canadian Citizenship

P#	Delayed my application for citizenship	Made me consider not applying for citizenship	Made me decide not to apply for citizenship
P5	Disagree	Disagree	Disagree
P7	Disagree	Strongly Disagree	Strongly Disagree

Note. IELTS = International English Language Testing System; P# = Participant number.

The final group of questions asked participants to respond to a statement about whether a demonstration of speaking skills should be required regardless of national background and whether alternative methods of demonstrating one's skills should be accepted. Table 27 shows that three of six respondents in the EAP question set agreed that *everyone* should have to demonstrate English speaking skills for acceptance into the respondents' academic program and the other three disagreed. On the question of whether an alternative form of demonstrating one's speaking skills should be accepted for admission into their academic program, five out of six participants responded with some level of agreement. Only P4 disagreed with this statement.

Table 27

Whether Everyone Should Demonstrate English Abilities for Admission to Participant Academic Programs and Whether Alternative Forms of Demonstration Should be Accepted.

P#	Demonstrating English speaking skills should be required of anyone (including speakers of English as a first language) applying for admission to my program from outside of Canada.	I should be able to demonstrate my English speaking skills to pursue academic studies in Canada in a way other than taking the IELTS Speaking test or other standardized speaking tests.
P1	Disagree	Strongly agree
P2	Strongly agree	Agree
P3	Disagree	Strongly agree
P4	Strongly agree	Somewhat disagree
P6	Disagree	Strongly agree
P7	Strongly agree	Agree

Note. IELTS = International English Language Testing System; P# = Participant number.

Finally, Table 28 shows P5's and P7's responses the equivalent questions as pertaining to obtaining Canadian citizenship. P5 disagreed that anyone, including native English speakers should have to demonstrate their speaking skills when applying for citizenship while P7 agree with this statement. Both P5 and P7 agree that they should be able to demonstrate their skills in ways other than a test when applying to become citizens of Canada.

Table 28

Whether Everyone Should Demonstrate English Abilities when Obtaining Canadian Citizenship and Whether Alternative Forms of Demonstration Should be Accepted.

P#	Demonstrating English speaking skills should be required of anyone (including speakers of English as a first language) applying for citizenship.	I should be able to demonstrate my English speaking skills to obtain citizenship in a way other than taking the IELTS Speaking test or other standardized speaking tests.
P5	Disagree	Strongly agree
P7	Strongly agree	Agree

Note. IELTS = International English Language Testing System; P# = Participant number.

4.2 Report of Coding Results

As discussed in earlier chapters, I used McNamara and Ryan’s (2011) framing of the terms fairness and justice, with fairness being understood as construct validity and relevance, and justice being understood as consequential validity and value implications, after Messick’s (1989) facets of unified test validity. However, what becomes apparent from attempting to present the data is that much of what the participants described resists being neatly categorized into strictly Messickian quadrants. With this in mind, I have not attempted a granular interpretation of the data in this chapter. While I have had to make decisions with regards to which coded data to present and how, it has been my aim to present to the qualitative data collected from the coding process as openly and objectively as possible. Themes are presented with illustrative examples as they appear in the transcripts. I have occasionally used ellipses for readability and added bracketed explanations, but these were kept to a minimum.

Focused Coding produced 13 top-level codes and 133 child codes, resulting in 495 child code references and 1254 total references across all seven transcripts. Table 29 shows the number of distinct codes and the number of individual references identified in each the seven interview transcripts created for this study.

Table 29*Total Codes and References Labelled in During Focused Coding*

Participant Number	Total Codes	Total Child Code References
P1	54	55
P2	72	79
P3	68	83
P4	39	59
P5	47	58
P6	75	90
P7	51	71

From these codes, four themes were identified as having particular relevance to the present study's questions regarding test-taker perceptions of fairness and justice.

1. Describing Factors that Impacted Performance (33 codes, 86 references)
2. Appraising the Relevance of the Test to Academic and Everyday Contexts (2 codes, 62 references)
3. Describing the Costs and Benefits of the Test (15 codes, 21 references)
4. Considering whether there should be a test (11 codes, 50 references), what kind of test (16 codes, 30 references), and who should take it (1 code, 9 references).

These four themes are explored in detail in the subsequent section of this chapter along with subthemes identified during analysis.

4.2.1 Theme 1: Describing Factors that Impacted Performance

Of the themes most relevant to the question of test fairness, none emerged as strongly as construct external influences on test performance. All seven participants were able to identify as least one. The three most prevalent factors described by participants were affective filters (especially nervousness), test content (i.e., the topics discussed during the speaking test), and test timing (i.e., the length of speaking turns, speaking tasks, and the IELTS test as a whole. While

they are described separately here, it is apparent that the source of negative or positive reactions is frequently described as being related to the timing or content of the test.

Sub-theme 1.1: Affective Filters. The single most frequently arising code pertained to a cluster of affective filters which I have labelled as *Feeling Anxious, Nervous, Stressed, Intimidated or Under Pressure*. Six of the seven participants described these feelings having an impact on either their performance on a real IELTS test or on the mock test they had just completed. Unsurprisingly, when participants described feeling anxious or nervous, these feelings were described as having a negative impact on their performance. However, *how* the participants were impacted by these feelings and *why* they arose revealed considerable variety.

Nervousness caused by testing anxiety. Only P3 explicitly framed her experience of nervousness as being the result general test anxiety. However, one of the ways that her description stands out is in how she connected these feelings to the representativeness of the test itself:

I don't think it was an accurate representation, cause for me personally... I tend to have anxiety when it comes to testing situations, especially speaking to people... I have lots of ideas - my brain, but my mouth just cannot say it and after the exam, I'm always like, why didn't I say that? Oh, why did I say that? So, I feel like there's a lot more in here that I can't say on the spot. I just make things up on the spot (P3)

P3 went on to generalize from her personal experience of test anxiety to question the representativeness of IELTS speaking tests in general:

I think that's more so for - I was thinking students who may have performance anxiety like myself, and like I say, it wasn't it's not a representative.... it's not representative of

what they can do in person, so it's like if someone just blanks out when they're in the exam and end up getting like a really bad score and then deters them or just stops them from entering university. I don't think that's fair. Because anything can happen on the day itself, you get one chance to prove yourself. And that's it. (P3)

P3 connected her feelings of performance anxiety to the experience of “blanking out” which in turn undermines the representativeness of the test in her view. Moreover, the above excerpt shows P3 connecting her feelings of nervousness to another key component of the test: the heightened stakes.

Nervousness caused by high stakes. In comparing the experience of taking a mock exam to a real one, several participants described the mock/real distinction as meaningful in ways that impacted their performance. For example, P2 found that part one of the mock speaking test helped him “to go with much confidence into the bigger conversation that we have in the next part” (P2) and specifically mentioned that this allowed them to become “familiar with... the examiner and be... friendly” (P2). However, P2 went on to draw a distinction between this experience in the mock test and the experience of doing part one in a real test, which he took in 2018. P2 noted that he did not feel the same level of confidence when taking a real IELTS exam because “the pressure was... kind of... different when you are in a test and when you are not” (P2). While P2 would go on to describe the mock IELTS exam and a real IELTS exam as “fair in their own different ways” (P2), he does imply that the difference between the two was reflected in his performance. P2 noted that “in a stimulated [sic] test, I’m in more in like a relaxed and kind of like - I’m free to like speak also like not in like a pressure while in like a real test” (P2).

P1 described a similar experience when comparing a mock test to a real one, implying that the affective difference between the two would impact her strategy use during the test. When

asked if she thought the experience of taking the mock test was different from a real testing experience, she replied:

I'll be very scary, I think, [laughs] if I were in a real test. I'll be very, like, "Oh my God" I need to say then, the format that they want, using the vocabulary that way from me. Yes, I was more relaxed [in the mock test]. Which is sad, I think. (P1)

This comment from P1 is interesting for a few reasons. She suggested that in a real exam she would be more mindful of the exam's criteria, but also that the fact that she *didn't* feel she had to do this in the mock exam meant that she was more relaxed. While this comment could be interpreted as an expression of general test anxiety, P1's description of nervousness while taking a real TOEFL exam offers some insight into *why* she would be more nervous in a real test. Namely, one of the stakes when taking a standardized English test is the cost associated with taking and potentially retaking the test:

So, and for example, for the second time I did the [TOEFL] test, I went with this to the exam because I was so nervous. So, I think, like, it's stress, something- it's a lot of things that can, you know, because if you fail, you need to pay more money to get a test. (P1)

While this comment was not related to IELTS, it is an example of how knowledge of the financial stakes of a high-stakes English test can influence test-takers during the act of testing.

Nervousness caused by social anxiety/public speaking anxiety. An interesting detail of P3's earlier description of her experience of test-taking anxiety was her comment that this anxiety was especially pronounced when speaking to people. While this case was framed as part of her test-taking anxiety, it mirrors a comment made by P7:

I think I have the hardest time with part one in the mock speaking test. Maybe because it was like at the start of the test, and I was still kind of getting used to things, and I get quite nervous when I'm being asked questions, like, in a public setting. So yeah, I kind of... hesitated a lot and, uh, maybe I pronounced things wrong? (P7)

P7's response here is interesting in part because she is addressing the mock test rather than describing a real testing experience, suggesting the stakes were at least different from the earlier participant descriptions of feeling nervous during a real test. Another noteworthy aspect of this comment is how P7 described the impact of feeling nervousness on her performance: hesitation and "wrong" pronunciation. These are significant as pronunciation and fluency are two of the four aspects of speaking by which a test-takers are graded.

Nervousness caused by the test-taking environment. P2's earlier positive comment about getting a chance to feel comfortable and friendly with the examiner relates to another theme that participants connected to nervousness: the test-taking environment, that is, the emotional experience of being in a physical and social space with the examiner. P3 described her real test-taking experience quite vividly.

the way they set it up, it's like they just have you in a room to wait and you have to walk down this long hallway and they make you just sit outside the room and wait for a turn. And it just kind of feels like, I don't know like maybe waiting to enter like a principal's office or something. It just seems like they just made it so intimidating, might even go in, like the room is dark and it's just one examiner there. It's just not a very comfortable experience. (P3)

Nervousness caused by speaking test topic. The topic of the test was also identified as a cause of nervousness. P5 identified uncertainty about the "kind of questions she [the examiner]

will ask from me” (P5) when describing why she felt nervous when taking a real IELTS speaking test. P6 also mentioned the test topic as a source of nervousness:

I think for a lot of people who are still yet to be comfortable with English, that's what makes them feel nervous, because maybe they have studied for like a bunch of different topics, and they were all very random. And then, all of a sudden, you know, like a topic that they have to kind of like talk about is something that's they're still not unfamiliar with. (P6)

In terms of the fairness of the IELTS speaking test, P6 was generally critical of the variability of test topics, which he described as random. However, his comment does also suggest that this randomness factor is at least partially influenced by the test-takers overall comfort with speaking in English.

Sub-theme 1.2: Impact of Test Content on Performance. After affective considerations, the second most discussed potential influence on test-taker performance, according to the participants, was test content (i.e., the topics discussed in parts one, two, and three of the speaking test). Five of seven participants described the topic they received from the examiner having an influence on their performance. The participants described this influence as being either positive or negative depending on the topic assigned. In particular, two elements of the topic were described as impacting performance negatively or positively: familiarity with the topic given by the examiner generally, and relevance of the topic to the participants' field of study.

Impact of Familiarity of the Topic on Performance. Familiarity with the topic given by the examiner was broadly associated with better perceived performance by participants. The

issue addressed by several participants was the element of luck or randomness this introduced into the test. For example, P4 connected familiarity with the topic given to feelings of comfort, ability to demonstrate English abilities, and ultimately the mark received on the test:

there is some luck on it. If you don't get the right question or you don't like you don't feel comfortable, sometimes even for me,... if, like with... the same level with my English, that won't change, if <you ask me one question> I can feel comfortable, present like thousands of arguments or talk about something that like I like or I don't know (??). <And you can ask me questions that I don't find any interesting things to say so it will reflect on my on my marks later. (P4)

P4's use of the word "luck" is somewhat reflected by P6 who described perceived "randomness" of the testing topics as a factor in performance.

I feel sort of like it's not really the most reliable test ever, it's because.... the topic can be so random and sometimes it's something that you don't really like, what's that? I've never even heard of that term before. And you just end up like making things up just because, you know, like you don't know anything about. (P6)

As a positive example, P1 connected her familiarity with the topics she was assigned in part one with being a good demonstration of her pronunciation and vocabulary skills.

As I was talking in familiar - about familiar, they could assess maybe my real knowledge because were, like, familiar, like, terms. So, there are terms that I use often. So yeah, maybe they have, like, some real idea if I'm speaking well, or no? (P1)

Interestingly, P1 was also able to connect her randomly assigned topic in parts two and three of the mock speaking test to her area of academic research. This was in spite of not being directly asked about her field of study.

Impact of Perceived Relevance of the Topic on Performance. In describing the impact that the test content had on their performance, P3 and P5 spoke positively about their experiences with the mock test because of the topics they were given. These participants both compared their mock test experience to their real testing experiences. P3 was randomly given one of the versions of the test that focused on academic topics:

I found the speaking topics for the stimulated [sic] test easier to answer because it's more relevant I think, because for the topics they asked me [during the simulated test were] about university and I just graduated last year, so I can remember a lot of my university experience. I can bounce off on that. (P3)

P3 contrasted this experience with her memory of taking a real IELTS speaking test in 2018:

I think asked me to talk about.... tourist spots back home and I was like, I barely visited any of the tourist spots at home and I can't tell you anything, so anything that asked about family too, but there's only so much you can talk about family, right? There's really not that much to talk about. (P3)

P5 drew a similar comparison when reflecting on her experience taking a real IELTS in 2014:

The one which I have give now [i.e., the mock test], it was easy than that one [i.e., the real IELTS test I took in 2014] that was kind of difficult and it was irrelevant they were, you're asking irrelevant questions. I mean, they were creating some kind of scenarios and just asking to what will you do then that's that was. (P5)

P6 expressed a similar sentiment about the importance of the topic on performance but expanded this idea to considerations of context:

when you're attending, you know, ... like a class, a session, you know exactly what's going to happen. At least you will have sort of like, okay, this is an anatomy class, for instance. And I know, sort of like, what's gonna happen here. The teacher or instructor is going to talk about the human body, movement, whatever it is, right? So your brain is well prepared. Whereas with the IELTS Speaking can be pretty random. (P6)

While several participants discussed testing environment as impacting performance, P6 uniquely discusses the role of context in performance. In the IELTS speaking test, the general topics of the questions to be asked in parts one and three are “introduced” by the examiner, using phrases like “let’s talk about...” before asking the asking a set of questions. This offers some element of schema activation; however, this is quite different from the immersive context described by P6. P6 went on to question the generalizability of IELTS from the perspective of test content:

I think for a lot of people who are still yet to be comfortable with English, that's what makes them feel nervous, because maybe they have studied for like a bunch of different topics, and they were all very random. And then, all of a sudden, you know, like a topic that they have to kind of like talk about is something that's they're still not unfamiliar with. So, yeah, that can be like a, you know, quite a quite a daunting experience for a lot of people I feel like. So that's why it doesn't really translate well to like real life, you know, universal situation, if you will. (P6)

It is interesting to note the participant’s framing of this comment as being about “people who are still yet to be comfortable with English.” In a different context, this could be read as supportive

of the fairness of IELTS from a test content perspective: that is, one's ability to account for the randomness of the topic will be reflective of the test-taker's overall English ability. However, the ensuing comment that "it doesn't really translate well to like real life, you know, universal situation" along with the participant's earlier comments suggest a mostly negative appraisal.

Sub-theme 1.3: Impact of Timing on Performance. The third most prevalent theme relating to factors that may affect one's score was test format. Of the format-related topics addressed by the participants, timing, that is, the length of the test or individual tasks, and/or having or not enough time to complete a test or task, emerged as the most widely discussed.

Impact of Timing Constraints. Four participants, P1, P2, P6, and P7, described timing constraints with regards to performance. P1 and P7 described the experience of being timed more negatively. Describing her experience with the mock test, P7 said:

I also would point out that because the... the parts were timed, and they were quite short, that there wasn't enough time for me to think and also speak at the same time. So, there was a bit of pressure there... (P7)

P7's response suggests both an internal strategy (thinking and speaking) and a resulting affective factor (pressure). She reiterated this point later in the interview as well when describing the extent to which she felt able to demonstrate her fluency as an English speaker, saying "I would say [I felt] hindered [in my ability] because of the time constraint and because I know I was being put under, like, a timer, so that put pressure" (P7). However, P7 did note that of the three parts of the test, the time to prepare a response in part two positively impacted her performance:

Yeah, I had time to prepare for it and think about it, which is why I was able to, uh, present- like express it in a more, like, less convoluted way than I would in parts 1 and 3 when it was just, uh, putting me like on the spot. (P7)

While P7 described an improvement in performance when she felt less of a time constraint, the impact of time constraints was not exclusively framed negatively. P2 described the time constraints in part two of the mock test as encouraging him to convey his ideas more concisely:

when I was given like a time limit to convey like my amount of thoughts I want to say like, concise and like a specific, a specific manner and be like detailed and... not to create some long sentences, and carry forward the conversation and convey my point in less amount of time. So, yeah, it definitely helped me. (P2)

So far, the comments related to timing have focused on the shortness of the speaking test.

However, a contrasting perspective was offered by P2, this time when describing his experience with a real speaking test when, by his own appraisal, his English level was less advanced than at the time of the interview. Reflecting on that time, P2 described finding the length of the speaking challenging for the opposite reason (that is, it felt very long). It is interesting to note that the participant described the mock test as shorter. If P2's real exam was implemented adhering to the standardized timing set by IELTS, it is unlikely that the mock exam was meaningfully shorter in terms of length. However, that P2's experience of the mock exam and memory of the real exam were different is interesting in its own right. He also described the difference between these two experiences as partially being a product of shyness:

I was like shy back then like 3-4 years ago. It... is difficult for me even though like I wanted to speak and express myself, it is difficult for me to speak for so long. So, yeah,

in a way like it, it was different [taking the mock exam], but it still accessed like the same kind of like skills in a shorter amount of time (P2)

The final comment about accessing the “same... skills in a shorter amount of time” is noteworthy. It suggests a more positive appraisal to demonstrate one’s English ability within the timing allowed by IELTS.

P6 also spoke positively about part two in terms of length. In his case, he found the part two to be the best assessment of his skills in part because of the longer period of uninterrupted speaking time offered by the longer turn.

I think the second part, you know, where I had to speak for I mean two minutes was not really a long time, but I could flex a little bit. That would be something that I would say the most useful in terms of how one could actually, you know, like, alright, this is what I've trained for. This is what I've studied for. So this is time for me to shine right? That will be the best part of all three parts. (P6)

P3 expressed a similar sentiment but framed her opinion explicitly in terms of her ability to demonstrate fluency during part two.

I guess that also demonstrated my fluency, especially in Part 2 when I was able to talk for more than like - for one to two minutes. I did kind of stutter at some point because I ran out of things to talk about. But yeah, but I think Part 2 was probably the best method for me to, um, for them to assess my fluency. (P3)

Being Interrupted by the Examiner. Related to the issue of timing and time constraints in the speaking test is the experience of being interrupted by the examiner. In an IELTS speaking test, the examiner may interrupt a speaker in order to keep each part of the test within the set

time limits. This was reflected by P6 and P7, who both described the experience of being interrupted during the mock speaking test. When P7 brought up being interrupted, she initially framed the experience in terms of her own perceived misunderstanding of the expectations of part one:

I also did notice that ... the moderator cut me off during, like, during a couple parts in part 1 [of the mock speaking test] because I was talking. I was going on a tangent, and I didn't realize that it was a short question. (P7)

When asked if this affected her performance, P7 expressed a belief that it did:

I was going on this train of thought and then he cut me off and then I had to, and then I was asked a different question and then I had to restart my thought process again [laughs]. (P7)

While P7 focused on the experience of being interrupted during part one of the speaking test, P6's comments related to being interrupted at the end of part two. P6 did not believe that being interrupted impacted his performance but he does describe modifying his approach to part two to avoid being interrupted.

I just didn't really want to be in that awkward situation where I was gonna like in the middle of, like something that I was so passionate about and like, alright, let me cut you there, let's move on. Like damn, alright, so it's like, alright cool let's just do it a bit slowly. I kind of like <time> myself. I would rather do it just below two minutes maybe like 1:30 or something like that instead of being like in that awkward situation where, you know, again, I was talking and I gotta stop because, you know, it's like, all right, it's two minutes time to move on. Yeah. (P6)

Length of the Test in its Entirety. While the above quotations focused on the length of the speaking test in particular, P3 and P6 described the impact of test length from the perspective of the full IELTS test (i.e., all four sections of the test when taken in succession). When describing factors that influenced her performance on a real IELTS test she took in 2018, P3 mentioned the length of the IELTS test in its entirety:

I would say probably fatigue [impacted my performance on the IELTS test] because we had to wait for a long time. There was a lot of us packed in one room and obviously we have to take turns, so we could be there for like 4 or 5 hours. And by the time it's your turn, you're just tired and you don't want to be there, yeah. (P3)

Speaking more generally, P6 expressed a similar experience:

it was long, it was boring, and so repetitive as well, right? So when you are like so physically and mentally tired, it doesn't really matter how good you are, it may impact your performance. (P6)

Both P3 and P6 identify the length of the IELTS test in its entirety, which is often completed in a single day depending on the testing location, as a cause of fatigue which in turn negatively impacted their performance.

4.2.2 Theme 2: Appraising the Relevance of the Test to Academic and Everyday Contexts

All seven of the participants in this study had experience living in Canada, and most had experience studying at a Canadian post-secondary institution. As such, they were able to offer their insight into the relevance of the IELTS speaking test to both everyday and academic settings. Part one of the test was the most widely commented on. Generally, participants found part one to be relevant to their English use in day-to-day communication but found the relevance

of this part of the test mainly limited to the social interactions required in university life. Considering the test more broadly, participants did note the relevance of IELTS to persuasive communication and applying for co-op positions. However, P2 and P4 also noted that the actual speaking requirements of their specific academic programs were quite low.

Sub-theme 2.1: Relevance to Communication Needs in Everyday Settings.

Regardless of whether the participant described their need to take a standardized English test as being for academic or immigrations purposes, as residents of Canada, all participants had experience of living and communicating in everyday settings.

Communication Needs in Everyday Setting. When describing their everyday needs, the participants indicating a mix of social communication (such as making friends), running errands (such as grocery shopping), interacting with professionals (such as doctors and teachers), and fulfilling workplace communication needs (such as interacting with coworkers). Discussions of the relevance of the content and skills tested in IELTS revealed a considerable amount about how the participants use English and what they perceived as being meaningful or necessary skills to have while living in Canada, with P1, P3, P5, P6, and P7 commenting on the relevance of the IELTS test to their every day communication. The exact meaning of every day communication was left open to the participants' interpretation. The kinds of everyday communication acts mentioned by the participants are given in Table 30.

Table 30

Everyday Oral English Communication Acts Identified by Participants

P#	Everyday communication acts
P1	presenting ideas, arguing ... using the terms I need to
P3	make new friends; speak[ing] to people and ... respond[ing] to people; doing my job well; going to the grocery store to buy groceries
P5	[telling someone about] me and my sister[‘s] relationship; [communicating] with government, whenever I used to appoint vaccine for my child for my baby, so I have to speak in English that time. For appointments, for going to the doctor.; [communicating] with teachers or and or the doctor other than my family doctor and at the desk for the receptionist, like for appointments and immunizations, like these kind of things.
P6	[going] grocery shopping, or I don't know, gas, for instance, or if you I don't know anything really, you go to the gym, and then you just kind of hang out with your friends.
P7	having a regular conversation with a friend or just a random person; back and forth communicating with my coworkers or um friends in general; [discussing] day-to-day life, just like, my daily routine and- “how did that go?” ... having, like, a casual conversation

Note. P# = Participant number.

None of the participants in the present study took the IELTS test specifically for access to the job market. However, P3 and P4 did discuss the communication needs of the workplace. P3 described some of the communication needs of her position working in a medical office, stressing the importance of accurate communication when performing these tasks:

I mean you need to learn how to communicate (???) accurately because in my line of work I'm required to do scheduling and booking a lot and you need to let people know like, you know, you're booked in and alright, you have to cancel this with that. So, if you're not able to communicate with them properly, they'll just turn up on the spot when you're not booked in and it's gonna mess up the entire schedule. I think communication is the main thing. (P3)

P4, who was studying to work in human resources, also stressed the importance of accuracy in his intended line of work, saying:

So for example an engineer, then you can you can make, many much more mistakes when speaking on when writing, but if you want to be human resources and you'll be the head of the company, maybe or like strategic position so, you're not allowed to do many mistakes in English. (P4)

Relevance of Parts of the Speaking Test to Everyday Communication. Participants who discussed the relevance of part one of the test to everyday communication were, on the whole, quite positive. P5 in particular emphasized the kinds of questions asked in this part:

So that's why and yeah, those questions really matter, because if I communicate with any other person who spoke English, then they are really very important I don't- I just know how to speak them correctly, yeah. (P5)

P3 and P7 discussed the relevance of part two to everyday communication. P7 was more positive in their appraisal, with the noteworthy detail that it was their response was motivated by the content of the question they were asked in the mock test, saying:

I guess part two [was relevant], um... I talk about that store a lot, like Makers, because people would ask me "oh, where do you get gifts from" or whatever, and I would just say that that's my favorite store to go to. So, because I bring it up a lot, I ticked strongly agree for- for daily communication (P7)

P3 was more critical of the relevance of part two. In her case, the focus was on the format of part two rather than the topic:

I'll talk about this topic for two minutes, who is going to ask you to do that in the real world? Like, that's not gonna happen (P3)

P7 identified the interactive nature of part three of the speaking test as particularly relevant to their daily communication needs, saying:

think part 3 is probably more relevant because like I do a lot of just back and forth communicating with my coworkers or um friends in general, and it's just- we talk about random things. So, it's... it's definitely more relevant (P7)

Sub-theme 2.2: Relevance to Communication Needs in Academic Setting. Six of the seven participants reported that at least one of their reasons for taking IELTS or another standardized English test was for admission to a post secondary institution. With these participants also reporting at least some experience studying in Canada, there were able to offer both a perspective on the kinds of oral communication required of them while studying and the relevance of the IELTS speaking test to their academic oral communication needs.

Communication needs in an academic setting. As is shown in Table 31, the six participant who discussed academic communication all mentioned at least one specific communication act during their interview. Giving presentations, presenting arguments, and participating in discussions were identified by multiple participants as part of the oral English communication requirements of their academic program.

Table 31

Academic Communication Needs Described by Participants

P#	Academic Setting Communication Acts
P1	engage in discussions, also argue your points,
P2	presentations and... group work... in like the later part of, like, our undergrad; preparing for my co-op interviews
P3	discuss on certain topics; presentations
P4	disagree or agree and or you have, you have to convince your professor or your clients; discussion with the with clients or with a colleague and I have to present arguments and to give my position.
P6	lead the whole conversation
P7	talking to your classmates

Note. P# = Participant number.

Relevance of Part One to Academic Communication. P3, P4, and P7 commented on the more personal questions typical of part one. P3 noted that these questions were perhaps relevant to the social requirements of university life but not the academic requirements.

Yeah, I feel like for familiar topics, I feel like it's only relevant probably for the first few days when you're just meeting like, the ice breaker period, where you can talk and make new friends, but other than that, once we start getting into like academic topics, no one's gonna ask you like you know, what do you like about home or like, what's your favorite subject, like that is not relevant anymore. (P3)

P7 similarly drew a distinction between social conversation and academic communication requirements.

I think it's because in my program it's very... Because I took communications ... in university and it's a very like, uh, like theory-based program, I would say. So, there's not

a lot of like day-to-day conversations that you would have um when you're in a lecture or when you're talking to your classmates, um it just wasn't really relevant, I guess, the topic that we talked about in the test. So, I guess academically, for me anyways, it just wasn't... um it just isn't a topic that would get brought up, but I guess like if you're having a regular conversation with a friend or just a random person, you would bring those topics up for sure. (P7)

While P3 and P7 noted some utility of these questions, P4 found the more personal questions in part one of the test were not reflective of the speaking requirements of his program at all.

Because during the... three months, maybe, during the four months studying this program I didn't see like, I wasn't never asked like to... talk about anything that I like or about any kind of those questions at all. Even to present myself, nobody has like [laughs]. (P4)

Relevance of Speaking to Academic Communication Needs Generally. One issue that a general (rather than program specific) academic assessment will encounter is the differing communication demands between programs. P2 and P4 both indicated that speaking was not a major component of their academic programs. P4 downplayed the importance of speaking in their academic program, saying “I don't feel that we worked on we work on... speaking and communication so it's more about writing and it's not about presenting” (P4). However, he did find the questions from the mock IELTS test that required persuasive argumentation relevant to his academic English requirements, noting that “we go in these kind of discussions, and you'll disagree or agree and or you have, you have to convince your professor or your clients, or anybody about anything” (P4).

P2 was also able to identify some oral English requirements in his academic program, but like P4, considered them a skill of secondary importance:

in my program, speaking is not like, major part because I am in computer science. It's like mostly like coding, or if you were like a good programmer, that's what you need, not like maybe, we do, like, have some like presentations and like, group work, but that comes in like the later part of, like, our undergrad. (P2)

While both P2 and P4 indicated that speaking was not a major component of their programs, both participants identified speaking needs when transitioning into the job market. P4 discussed the importance of speaking skills when interacting with clients and colleagues. P2 described part two of the IELTS speaking test as relevant to preparing for co-op interviews, which many university students in Canada do as part of undergraduate and graduate programs.

I was also preparing for my co-op interviews and something they would allow me to give some time to think about a question. So, it's like in good exercise to go back and forth to be, to think about a question and just to be come up with like a scenario and to create an example in my mind. So yeah, it's a good exercise for my program. (P2)

4.2.3 Theme 3: Describing the Costs and Benefits of the Test

The third theme identified from the coding data was participant discussions of the costs and benefits they associated with taking or having to take the IELTS test. The three most commented on were affective (i.e., taking the test or needing to take the test was a source of a positive or negative emotion), monetary (i.e., the cost to take the test), and informational (i.e., taking the test or needing to take the gave the participant information or preparation which they found beneficial).

Sub-theme 3.1: Affective Impact. All seven participants described an emotional impact of taking or needing to take a standardized English test at least once. These are shown in Table 32.

Table 32

Participant Descriptions of the Emotional Impact of Testing

Emotional Impact	Participants Describing Experiencing Impact
Anxiety, Pressure, Stress, Nervousness, or Intimidation	P1, P2, P3, P5, P6, P7
Boredom	P2
Confidence	P2, P5
Demotivation	P2
Not Belonging	P2
Resiliency	P6

Note. P# = Participant number.

Feelings of anxiety, pressure, stress, nervousness, or intimidation were the most commonly reported negative emotional impact. The impact that these feelings had on test performance have already been discussed. While participants largely focused on the impact of these feelings in the moment of testing, P3 described a period of post-test rumination after both taking a real IELTS test and after taking the mock IELTS, saying “after that [mock] test I was pretty much going through the same process as I did after the [real] IELTS. I was like, why did I say that? Why didn't I do that?” (P3).

In general, participants described the emotional impact of taking the test rather than preparing of it. P2, however, did include the process of studying and taking practice tests in his description of how preparing for and taking the IELTS test impacted him:

sometimes you get like demotivated when you don't feel like you... are getting like a good score when you give like a mock test. And it's also like an anxiety, and the stress,

it's mentally draining sometimes when you have to study and repeat the same process for like a longer period of time and reduce like motivation and you don't feel like.. you belong somewhere and you feel like bored enough. (P2)

While the role of anxiety was largely viewed by the participants as negatively impacting test-taker performance in the moment, P2 also identified potential benefits of this pressure in terms of longer term impact:

I think it definitely creates a pressure on my mind when we have to, like, get some standardized tests, but definitely, like, less pressure I know it's like now graduated from school that prepares you well for to take like more difficult and difficult standardized tests. So, definitely, like, the pressure is like still but it's like less and I have like more confidence now than before. (P2)

Similarly, P5 also described IELTS as giving her confidence, saying: "I came to know about me that how much can I understand and how much can I speak or how much confident I am. So yeah, it helps me positively" (P5). When asked about the ways that preparing for and taking the IELTS affected him positively, P6 expressed a somewhat similar sentiment, saying: "if I have managed to overcome, like all of these different obstacles, and some of them are just dumb, then, you know, honestly, I can just take on the world" (P6).

Sub-theme 3.2: Monetary Cost. Three participants discussed the cost of taking the IELTS test. P3 describes the test as expensive, particularly for international students who may wish to immigrate. She connects this to the fact that IELTS tests only remain usable for two years, saying: "we don't need extra financial problems for something that's only gonna last like 2 years and just for university admission." (P3)

P1 did not explicitly refer to the test as expensive, but she does connect the fact that the test does cost money as potentially impacting test performance.

So, and for example, for the second time I did the test, I went with this to the exam because I was so nervous. So, I think, like, it's stress, something, it's a lot of things that can, you know, because if you fail, you need to pay more money to get a test. (P1)

P6 was perhaps the most pointed in his critique of the cost of taking the IELTS test. P6 also described the test as expensive but goes on to question the motivations of the test users and test owners because of a financial incentive to increase the number of test-takers:

So, I think it was just kind of like again, just the system that sort of like made people do that just because they knew they would pay. They would have to pay for that. And it wasn't cheap. It's still not, right? So, it's just like such an unfair approach to the whole thing here I feel like. (P6)

It is interesting to note that all three participants connect the monetary cost of the test to another aspect of the test.

Sub-theme 3.3: Informative Benefits. When considering the benefits of the test, P2, P5, P7 described the positive impact of needing to prepare for and take the IELTS test in terms of providing information. P5 described how preparing for IELTS helping her increase her English ability, saying, “It helps me to... know many words, many vocabulary which I can use over here” (P5). P7 highlighted the benefits IELTS provided by preparing her for her future life in a primarily English-speaking environment.

I guess in general it has affected me positively cause then I know sort of what to expect when I'm in a foreign country. For example, when I first moved here and I took the

IELTS test beforehand, like, I know what to expect when I'm approached by a stranger when they're asking me about, like, random things and it was also a good way for me to, I guess, practice learning how to have day-to-day conversations in English, because prior to that I was speaking my primary language in in Hong Kong, which was Cantonese, with my family and friends (P7)

P2 expressed a similar sentiment, focusing on how preparing for the IELTS provided information about different cultures as well as preparing him for the experience of studying in an English-language post-secondary environment.

I think in a positive way, it's definitely like learning about the different culture and practicing like how I would fit into a different culture and to be like- make some decisions about my life with, uh, I would fit into my - into a different culture while learning about a different language and preparing for a test, and how like people study here and what are like their study techniques. So, getting familiar with the process and it was like very engaging. (P2)

These comments are reflective of a larger trend emerging from the interviews. Namely, while participants were critical of aspects of the standardized tests that had taken or were about to take, they also viewed having English language skills and being able to engage fully in an English-language environment as important both in academic and non-academic contexts.

4.2.4 Theme 4: Considering whether there should there be a test, what kind of test, and who should take it.

In general, the participants saw a gatekeeping English assessment as necessary both for immigration and academic purposes. While the participants generally viewed IELTS neutrally or

favourably as compared to its competitors, there was a strong desire for alternative ways of demonstrating English abilities to be made available. On the question of L1 English speakers, participant opinion was split with some rejecting the idea as unnecessary and others supporting the idea both out of a sense of fairness and speaking from the experience of L1 English speakers communication abilities not uniformly being as strong as they may need to be.

Sub-theme 4.1: Considering whether the test be necessary for PR. P5, P6, and P7 discussed the necessity of taking the IELTS test for permanent residency. P5 and P7 participants believed that the inclusion of IELTS in the PR application process was appropriate. P5 focused on the length of time in Canada that is usually required before applying for PR or citizenship:

[PR applicants] should take the test. And for citizenship, because they have been living over here for three or more years, so they must know English, like kind of they must be....

I mean, they come across for so many people, so they learn English how to speak. (P5)

The emphasis on the importance of being able to communicate was also reflected by P7's comments.

I think having at least a working proficiency or being able to have a small conversation with someone is important, especially if you're trying to make a living here, or just live in a primarily English speaking country, I think it's important. (P7)

P6 did not the argue against the inclusion of IELTS in the PR process, but they were critical of the current PR system in which most applicants will need to take an English proficiency test. P6 argued for a more "case-by-case" approach to the inclusion of English testing, using their own situation as an example:

having two degrees from two English-speaking countries, I was actually asked to do it again for my PR application. Which I thought was pretty dumb, because, you know, I was actually like having a conversation ... with someone from the Canadian immigration authority, I just asked them like, hey, like last time I checked like Canada, U.K. spoke English, I don't know if it's changed. (P6)

Sub-theme 4.2: Considering whether the test should be necessary for citizenship. P7

expressed the opinion that reaching a minimum IELTS score was even more important for obtaining Canadian citizenship than gaining admission to an academic program because:

some people are applying for citizenship because, like, they want to stay in Canada, permanently, um, find a job you know, have- build a life and whatnot, and I think that's even more important to be fluent in English (P7)

However, P7 did clarify that she thought requiring applicants to take an English test for both PR and Citizenship was unnecessary:

if you're applying for PR and you're required to do an English speaking test, then when you apply for citizenship, you won't be, you shouldn't meet to do one, because it's probably within, like, two or three years that you're applying for PR and citizenship, so, and I don't think your English skills would, like... decrease it within that period of time. So yeah, I just don't think it's necessary to take English tests for both PR and citizenship, but definitely PR. For sure. (P7)

Sub-theme 4.3: Consider whether the test should be necessary for University

Admission. P3, P4, and P7 all described IELTS, or another English-language proficiency test as necessary, at least in some cases, for university admission. P3 was critical of IELTS as it is

currently administered due to the effects of anxiety discussed earlier, but did describe some kind of English-language proficiency test as necessary. P4 focused primarily on reading and writing, saying:

We need like writing skills. Also, also like vocabulary, grammar like you, it's for, per session, maybe we have to review uh, 30 page, 30 pages in our textbook. So if you can't like with sometimes with specific vocabulary. So if you're not like, if you can't, if you can't read quickly, if you can't understand things quickly, and if you can't feel comfortable when writing, then you can't succeed. (P4)

P7, who listed both English and Cantonese as her L1, echoed this sentiment, using her own experience as a student:

I do notice that sometimes when I'm put in a group with some of these international students, their English skill level is very limited and I think it probably, like, hinders them from actually learning or receiving education, like, in the most, in an efficient or effective way through school, like, through classes and whatnot, because they're all English speaking.

P7 also reflected on how the experience of interacting with lower-proficiency English-users affected her personally:

Yeah, and I also think it's, like, it's harder to work with these international students sometimes because their proficiency is so limited and it just required; it just causes a language barrier for some reason which probably shouldn't happen in an English-speaking university, so yeah. (P7)

While the participants were broadly in agreement that having English-language proficiency was important for academic studies at English-language university, P2 questioned the need for the test on the grounds of his own previous academic experience:

is there like a really a need for a standardized test even though we like study English?

And they were like something, as I got a good score in like since my like high school,

also they asked for my academic transcript. I had some really good score in grade 11 and grade 12. (P2)

Sub-theme 4.4: Considering whether L1 English speakers should have to take the test. As of the time of writing, L1 English speakers wishing to attain PR status in Canada will typically be required to provide a standardized language test score in the same way as any other applicant. For citizenship and most university applications, however, obtaining a high school diploma from a predominately English-language secondary school, in an approved list of majority English-speaking countries will allow the applicant to bypass this requirement.

P5, P6, and P7 expressed the opinion that testing L1 English speakers should not be required to take a standardized English test. In answering, both P5 and P7 focused on the functional rationale behind requiring the test. P7 expressed the opinion that testing L1 English speakers was unnecessary:

I do see why an English test would be required, particularly for citizens of countries that- where English isn't the first language, I completely agree, but I just feel, like, for the United States, I don't think ... it should be a requirement because yeah, like you said, it's given like the proximity and how similar, um, both countries are, ... it's just redundant, I think. (P7)

P5 shared this opinion, saying:

Because I think that... if they have... their... native language is English, so I don't think that they do need a kind of test because they already know how to speak. (P5)

P6 framed the question L1 English speakers specifically in terms of the fact that IELTS is a money-making venture for the test-owners:

the whole system gradually makes so much money out of this. Because, regardless of like what type of IELTS you take, you would still have to pay, I believe 320 dollars for a single test. So just imagine, if, like.... for like, when the same day you have like 10 people taking the test, you would make like 3,200 dollars. So that's like a lot of money. So I feel like the system was made just to kind of like to facilitate that. And they're like, well, if we can make more dollars out of this, why not just make them do that regardless of their nationality? I don't care if they're Irish, Australian, American. Make him do it [laughs].
Yeah. (P6)

P2, P3, and P4 expressed the opinion that, at least in some cases, L1 English speakers should be required to take a standardized English test. On the whole, these opinions were justified with an appeal to a sense of justice, fairness, equality, and so on.

Yeah, in order to create like fair and equal like amount of chance for everyone. If I have to give like a certain test then why I, ONLY I, have to give? Cause I want like everyone to give it and to be in my like shoes and be like- have some like kind of like understanding... (P2)

How to evaluate someone who is coming from X country in X continent and other country in other continent to not mention any country. Yeah. What, what fac ... it's too

complicated, and it's not objective at all. And maybe he spent there five years and spent in the other country 10 years. Then what we're going to do?... Just do the test for anyone. (P4)

While P3 similarly justified her opinion with a call for greater fairness and objectivity, she also included the need for English language skills as part of their reasoning, saying:

I just don't think it's fair. It's - I think it's quite biased, in particular, like why is it that only international students have to prove that they can, you know, enter a university program, that they can speak English? If it's a matter of English proficiency - like proficiency, shouldn't like locals have to do it too? (P3)

P3 expanded on this idea by noting that, from her experience, L1 English speakers are not necessarily more effective communicators, calling into question the motivations behind the use of the test:

I've seen many international students who could probably speak better than locals here too. So, why do they have to prove it just because they're born in a different country, even though their official language in that country is English? What makes them any different from people who were born here? I don't think it's fair. I think it's an additional barrier for international students. (P3)

Sub-theme 4.5: Consider whether alternatives forms of assessment be accepted.

While participants broadly viewed IELTS favourably as compared to other standardized English exams, there was also broad agreement that other methods of proving one's English ability should be made available. Table 33 shows the wide range of proposed alternatives, with the three most frequently suggested alternatives involve allowing some level preparation (for example by

allowing test-takers to give a prepared presentation), find methods of less formal assessment, and in general finding ways of reducing test-taker anxiety.

Table 33

Proposed Alternative Ways of Proving English Abilities

Proposed Alternative Ways of Proving English Abilities	Participants Proposing This Alternative
Combine Multiple Skills	P4
Consider Courses, Workshops, Certifications, or Volunteer Experience as an Alternative to a Standardized Test.	P2
Consider Test-Taker Backgrounds	P6
Find Ways of Reducing Test-Taker Stress or Anxiety	P2, P3, P6
Give Prepared Presentation	P1, P3
Give Time to Prepare in Advance	P3
Give Topic in Advance	P3
Incorporate Video Stimulated Discussion in Speaking Test	P4
Use an Interview to Prove English Abilities	P3
Let Test-Taker Choose How They are Assessed	P7
Let Test-Taker Choose Topic of Assessment	P4
Let Test-Takers Provide a Recording	P7
Observe Test-Takers in a more “Real World” Setting	P3
Format Assessment as a Less Formal Conversation or Chat	P1, P3, P5, P6
Spread Out Assessment Over a Longer Period of Time	P6
Use A Combination of Prepared and Unprepared Elements	P4

Note. P# = Participant number.

Allow for greater preparation. When proposing alternatives to the currently used testing methods, several participants described means of proving one’s English that incorporate some element of pre-test preparation. P1 proposed giving an oral presentation about a familiar topic. The proposal was echoed by P3, who clarified that this kind of demonstration of English speaking ability would be particularly appropriate for academic admission purposes:

For academic purposes, I mean they - if they want to, they could give you a topic in advance and make you prepare for a presentation because that gives you a chance, a

chance to prepare and you need those skills for university degrees, that presentation skills. So, I think that would be very relevant for school, but for, like, immigration purposes, for, like, PR. Not so much. (P3)

P3 defended her proposal not only on the grounds of relevance to the communication needs of a university setting, but also on the grounds of representativeness of their overall abilities.

According to P3: “I feel like there's a lot more in here that I can't say on the spot. I just make things up on the spot. But if you give me some time to prepare, I can do a lot better for it.”

Find methods of less formal assessment. A trend that emerged from several of the participants' proposals for alternatives was reducing the perceived formality of the assessment. When discussing how this could be accomplished, P1 described it as “meetings... conversations, interviews, like where you can express yourself.” P3 described an alternative that was “a chat instead of a full on interview” and imagined what this assessment might look like: “maybe stage it as something where you go in the shop and you want to buy something or you need to, you know, talk to someone to discuss, but not make it so rigidly formal.” P1's comments about self-expression and P3's description of a more naturalistic assessment were mirrored in comments made by P6:

I think I would benefit more if there was like a program where I could be paired with like a native speaker, and just hang with them. Really, you know, like, just talk about things and stuff and they could sort of like point... out like mistakes I would make or like I made and then, you know. It's more like a casual setting where, you know, the pressure is not as immense, and then y- I would be able to kind of like.... express myself a bit more
(P6)

P6's comment about how this alternative method of demonstrating English abilities would reduce pressure of a broader theme within several participants' proposed alternatives: finding ways of reducing test-taker anxiety.

Reduce test-taker anxiety. In addition to the above comment made by P6, P2 and P3 both explicitly described reducing test-taker feelings of anxiety or pressure as a change that would be advantageous for a proposed method of demonstrating English abilities. Reflecting the statements quoted above about reducing the formality of the assessment, P2 described the need for an alternative that was more "practical" in which test-takers could:

enjoy themselves, where they could feel like they're not giving like a test, where they're not bound in a time-based scenario and they would - it would give them like equal opportunities to discuss their opinions and perspectives. (P2)

Similarly, P3 described the advantages that having a lower-pressure form of assessment could have on both test-taker experience and test-taker performance. In their words, a preferable means of assessment would be one that:

puts the interviewee at ease, so they won't have, you know, they won't panic. They won't have anxiety to make them forget that this - you have one shot to do this. Because if you just make them at ease likely, you know, conversation will flow easier. They won't freeze up, or panic, or have anxiety. A lot of people, not just people with anxiety, right. If you feel comfortable, obviously your speech will be smoother and ideas will flow faster. (P3)

It is noteworthy there was such wide agreement that alternative forms of assessment should be made available and that so many of the participants would identify reducing the formality of the assessment and/or reducing test-taker anxiety as a change they would like to see. This trend

seems in-keeping with the significance participants placed on anxiety and nervousness when describing both factors that influenced their performance and negative consequences they experienced from taking the test.

4.3 Summary

This chapter presented the seven participants' responses to the questionnaire and the themes identified during the coding process which were most relevant to addressing the questions of the fairness and justice of the IELTS test as used for immigration and academic purposes.

Questionnaire and coding data reveal that there is wide agreement that factors other than one's English ability may affect performance on the test. In particular, the participants identified feelings of anxiety, pressure, stress, nervousness, or intimidation as a major factor impacting their performance on the test. The perceived variability of test topics was also identified as a factor in performance. Lastly, participants identified timing as a factor in how well they did on the test. Some participants noted that speaking at length in part two of the test could be challenging. However, the participants generally found that having a longer amount of time to speak uninterrupted impacted their performance positively.

While the relevance of the assigned topic given to one's field was viewed as impactful on performance, participants viewed the relevance of the test overall as variable. Part one in particular was viewed as relevant to everyday communication and the social requirements of attending a mostly English-language university. While participants did see the persuasive communication required in an academic setting reflected in the IELTS speaking test, some participants found that the actual speaking requirements of their academic programs were fairly limited.

Participants viewed the impact the test had on them both positively and negatively. The primary negative impacts described were emotional, with anxiety, pressure, stress, nervousness, or intimidation being the most prevalent negative emotions described. While the participants tended to focus on the negative emotions they experienced, feelings confidence and resiliency were described as positive emotional impacts of taking the test. The other positive impact described by the participants was information gleaned from taking the test. IELTS was identified as having providing participants information about their English level and helping prepare participants for life in an English-speaking environment.

Lastly, participants were largely in favour of keeping an English proficiency requirement for immigration and academic purposes but were split on whether L1 English speakers should be included in this requirement. While questionnaire data showed a generally favourable view of IELTS compared to its competitors, there was broad agreement that alternatives to current standardized English tests should be made available to serve the purpose of proving one's English language ability.

In the next chapter, I discuss how these findings help answer the questions of test-taker perspectives on the fairness and justice of the IELTS speaking test. I then discuss the theoretical, methodological, and practical implications of these findings as well as limitations of the current study and future areas of research.

Chapter 5: Discussion and Conclusions

In this chapter, I discuss and interpret the findings from the previous chapter as they pertain to the questions of justice and fairness in language testing. In section 5.1, I discuss how the findings answer the research questions of this study. In section 5.2, I describe some of the methodological, theoretical, and practical implications of the study as well as some of its implications and suggest some areas for future research. In section 5.3, I conclude this thesis.

The research questions put forward in Chapter 2 were:

1. Do test-takers view the IELTS speaking test as fair (as conceived as both construct validity and relevance/utility)?
2. Do test-takers view the IELTS speaking test as just (as conceived as both value implications and social consequences)?

The present study required both the interview process and coding process be approached with a spirit of openness. While of course, researcher positionality is always present at all stages of the research process, the approach was to give the participants a chance to express their opinions about the aspects of the test that mattered to them. Likewise, the coding process tried to avoid artificially creating codes or themes that fit neatly into the language or theoretical framing of the research questions. As such, there are countless aspects of fairness and justice that the participants did not address. Similarly, there are topics that arose which challenge the neatness of the divide between fairness and justice or the quadrants of Messick's aspects of unified validity. Rather, construct validity, relevance/utility, social consequences, and value implications are used as a language and theoretical underpinning with which to unpack and explore the often complex ideas and experiences of the seven participants in this study.

5.1 The question of fairness.

5.1.1 *Do the participants view IELTS as fair from the perspective of construct validity?*

Construct-irrelevant variation emerged as the major challenge to the fairness of IELTS in this study. Participant focus on the role that nervousness plays in performance on the speaking stood out at a particularly strong challenge to the construct validity of the test, in particular Messick's Generalizability aspect of construct validity. This is especially relevant to the IELTS speaking test grading criteria of fluency, where a test-taker's ability to speak at length on a given topic with minimal hesitation can have a direct impact on that test-taker's score. When participants reported "going blank" or worrying about what they "should say" during the test, what they are describing is, at least in part, construct-irrelevant variation. The counter-argument to this could be that nervousness while speaking English generally is relevant to the constructs evaluated by the test. If a speaker is hindered in communicating in English because of nervousness, a test-user could make generalizations about that speaker's ability to communicate orally within the real-world context of interest (i.e., university, life in Canada, and so on). However, the participants in this study overwhelmingly identified the fact that the IELTS speaking test *is* a test, with associated costs and stakes, as an explanation for *why* they felt nervous. In other words, the participants argue that the generalizability that can be made from their performance during the speaking test is hindered by feelings of nervousness, which is a challenge to the fairness of the test.

That the variability of test topics is relevant to test fairness in multiple, sometimes compounding ways. There was a broad consensus among the participants that the topics given by the examiner during the test have an impact on the test-taker's performance, reflecting findings by Hoang (2019) and Liu and Hamid (2024). This is not surprising given that test-takers are

evaluated on fluency and lexical resource. Being given a topic about which the test-taker has something to say, and the vocabulary to do so, will give that test-taker an advantage. This is a perspective reflected by several of the participants. However, the participants identified the test topic as impacting performance in indirect ways as well. Several participants noted that their knowledge that the topic could influence their performance on the test was itself a source of nervousness which, as already stated, was identified as a factor in test performance.

Lastly, the role of timing (i.e., the impact of the length of the test, or not having enough time to complete a test or task) emerged as a consideration from the perspective of construct validity, and therefore fairness. The tension described by the participants in this study was, in the very broadest of terms, whether a greater amount of time to demonstrate one's abilities would increase the overall fairness of the test. There was some acknowledgement that speaking at length in part two of the speaking test can be challenging for some test-takers. Challenges pertaining to the length of part two, along with the multiple hours required to take all four modalities of the IELTS test, might suggest that any additional lengthening of assessment is inadvisable from the perspective of fairness. However, the more prevalent opinion expressed by participants was that having ample time to demonstrate their language abilities, including the opportunity to speak at length without interruption, was associated with a higher degree of test fairness, in line with similar findings from Cheng and DeLuca (2011). The association between greater time to speak with greater fairness was further supported by the fact that several participants, when describing an alternative method of proving their English abilities, described scenarios that would likely require a greater amount of time to complete.

5.1.2 Do the participants view IELTS as fair from the perspective of relevance/utility?

The question of test fairness paints a more complex picture from the perspective of relevance/utility. From the perspective of test content, there is a sense that IELTS can be relevant to the testing universe in which it is used. That this is a variable viewed as a product of “luck” is a challenge to the fairness of the IELTS speaking test, a sentiment also found by Hoang (2019).

When the issue of test content is put aside and focus is shifted to the *kinds* of communication employed in the three parts the test, IELTS fared quite a bit better in the eyes of the participants of this study. When addressing their communication needs in Canada, several participants described various kinds of informal interactions that require sharing personal information and experiences in contexts that often require shorter turns from the speakers. It is appropriate then that the participants found part one of the speaking test particularly relevant to their daily communication needs in Canada.

Considering the relevance/utility of the test for academic purposes as three separate sections, all three parts were generally viewed as relevant. In the interviews, part two was the most widely discussed in terms of its relevance to academic communication needs of the participants. This is in part because it required participants to speak at length, as might be done in a presentation, and required participants to demonstrate their ability to develop ideas and arguments. The appraisal of part two’s relevance to academic communication is supported by the fact that several participants identified presentations or an otherwise prepared speaking activity as a possible alternative to the IELTS speaking test.

When considering the speaking test as whole, the participants generally asserted that certain parts are relevant to some of the communication needed for everyday life in Canada,

other parts are more relevant to academic communication, and all parts' relevance is contingent on the kinds of topics that are given to the test-taker. An issue that this raises is the fact that all IELTS test-takers, regardless of their reason for taking the test, will be given the same speaking test. This is because the speaking portion of the test is identical for IELTS Academic and IELTS General Training, exposing the IELTS speaking test to the challenge that it is the product of a change retrofit as described by Fulcher and Davidson (2009). However, this is complicated by the fact that the participants found some aspects of the speaking test relevant to an academic setting (e.g., speaking at length in part two of the test) and other aspects more relevant to daily communication (e.g. the familiar topics discussed in part one of the test).

5.2 The question of justice.

5.2.1 Do the participants view IELTS as from the perspective of social consequences?

The social consequences of a test can be considered in many ways. These range from negative washback in the education system, to the broader impact of qualified professionals being blocked from working in their adopted country, to current or prospective immigrants having their application process delayed. The term social consequences lends itself to big picture thinking, but this can sometimes cause researchers to focus on broader social trends at the expense of the personal experiences of stakeholders. There is the epistemological problem in that personal experiences often point to broader social trends that warrant big-picture investigation but may have been missed. There is also the ethical problem suggested by Shohamy (2001b) and Hamid and Hoang (2018). Namely, ignoring the individual consequences of a test perpetuates testing policies that ignore and dehumanize already marginalized stakeholders. Undemocratic and dehumanizing testing policies are themselves a social consequence that need to be considered when examining the justice of a test. That the participants in the present study focused mainly on

how the test affected them personally is not an indication that their experiences should not be considered using Messick's term social consequences. On the contrary, it is vitally important to the discussion of the social consequences of IELTS include these individual experiences.

That the participants of this study focussed so heavily on the emotional consequences of the test is revelatory. The seven participants of the present were largely able to successfully achieve the scores they needed for their previous academic or immigration goals and the stories of multiple retests between lengthy stretches of studying were not as prevalent as in other studies. This is reflected in questionnaire responses pertaining to delays and other negative personal consequences having some of the lowest agreement rates.

However, this did not stop the participants in this study from describing both feelings of accomplishment upon their successful completion of the test and also feelings of anxiety, pressure, and self-doubt before during and after sitting for their exam. The centrality of emotional and psychological consequences echoes the earlier findings reported by Cheng and DeLuca (2011) and Hoang (2019).

5.2.2 Do the participants view IELTS as just from the perspective of values?

There are numerous ways to approach the question of values within the test. To aid with this, I return to McNamara's (2013) from of the question: "whose values are represented in the test, and whose interests are advanced through its existence and operation?" (p. 1). It is the second half of McNamara's question that is most relevant to the value implications addressed by the participants of this study. There is an argument to be made that whether or not English assessment should be included for university admission, PR applications, or citizenship applications can be discussed as an issue of values.

The general trend emerging from the interviews is that the participants view the inclusion of language assessment generally in the academic and immigration process is just insofar as an assessment of English is viewed as necessary. While several of the participants were critical of aspects of the IELTS test, there being some kind of assessment was deemed as not only necessary but was also framed as being in the best interest of prospective international students and immigrants themselves.

The strongest criticism of the justice of the test from the perspective of values was the perception levelled by some participants surrounding IELTS as a bureaucratic necessity and as a money-making venture, similar to the findings of Hamid et al. (2019). That language assessment is viewed on the one hand as necessary but on the other as placing a burden on the test-taker is reflected in the mixed opinions expressed relating to L1 English speakers being required to prove their language abilities in some contexts but not others. Test-takers are able to consider the necessity of the test on practical terms, possibly suggesting that giving the test to L1 English speakers is not necessary for the purpose of ensuring adequate language abilities. However, they are also able to see the test as an instrument of power and arguably financial exploitation. In which case, insisting on equitable treatment could logically take the form of treating L1 and L2 English speakers equally in the eyes of the testing policy.

5.3. Implications

Having discussed findings of the present study, I now turn to theoretical, methodological, practical, and reflective implication of these findings. The theoretical implications pertain to the contribution this study makes to unified validity, the role of test-taker experience in test validation, and the ongoing discussions of democratizing and humanizing language testing. The methodological implications pertain to the present study as part a work of qualitative research

within test validation research. The practical implications pertain to how the findings of this study might be translated into action by test owners and developers, test users, and test takers (as well as educators supporting these test takers). Lastly, the reflective implications pertain to directions this study points to for my own practice as a teacher and researcher.

5.3.1 Theoretical Implications

Support for a unified approach to validity. Deygers (2017) notes the challenging of “theoretically disentangling” (p. 147) justice from fairness in validation research. The findings of the present study suggest that a complete disentangling is not possible. Aspects such as high stakes, variable content, limited timing and so on interact dynamically with the emotional experience of the test-taker which can impact overall performance, strategies used during the test, and attitudes about the test. The history of validation research may have focused primarily on topics that could be classified under fairness, but to ignore questions of justice means ignoring factors that affect test fairness and vice-versa.

Support for the value of test-taker expertise. The present study supports claims like those made by Cheng and DeLuca (2011) that “test-takers can bring a critical perspective to validation inquiries” (p. 117). Test-takers have experiences that no other stakeholder can claim. That is, the participants of this study know what it *feels like* to take the test with real consequences at stake. This fact emerged as an important part of the stories the participants chose to tell. What these stories show the test validation researcher is that what taking the test feels like is not only relevant to questions of justice, but fairness as well.

Support for the continuing democratization of language testing. When calling for the democratization of language testing, Shohamy (2001b) wrote that this was “not a call for the

abolition of tests, but a call for practising democratic testing that requires shared authority, collaboration, the involvement of different stake holders (test-takers included) and the monitoring of undemocratic uses of tests" (p. 389). The present study suggests that this is a sentiment also shared by many test-takers.

5.3.2 Methodological Implications

For the foreseeable future, IELTS will continue to be used as part of post-secondary admission, immigration, and international credential recognition in Canada. As such, ongoing research into all aspects of the test's validity for these purposes remains necessary, including collecting and integrating test-taker perspectives of the fairness and justice of the test. Qualitative research must be part of this process as it allows for a richness of description and is a way of including test-takers own words into academic discussions. Additionally, the present study adds at least a few more voices to this discussion while strengthening and amplifying a number of the themes independently identified by earlier researchers into IELTS such as Cheng and DeLuca (2011), Hamid and Hoang (2018), Hoang (2019), and Liu and Hamid (2024).

5.3.3 Practical Implications

For test owners and developers ongoing validation research remains critical in order to maintain the *face validity* required to continue to operate. Based on the findings of this study educating test-taker anxiety should be a high priority for the test developers. This would address issues of construct-irrelevant variance and perceived negative impacts of the test. One way to address this could be taking steps to lower the stakes of the test. An example of a measure already taken by the IELTS owners was the introduction of One Skill Retakes in 2023 (British Council, 2023). Allowing test-takers greater access to completing parts of the test individually could address of number of the anxiety-related issues mentioned by the participants in this study.

In general, allowing participant to take parts of the test individually can reduce the pressure to achieve a required score on all four parts of the test in a single sitting. Beyond the pressure of needing to do well in four separate modalities in the same sitting, allowing one-skill testing could address the financial burden of needing to take the entire test multiple times.

For test users, such as post-secondary institutions and immigration authorities, the present study supports ongoing consideration of the ways IELTS is used and interpreted as a high-stakes gatekeeping assessment. The results of this study do not suggest that test-takers overwhelmingly favour removing English assessment generally or IELTS specifically from either the post-secondary admission or immigration process. What the findings of this study *do* suggest is that IELTS may not be a universally fair representation with which to generalize about the English language abilities of all applicants for university admission or immigration. Considering factors such as test-taker background, level of test anxiety, and intended profession or academic program reveals that universal judgements of the fairness of IELTS for a given context remain problematic and should be accounted for.

The major practical implication of this study for test users is the importance of accepting multiple forms of English language proficiency proof. Currently, most major Canadian universities accept at least two standardized English tests for academic admission. Likewise, the Government of Canada accepts at least two standardized English tests as part of the Permanent Residency (PR) and Canadian citizenship application process. This is a step in the right direction, allowing test-takers to select a test based on their personal preferences regarding format, mode of delivery, and so on. However, the findings of this study suggest that test-users should increasingly consider non-test alternatives when possible.

There is a strong desire among participants for an opportunity to prove their English-language abilities that is more naturalistic, less time-restricted, and less prone to the variability caused by stress and the “luck” of whether one receives a topic which they are comfortable talking about. There is an argument to be made that bridging programs currently offered in some post-secondary institutions and coursework options for citizenship applications offer a solution to some of the issues mentioned by participants in this study. This study suggests that course-based options should continue to be promoted. However, such programs often represent a major commitment in terms of both time and money, resulting in the “IELTS-out” phenomena described by Johnson and Tweedie (2021). Additionally, even when the tuition for a course-based alternative is funded by the government, as with the Language Instruction for Newcomers to Canada (LINC) program, access to such programs is contingent both on the prospective student having the time away from other commitments to take the course and on the availability of the course itself as demand for space in the LINC program can outpace availability of such programs.

For some applicants, a standardized test may be the most appropriate means of demonstrating English-language abilities. For others, a course-based approach will be the better option. For still others, a middle ground between the two options should be explored. This could take the form of prepared presentations, less formal interviews, or exploring other methods that would allow test-takers more choice in order to better demonstrate their English-language abilities.

For test-takers, there are several implications. Managing test-taker anxiety should be a priority. Cheng and DeLuca’s (2011) assertion that greater preparation for the test and use of test-taking strategies may correlate with reduced feelings of anxiety may provide a general way

forward for test-takers who are particularly concerned about managing anxiety. While individuals' experience with test-related anxiety will vary, there is a considerable body of work related specifically to anxiety-reduction strategies during the IELTS speaking test (see Bai & Xian, 2024).

The perceived impact that test content has on performance should also be a consideration for test-takers. Test preparation could increase its focus on transferability regardless of subject matter (for example vocabulary items with a wide range of applications). Increased strategy development and practice preparation for situations where the topic is not an area where the test-takers are knowledgeable could help mitigate both the impact of receiving a “random” topic as well as potentially mitigating the anxiety associated with topic variability.

While the current testing policies are in place, test-takers and their teachers should also explore all the testing options that are available to them. While the test one chooses to take may be limited due to geographic location, there is an increasing trend towards online testing that could offer test-takers more choice in the interim. Finding the format with which the test-taker is most comfortable may mitigate some of the concerns raised by the participants in this study.

5.3.4 Reflective Implications

This study was conducted with a high level of collaboration during both the development of the data collection instruments and during data collection itself. During most of this process, I continued to work at various times as an English as an Additional Language instructor and as a sessional instructor teaching undergraduate reading and writing courses. The experience of doing this work while also collaborating with colleagues, speaking with participants, and coding and analyzing their experiences has had a profound impact on my practice as a teacher. In particular,

it has challenged me to become considerably more reflective about the kinds of assessments I give my own students.

As a work of critical applied linguistics, this study suggests a great deal of promise in smaller-scale, collaborative research, including practitioner research. The focus on much of this study has been on the largest standardized English test as it is used by large organizations like governments and universities. It is hoped that this study, when considered alongside other studies will contribute to broader discussions about how to make large-scale standardized English-language testing as fair and just as possible. However, an unanticipated byproduct of completing this study encouraged me, the practitioner-researcher, to strive to democratise my own use of tests.

5.4 Limitations

As a predominantly qualitative study with a small number of participants, the aim of this study was to describe and interpret participant opinions and experiences with a level of richness. This study can contribute to the ongoing discussions surrounding the fairness and justice of IELTS. While the participants in this study are diverse in terms of their countries of origin and fields of study, they represent a small sample of the wider test-taker population. IELTS is taken for post-secondary admission, permanent resident and citizenship applicants, and internationally educated professionals. However, the seven participants in this study were overwhelmingly academic test-takers and none reported taking the test for professional credential recognition.

Additionally, all seven of the participants had, at the time of their participation, already successfully achieved at least one of their reasons for taking a standardized English test. While these are voices that it is critical to hear from, this study did not give voice to any participant

who was struggling to achieve a required score, or indeed any participant who had tried and given up. It should also be noted that, because this was not a longitudinal study, the picture presented in this study can only truly represent the participants views and attitudes at a single moment in time.

5.5 Recommendations for Future Research

As noted earlier, the present study uses data collected as part of a larger study from which will include a larger and more diverse group of participants. However, there remains a need for continued research with more test-takers, representing a wide range of the test-taking population, and in as many test-taking contexts as possible. In particular, there is a need for continued work with IELTS test-takers residing outside of English-speaking countries as well as a need for more opportunities for these test-takers to express themselves in languages other than English.

Relatedly, there is also a need for more longitudinal work, as one's relationship with IELTS is often an extended process rather than a single sitting for the test. Longitudinal work would allow for comparisons between attitudes towards the test while a test-taker has not yet achieved their desired score and after they've done so (or after they've given up on trying to achieve said score).

Lastly, while test-takers are arguably the IELTS stakeholders most affected by the test on a personal level, other stakeholders are better positioned to enact the recommendations mentioned earlier in this chapter. In particular, policy-makers in post-secondary admissions, professional organizations, and government organizations responsible for immigration are critical to ensuring test validation research results in better language testing policies. There is a

need then not only to engage with policy-makers but also to include them in future research as stakeholders in their own right.

5.6 Conclusion

The present study sought to respond to the call for a more democratic and humane discussion surrounding standardized testing. Of course, no individual study can claim to settle so large an issue. The complexity of the use of standardized, high-stakes English testing in Canada requires continuous, critical study and discussion. To be done thoroughly, rigorously, and ethically, the democratization of language testing requires multiple perspectives and approaches. The present study has sought to contribute to the ongoing discussion by focussing on the experiences, opinions, and expertise of seven IELTS test-takers. I have situated my presentation and analysis of these test-takers' perspectives in the broader study of test validity.

Questionnaire and interview data reveal an appraisal of the fairness and justice of IELTS that is complicated, but also worthy of serious consideration. The question of fairness revealed an appraisal of the IELTS speaking test that is problematic but not without redeeming qualities. The role of test-taker emotions on performance, particularly nervousness, was identified as a major challenge to the fairness of the IELTS speaking test.

The insights gleaned from this study represent an attempt to amplify a small number of voices to a larger conversation about the fairness and justice of the use of IELTS for academic admission and immigration in Canada. In turn, this conversation is one part of an even larger one around the fairness and justice of testing more generally. It touches on the epistemological questions of how powerful institutions make claims to knowledge through the tool of testing. It

also touches on the ethical questions of how these powerful institutions use tests to make decisions affecting the lives of often marginalized individuals. More than two decades before the time of writing, Shohamy (2001b) claimed that “test takers have no say about the content of tests and about the decisions made based on their results” (p. 375). There is a case to be made that this statement remains true for IELTS test-takers. However, in focusing on what test-taker *do* have to say, and incorporating this into the wider study of validity and language testing, applied linguists, as well as researchers in other fields, can play a role in ensuring that test-takers’ voices are heard. In the spirit of test-taker centredness, it seems appropriate to give one participant from this study the last word:

I feel like this is such a, you know, it's ... an interesting project. I'm quite passionate about this, because I feel like the system itself is just, I don't know, there's a lot of things that need to be improved. And again, we're just milking people because they just know right? Like, well, these people are gonna pay, anyway, so let's just keep increasing the price tags. They want to stay in Canada anyway. So again, keep increasing a price tag. Let's just make them do this. So I hope, you know, a lot of people out there are going to be able to listen to the results of whatever it is that you guys are doing together now.... and keep going [laughs]. Thanks for having me. (P6)

References

- Ahern, S. (2009). "Like cars or breakfast cereal": IELTS and the trade in education and immigration. *TESOL in Context*, 19(1), 39–51.
<https://doi.org/10.3316/ielapa.065328178845270>
- Angoff, W. H. (1988). Validity: An evolving concept. In H. Wainer & H. I. Braun (Eds.) (Ed.), *Test validity* (pp. 19–32). Lawrence Erlbaum Associates, Inc.
- Ata, A. W. (2018). An investigation into the knowledge, education and attitudes of male and female international students in Australia to the International English Language Testing System (IELTS). In A. Ata, L. Tran, & I. Liyanage (Eds.), *Educational Reciprocity and Adaptivity* (1st edition, pp. 224–241). Routledge. <https://doi.org/10.4324/9781315201412>
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests (Vol. 1)*. Oxford University Press.
- Bachman L. F., & Palmer A.S. (2010). *Language assessment in practice (2nd ed.)*. Oxford University Press.
- Bai, Y., & Xian, H. (2024). Exploring the interplay of digital storytelling, L2 speaking skills, self-regulation, and anxiety in an IELTS preparation course. *Humanities and Social Sciences Communications*, 11(1). <https://doi.org/10.1057/s41599-024-04109-8>
- British Council. (n.d.). *IELTS: The most popular English proficiency test in Canada*. Retrieved March 30, 2025, from <https://www.ielts.ca/find-out-more/english-proficiency-tests-in-canada/>
- British Council. (2019). *Guide for education institutions, governments, professional bodies and commercial organisations*. <https://takeielts.britishcouncil.org/sites/default/files/ielts-guide-for-organisations.pdf>
- Brown, J. D. (2000). How does construct validity differ from other forms of validity? *Shiken: JALT Testing & Evaluation SIG Newsletter*, 4(2), 8–12.
<https://teval.jalt.org/test/PDF/Brown8.pdf>
- Canadian Bureau for International Education. (2023). *Canadian Bureau for International Education*. <https://cbie.ca/media/facts-and-figures/international-students/>

- Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1(1), 1–47.
<https://doi.org/10.1093/applin/I.1.1>
- Carlsen, C. (2007). The place of reliability within a unified (Messickian) view of construct validity. In C. Carlsen & E. Moe (Eds.), *A Human Touch to Language Testing* (1st edition).
<https://www.researchgate.net/publication/333728156>
- CELBAN. (2024). *The CELBAN Centre*. <https://www.celbancentre.ca/>
- Centre for Canadian Language Benchmarks. (n.d.). *language.ca*. Retrieved March 30, 2025, from <https://www.language.ca/home/>
- Cheng, L., & DeLuca, C. (2011). Voices from test-takers: Further evidence for language assessment validation and use. *Educational Assessment*, 16(2), 104–122.
<https://doi.org/10.1080/10627197.2011.584042>
- Collins, P. (2015, September 27). *English test derailing Irish dream of Australian citizenship*. *Irishtimes.Com*. <https://www.irishtimes.com/news/world/asia-pacific/english-test-derailing-irish-dream-of-australian-citizenship-1.2368668>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 54(4), 281–302. <https://doi.org/10.1037/h0040957>
- Deygers, B. (2017). Just testing. *ITL - International Journal of Applied Linguistics*, 168(2), 143–163. <https://doi.org/10.1075/itl.00001.dey>
- Deygers, B. (2019). Fairness and Social Justice in English Language Assessment. In X. Gao (Ed.), *Second Handbook of English Language Teaching* (pp. 1–29). Springer International Publishing. https://doi.org/doi.org/10.1007/978-3-030-02899-2_30
- Duolingo. (2023). *Duolingo English Test*. <https://englishtest.duolingo.com/applicants>
- ETS. (2024). *What is the TOEFL iBT test?* <https://www.etsglobal.org/ke/en/test-type-family/toefl-ibt-test>
- Foucault, M. (1977). *Discipline And Punish: The Birth Of The Prison* (A. Sheridan, Trans.). Allen Lane.
- Frost, K. (2017). *Test impact as dynamic process: Individual experiences of the English test requirements for permanent skilled migration in Australia* [PhD thesis, University of Melbourne]. <http://hdl.handle.net/11343/194277>
- Fulcher, G., & Davidson, F. (2009). Test architecture, test retrofit. *Language Testing*, 26(1), 123–144. <https://doi.org/10.1177/0265532208097339>

- Gribble, C., Blackmore, J., Morrissey, A.-M., & Capic, T. (2016). Investigating the use of IELTS in determining employment, migration and professional registration outcomes in healthcare and early childcare education in Australia. *IELTS Research Report Series*, 4. www.ielts.org/researchers
- Hamid, M. O. (2014). World Englishes in international proficiency tests. *World Englishes*, 33(2), 263–277. <https://doi.org/10.1111/weng.12084>
- Hamid, M. O., Hardy, I., & Reyes, V. (2019). Test-takers' perspectives on a global test of English: questions of fairness, justice and validity. *Language Testing in Asia*, 9(1). <https://doi.org/10.1186/s40468-019-0092-9>
- Hamid, M. O., & Hoang, N. T. H. (2018). Humanising Language Testing. *TESL-EJ (Berkeley, Calif.)*, 22(1), 1–27. <https://tesl-ej.org/wordpress/issues/volume22/ej85/ej85a5/>
- Hoang, N. T. H. (2019). Building a validity argument for the use of academic language tests for immigration purposes: Evidence from immigration-seeking test-takers. *Language Education and Assessment*, 2(3), 135–154. <https://doi.org/10.29140/lea.v2n3.148>
- Hoang, N. T. H., & Hamid, M. O. (2017). 'A fair go for all?' Australia's language-in-migration policy. *Discourse*, 38(6), 836–850. <https://doi.org/10.1080/01596306.2016.1199527>
- IELTS. (n.d.). *IELTS Speaking Band Scores*. Retrieved April 2, 2025, from <https://ielts.idp.com/canada/results/scores/speaking>
- Immigration Refugees and Citizenship Canada. (2025, February 21). *Apply for citizenship: Who can apply*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/canadian-citizenship/become-canadian-citizen/eligibility.html>
- Immigration Refugees and Citizenship Canada [IRCC]. (2025a, February 12). *Express Entry: Language test results*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada/express-entry/documents/language-test.html>
- Immigration Refugees and Citizenship Canada [IRCC]. (2025b, February 14). *Find out if you have the language proof for citizenship: Step 5*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/canadian-citizenship/become-canadian-citizen/eligibility/language-proof/step-5.html>
- Immigration Refugees and Citizenship Canada [IRCC]. (2025c, February 19). *Express Entry: Rounds of invitations*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada/express-entry/rounds-invitations.html>
- Immigration Refugees and Citizenship Canada [IRCC]. (2025d, February 21). *Apply for citizenship: Who can apply*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/canadian-citizenship/become-canadian-citizen/eligibility.html>

- Immigration Refugees and Citizenship Canada [IRCC]. (2025e, March 10). *Work in Canada after you graduate: How to find your language level based on your test results*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/study-canada/work/after-graduation/eligibility/language-results.html>
- Immigration Refugees and Citizenship Canada [IRCC]. (2025f, March 25). *Express Entry: Comprehensive Ranking System (CRS) criteria*. <https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada/express-entry/check-score/crs-criteria.html>
- Jennings, M., Fox, J., Graves, B., & Shohamy, E. (1999). The test-takers' choice: an investigation of the effect of topic on language-test performance. *Language Testing*, 16(4), 426–456. <https://doi.org/10.1191/026553299667813224>
- Jin, Y. (2023). Test-taker insights for language assessment policies and practices. In *Language Testing* (Vol. 40, Issue 1, pp. 193–203). SAGE Publications Ltd. <https://doi.org/10.1177/02655322221117136>
- Johnson, R. C., & Tweedie, M. G. (2021). “IELTS-out/TOEFL-out”: Is the End of General English for Academic Purposes Near? Tertiary Student Achievement Across Standardized Tests and General EAP. *Interchange*, 52(1), 101–113. <https://doi.org/10.1007/s10780-021-09416-6>
- Kane, M. (2013). The argument-based approach to validation. *School Psychology Review*, 42(4), 448–457. <https://doi.org/10.1080/02796015.2013.12087465>
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319–342. <https://doi.org/10.1111/j.1745-3984.2001.tb01130.x>
- Kubota, R. (2016). Studying up, down, or across? In J. McKinley & H. Rose (Eds.), *Doing Research in Applied Linguistics* (pp. 17–26). Routledge. <https://doi.org/10.4324/9781315389608-2>
- Kunnan, A. J. (2004). Test Fairness. In M. Milanovic & C. Weir (Eds.), *European language testing in a global context* (pp. 27–48). Cambridge University Press.
- Liu, X., & Hamid, M. O. (2024). “English is important, but perfect English is not”: The disjuncture between the IELTS and language experiences from the perspectives of international students in Australia. *International Journal of Applied Linguistics (United Kingdom)*, 35(1), 1–20. <https://doi.org/10.1111/ijal.12633>
- MacDonald, J. J. (2019). Sitting at 6.5: Problematizing IELTS and Admissions to Canadian Universities. *TESL Canada Journal*, 36(1), 160–171. <https://doi.org/10.18806/tesl.v36i1.1308>

- Manias, E., & McNamara, T. (2016). Standard setting in specific-purpose language testing: What can a qualitative study add? *Language Testing*, 33(2), 235–249. <https://doi.org/10.1177/0265532215608411>
- McNamara, & Roever. (2006). Validity and the Social Dimension of Language Testing. *Language Learning*, 56(S2), 9–42. <https://doi.org/10.1111/j.1467-9922.2006.00379.x>
- McNamara, T. (2000). *Language testing*. Oxford University Press.
- McNamara, T. (2013). Values in Language Assessment. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd.
- McNamara, T., & Ryan, K. (2011). Fairness versus justice in language testing: The place of English literacy in the Australian citizenship test. In *Language Assessment Quarterly* (Vol. 8, Issue 2, pp. 161–178). <https://doi.org/10.1080/15434303.2011.565438>
- Merrifield, G. (2016). An impact study into the use of IELTS by professional associations in the United Kingdom, Canada, Australia and New Zealand, 2014 to 2015. *IELTS Research Report Series*, 7. <https://ielts.org/researchers/our-research/research-reports/an-impact-study-into-the-use-of-ielts-by-professional-associations-in-the-united-kingdom-canada-australia-and-new-zealand-2014-to-2015>
- Merrylees, B. (2003). *International English Language Testing System (IELTS) Research Reports 2003: Volume 4*. IELTS Partners. <https://ielts.org/researchers/our-research/research-reports/an-impact-study-of-two-ielts-user-groups-candidates-who-sit-the-test-for-immigration-purposes-and-candidates-who-sit-the-test-for-secondary-education-purposes>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Macmillan Publishing Co, Inc; American Council on Education.
- Messick, S. (1992). *Validity of Test Interpretation and Use*.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Messick S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5–12. <https://doi.org/10.3102/0013189X023002005>
- Murray, J. C., Cross, J. L., & Cruickshank, K. (2014). Stakeholder perceptions of IELTS as a gateway to the professional workplace: The case of employers of overseas trained teachers *IELTS Research Reports Online Series*. <https://ielts.org/researchers/our-research/research->

[reports/stakeholder-perceptions-of-ielts-as-a-gateway-to-the-professional-workplace-the-case-of-employers-of-overseas-trained-teachers](#)

Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. *Canadian Journal of Education*, 25, 262–279. <https://doi.org/10.2307/1585850>

OET. (2025). *oet.com*. <https://oet.com/>

O’Neill, T. R., Buckendahl, C. W., Plake, B. S., & Taylor, L. (2007). Recommending a Nursing-Specific Passing Standard for the IELTS Examination. *Language Assessment Quarterly*, 4(4), 295–317. <https://doi.org/10.1080/15434300701533562>

Paragon Testing Enterprises. (2025a). *cael.ca*. <https://www.cael.ca/>

Paragon Testing Enterprises. (2025b). *CELPiP: Home*. <https://www.celpip.ca/>

Pearson PLC. (n.d.). *PTE English Language Tests*. Retrieved March 30, 2025, from <https://www.pearsonpte.com/>

Pretorius, L. (2024). Demystifying Research Paradigms: Navigating Ontology, Epistemology, and Axiology in Research. *Qualitative Report*, 29(10), COV1-2715. <https://doi.org/10.46743/2160-3715/2024.7632>

Rajendram, S., Larson, E. J., & Sinclair, J. (2019). International Graduate Students’ Perspectives on High-Stakes English Tests and the Language Demands of Higher Education. *Language and Literacy*, 21(4), 68–92. <https://doi.org/doi.org/10.20360/langandlit29428>

Read, J. (2022). Test Review: The International English Language Testing System (IELTS). In *Language Testing*. SAGE Publications Ltd. <https://doi.org/10.1177/02655322221086211>

Read, J., & Wette, R. (2009). Achieving English proficiency for professional registration: The experience of overseas-qualified health professional in the New Zealand context. *IELTS Research Report Series*, 10, 1–42. <https://ielts.org/researchers/our-research/research-reports/achieving-english-proficiency-for-professional-registration-the-experience-of-overseas-qualified-health-professionals-in-the-new-zealand-context>

Rose, H., McKinley, J., & Briggs Baffoe-Djan, J. (2020). *Data collection research methods in applied linguistics*. Bloomsbury Academic.

Rumsey, M., Thiessen, J., Buchan, J., & Daly, J. (2016). The consequences of English language testing for international health professionals and students: An Australian case study. *International Journal of Nursing Studies*, 54, 95–103. <https://doi.org/10.1016/j.ijnurstu.2015.06.001>

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). SAGE.

- Sedgwick, C., & Garner, M. (2017). How appropriate are the English language test requirements for non-UK-trained nurses? A qualitative study of spoken communication in UK hospitals. *International Journal of Nursing Studies*, 71, 50–59. <https://doi.org/10.1016/j.ijnurstu.2017.03.002>
- Shohamy, E. (1998). Critical language testing and beyond. *Studies in Educational Evaluation*, 24(4), 331–345. [https://doi.org/10.1016/S0191-491X\(98\)00020-0](https://doi.org/10.1016/S0191-491X(98)00020-0)
- Shohamy, E. (2001a). Democratic assessment as an alternative. *Language Testing*, 18(4), 373–391. <https://doi.org/10.1177/026553220101800404>
- Shohamy, E. (2001b). *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. Routledge. <https://doi.org/10.4324/9781003062318>
- Statistics Canada. (2017a). *Focus on Geography Series, 2016 Census*. <https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Index-eng.cfm>
- Statistics Canada. (2017b). *Immigration and Diversity: Population Projections for Canada and its Regions, 2011 to 2036*. <https://www150.statcan.gc.ca/n1/pub/91-551-x/91-551-x2017001-eng.htm>
- Tapper, J. (2017, September 24). *Nurses who failed English test aimed at curbing immigration set for a reprieve*. The Guardian. <https://www.theguardian.com/society/2017/sep/23/nurses-language-tests-immigration-nhs>
- The Michener Institute. (n.d.). *Michener English Language Assessment*. Retrieved March 30, 2025, from <https://michener.ca/continuing-education/ce-courses/mela/>
- Weir, C. J., & O’Sullivan, B. (2017). *Assessing English on the global stage: the British Council and English language testing, 1941-2016*. Equinox Publishing Ltd.
- Xi, X. (2010). How do we go about investigating test fairness? *Language Testing*, 27(2), 147–170. <https://doi.org/10.1177/0265532209349465>
- Yucel, M., & Iwashita, N. (2016). The IELTS roller coaster: Stories of hope, stress, success, and despair. In *Narrative Research in Practice: Stories from the Field* (pp. 209–223). Springer Singapore. https://doi.org/10.1007/978-981-10-1579-3_10

Appendix 1: Codebook

Includes name and description of top-level codes.

Child codes also include an example. To reduce ambiguity, some examples include additional context with bold lettering to indicate the coded text. When the example is not bolded, the entire example is included in the code.

<p>AFFECTED PERF - Expresses Belief That [Test] Performance or Score Was Affected by Factors Other Than English Language Ability.</p>	<p>The participant describes factors <i>OTHER</i> than English language skills that they believe may influence the quality of a test-taker's performance on IELTS or the score they received. This may be a real experience or a general opinion. The influence may be positive or negative or unknown.</p>			
Name	Description	Example	Files	Refs
<p>Format: Ability to Refer Back to Earlier Parts of the Test (positive) (referring to mock test)</p>	<p>Referring to the mock test, the participant reports that the ability to refer back to earlier parts of the IELTS speaking test may positively influence the quality of a test-taker's performance on IELTS or the score they received.</p>	<p>I was able to reflect back on the earlier parts and added some on my own. P2</p>	<p>P2</p>	<p>2</p>
<p>Format: Inability to Refer Back to Earlier Parts of the Speaking Test (negative) (referring to real testing experience)</p>	<p>Referring to a real testing experience, the participant reports that the inability to refer back to earlier parts of the IELTS speaking test may negatively influence the quality of a test-taker's performance on IELTS or the score they received.</p>	<p>Interviewer: did you feel the same [about the positive impact of being able to refer back to your earlier answers] in 2018 [when you took a real IELTS test]? Or?</p> <p>Participant: I think it was not like, like a progressive also, like, there were like different dates. So the speaking is done on like, a different date in my country. And all the tests, like listening, reading and writing, they were given like the same dates. So, it was the kind of like a long, like, three hour- I can't remember the exact duration, but it was like a three hour or four hours, and like a 30-minute test for the speaking. So, and the dates are in like, week and a half, like, gap in between. Yeah, so I won't say like they're like, related. Yeah, it's, like, different. P2</p>	<p>P2</p>	<p>1</p>
<p>Format: Ability to Make Notes in Part 2 (positive)</p>	<p>The participant reports that the ability to make notes (i.e., in part 2 of the speaking test) may positively influence the quality of a test-taker's performance on IELTS or the score they received.</p>	<p>Yeah, the Q&A discussion [in part 3] was like the amount of notes I made, and I was able to follow like the conversations from the notes and added on my- some notes on my own at like, the same time. P2</p>	<p>P2</p>	<p>1</p>
<p>Format: Ability to Prepare Answer in Part 2 (positive)</p>	<p>The participant reports that the ability to prepare their answer in part 2 of the IELTS may positively influence the quality of a test-taker's performance on IELTS or the score they received.</p>	<p>Yeah, I would say Part 2 because I had time to prepare for it. Yeah, I had time to prepare for it and think about it, which is why I was able to, uh, present- like express it in a more, like, less convoluted way than I would in parts 1 and 3 when it was just, uh, putting me like on the spot. P7</p>	<p>P7</p>	<p>1</p>
<p>Format:</p>	<p>Referring to a real testing experience, the participant reports</p>	<p>I just felt like sometimes it was just awkward, like oh, man, like, I just have to give you like, maybe</p>	<p>P6</p>	<p>1</p>

Ability to Speak at Length in Part 2 (positive) (referring to real testing experience)	that the ability to speak for a longer period of time in part 2 (compared to parts 1 and 3) of the IELTS speaking test may positively influence the quality of a test-taker's performance on IELTS or the score they received.	one sentence, and then that's a full stop, you know. Whereas with the second part it was just really, you know, it was, I could flow better. P6		
Format: Ability to Use Non-Verbal Communication (positive)	The participant reports that the ability to use non-verbal communication (e.g. body language, facial expressions, etc) may positively influence the quality of a test-taker's performance on IELTS or the score they received.	And also like, speaking is like an examiner can like judge based on your like expressions or like, how do you feel or like, the point you are trying to convey like using your hand gest- uh hand gestures or expression or your body contact. P2	P2	1
Format: Being Interrupted by the Examiner (negative)	The participant reports that being interrupted by the speaking test examiner may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I was going on this train of thought and then he cut me off and then I had to- and then I was asked a different question and then I had to restart my thought process again [laughs]. P7	P6, P7	3
Format: Difficulty Speaking as Long as Required by the Test (negative) (referring to real testing experience)	Referring to a real testing experience, the participant reports that they found the length of time they were required to speak during the IELTS speaking test may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I mean, it [i.e., the first time I took the speaking test] was like definitely like longer, I would say in terms of like speaking or the amount of things I have to say for that. I can't remember like how many questions I was asked, but for each answer I have to like speak for a long time. And then like when I have to, like, speak in long time, because now I can like I can like talk too much and before if like me, I was like shy back then like 3-4 years ago. It- it is difficult for me even though like I wanted to speak and express myself, it is difficult for me to speak for so long. So, yeah, in a way like it, it was different, but it still accessed like the same kind of like skills in a shorter amount of time P2	P2	1
Format: Length of the IELTS Test in its Totality (negative) (referring to real testing experience)	The participant reports that the length of the IELTS test in its totality may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I would say probably fatigue because we had to wait for a long time. There was a lot of us packed in one room and obviously we have to take turns, so we could be there for like 4 or 5 hours. And by the time it's your turn, you're just tired and you don't want to be there, yeah. P3	P3, P6	2
Environmental Factors (negative) (referring to real testing experience)	Referring to a real testing experience, the participant reports that factors in the environment (i.e., the testing centre, room in which the test was taken, etc.) may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	Probably just environmental factors too, cause for me, I did mine back home in Malaysia and the way they set it up, it's like they just have you in a room to wait and you have to walk down this long hallway and they make you just sit outside the room and wait for a turn. And it just kind of feels like, I don't know like maybe waiting to enter like a principal's office or something. It just seems like they just made it so intimidating, might even go in, like the room is dark and it's just one examiner there. It's just not a very comfortable experience being.... P3	P3	1
Examiner Accent (negative)	The participant reports that the accent of the speaking test examiner may negatively influence the quality	So I got in a panic mode a little bit especially during the speaking test because the person who was, interviewing me I believe he was from New	P6	1

	of a test-taker's performance on IELTS or the score they received.	Zealand, and then the accent was so strong. So I gotta ask him to like can you please repeat that again? Can you please rephrase that? Because I don't understand, you know, like he would pronounce the word pen like PEN [pronounced 'peen'] like, how do you even understand pen [pronounced 'peen']? Like pen [pronounced 'peen'] like pins, you know? So it was very different compared to what I just did back in July, it was way easier. P6		
Exposure to English in an Informal Setting (positive)	The participant reports that their earlier exposure to English in an informal setting may positively influence the quality of a test-taker's performance on IELTS or the score they received.	So, things like that [i.e., being surrounded by English, in an English-speaking environment] would create like a major difference in bringing your scale up and in like listening to like English music or, yeah, just like practicing it on a like a daily basis, like reading books, things like that would help. P2	P2	2
Feeling Anxious, Nervous, Stressed, Intimidated, or Under Pressure (negative)	The participant reports that feeling anxious, nervous, stressed, intimidated or under pressure may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	test settings in general are a little bit intimidating for everyone. So, that also might hinder your ability to perform to the best. P7	P1, P2, P3, P5, P6, P7	16
Feeling Anxious, Nervous, Stressed, Intimidated, or Under Pressure (negative) (referring to real testing experience)	Referring to a real testing experience, the participant reports that feeling anxious, nervous, stressed, intimidated or under pressure may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	Probably just environmental factors too, cause for me, I did mine back home in Malaysia and the way they set it up, it's like they just have you in a room to wait and you have to walk down this long hallway and they make you just sit outside the room and wait for a turn. And it just kind of feels like, I don't know like maybe waiting to enter like a principal's office or something. It just seems like they just made it so intimidating, might even go in, like the room is dark and it's just one examiner there. It's just not a very comfortable experience being... P9	P2, P3, P5, P7	5
Feeling Anxious, Nervous, Stressed, Intimidated, or Under Pressure (negative) (referring to mock test)	Referring to the mock speaking test, the participant reports that feeling anxious, nervous, stressed, intimidated or under pressure may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	Um, I'd say probably part 1 cause I think I have the hardest time with part one in the mock speaking test. Maybe because it was like at the start of the test, and I was still kind of getting used to things, and I get quite nervous when I'm being asked questions, like, in a public setting. So yeah, I kind of... hesitated a lot and, uh, maybe I pronounced things wrong? P7	P7	1
Feeling Anxious, Nervous, Stressed, Intimidated, or Under Pressure (negative) (comparing mock test to real test)	Comparing the mock speaking test to a real testing experience, the participant reports that feeling anxious, nervous, stressed, intimidated or under pressure may negatively influence the quality of a test-taker's performance on IELTS or the score they received on the real test.	Oh yeah, for sure. I'll be very scary, I think, [laughs] if I were in a real test. I'll be very, like, "Oh my God" I need to say then, the format that they want, using the vocabulary that way from me. Yes, I was more relaxed. Which is sad, I think. P1	P1, P2, P3, P7	6
Feeling Uncomfortable (negative)	The participant reports that feeling uncomfortable may negatively influence the quality of a test-taker's	Because there is some luck on it. If you don't get the right question or you don't like you don't feel comfortable, sometimes even for me, if if, like with	P4	2

	performance on IELTS or the score they received.	with the same level with my English, that won't change, P4		
Feeling Comfortable (positive) (referring to mock test)	The participant reports that feeling comfortable may positively influence the quality of a test-taker's performance on IELTS or the score they received.	I was comfortable so like, I got the chance to to express my grammar skills. P4	P4	3
Feeling Confident (positive)	The participant reports that feeling confident may positively influence the quality of a test-taker's performance on IELTS or the score they received.	if people that enjoy public speaking, they might find themselves to be super confident during the speaking session , whereas for people that are not as confident that might hinder their overall score in the speaking test as well despite how fluent they are. P7	P2, P7	3
Feeling Confident (positive) (comparing mock test to real test)	Comparing the mock speaking test to a real test, the participant reports that feeling more confident on the mock may positively influence the quality of a test-taker's performance on IELTS or the score they received as compared to a real test.	Yeah, I think it was definitely helped me to go with much confidence into the bigger conversation that we have in the next part and it definitely- it's kind of like an icebreaker question, so getting to know-familiar with like the examiner and be, like, friendly and if I'm feeling like less confident, yeah, it has like more prepared, but when I was like in real test in 2018, I did not feel the same kind of, like, uh confidence or carrying forward like icebreaker like conversation P2	P2	1
Lacking in Confidence (negative)	The participant reports that lacking in confidence may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	if people that enjoy public speaking, they might find themselves to be super confident during the speaking session, whereas for people that are not as confident that might hinder their overall score in the speaking test as well despite how fluent they are. P7	P7	1
Feeling Confused (negative) (referring to a real testing experience)	Referring to a real testing experience, the participant reports that feeling confused may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I don't know. At that time I was just confused and you know, kind of I was, yeah, I was confused just because because of the marks we we I was and also because, yeah, I don't know I have to I have to just answer the questions I know. I just came in this thing. P5	P5	3
Feeling Distracted or Unfocussed (negative)	The participant reports that feeling distracted may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	So let's say if I was in a room full of people, and then they were like.... they really will try to focus on what's being said, and then my belly starts making that sound, it's so distracting, you know? So there's like a lot of.... I feel like external factors that also contribute to that [one's performance during the test]. P6	P6	1
Feeling Distracted or Unfocussed (negative) (referring to mock test)	Referring to the mock speaking test, the participant reports that feeling distracted may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I think in the mock speaking test it was harder for me somehow, um, maybe because it was through, like, a virtual setting that I got very distracted by other things happening around me. So, it was harder for me to focus. And also, the questions that were asked... It wasn't like- it required me to, like, think about things and not- I can't just say them, like, off the top of my head, if that makes sense. And so, because I have to focus on what I'm thinking, I kind of just... like get a little bit,	P7	1

		like, frazzled in my brain, um, in terms of, like, using more vocabulary or pronouncing the right words. P7		
Feeling Distracted or Unfocussed (negative) (referring to real test)	Referring to a real testing experience, the participant reports that feeling distracted may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	And I was also thinking about my son because I have just babysit him to her to his grand-grandmother. So I was just kind of worried for him that he was, he must be crying because he was on my breastfeed that time. So that's why. That was all the factors that were just circulating in my mind and I wasn't able to to, you know, to focus. So I was just in a hurry, just go and come, yeah. P5	P5	1
Feeling Tired (negative) (referring to real testing experience)	Referring to a real testing experience, the participant reports that feeling tired may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	I would say probably fatigue because we had to wait for a long time. There was a lot of us packed in one room and obviously we have to take turns, so we could be there for like 4 or 5 hours. And by the time it's your turn, you're just tired and you don't want to be there, yeah. P3	P3, P6	2
Luck (negative / positive)	The participant reports that luck may negatively or positively influence the quality of a test-taker's performance on IELTS or the score they received.	Interviewer: [...] why do you think that the score is only SOMEWHAT accurate of an indicator of your English-speaking skills? Participant: Because there is some luck on it. If you don't get the right question or you don't like you don't feel comfortable, sometimes even for me, if if, like with with the same level with my English, that won't change, if <you ask me one question> I can feel comfortable, present like thousands of arguments or talk about something that like I like or I don't know (???). <And you can ask me questions that I don't find any interesting things to say so it will reflects on my on my marks later. P4	P4	1
Thinking About Parental Obligations (negative) (referring to real testing experience)	Referring to a real testing experience, the participant reports that thinking about their obligations as a parent before, after, or during the test may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	Interviewer: [...] on the survey you stated that you somewhat agree that factors other than your English-speaking skills may have affected your IELTS score. Could you share with us what what some of those factors might have been that impacted your score in 2014? Or even during your experience with the simulated test? Participant: [...] And I was also thinking about my son because I have just babysit him to her to his grand- grandmother. So I was just kind of worried for him that he was, he must be crying because he was on my breastfeed that time. So that's why. That was all the factors that were just circulating in my mind and I wasn't able to to, you know, to focus. So I was just in a hurry, just go and come, yeah. P5	P5	1
Format - Time Constraints (negative)	The participant reports that time constraints (i.e., the shortness of the speaking test, individual speaking test tasks, or limited time to speak or answer questions) may	Also, if you don't do well under a timed environment where there's, like, you have to produce something within this amount of time, that also affects your ability, I think P7	P1, P6, P7	6

	negatively influence the quality of a test-taker's performance on IELTS or the score they received.			
Format - Time Constraints (negative) (referring to mock test)	The participant reports that time constraints (i.e., the shortness of the speaking test, individual speaking test tasks, or limited time to speak or answer questions) may negatively influence the quality of a test-taker's performance on IELTS or the score they received.	Um, and I also would point out that because the... the parts were timed, and they were quite short, that there wasn't enough time for me to think and also speak at the same time. So, there was a bit of pressure there... Uh, yeah. P7	P7	2
Impact of Topic on Performance or Score (negative / positive)	The participant reports that the topic of the questions asked in the IELTS speaking test may negatively or positively influence the quality of a test-taker's performance on the test or the score they received.	I get- [exhales] the whole thing here why I feel sort of like it's not really the most reliable test ever, it's because.... the topic can be so random and sometimes it's something that you don't really like, what's that? I've never even heard of that term before. And you just end up like making things up just because, you know, like you don't know anything about <beers>, for instance, like black bears, whatever it is right? P6	P1, P3, P4, P5, P6	12
Impact of Topic on Performance or Score (negative / positive) (comparing mock test to real test)	Comparing the mock speaking test to a real test, the participant says that the topic of the questions asked in the IELTS speaking test may negatively or positively influence the quality of a test-taker's performance on the test or the score they received	I found the speaking topics for the stimulated test easier to answer [than the ones I was asked to talk about during the real IELTS test] because it's more relevant I think, because for the topics they asked me about university and I just graduated last year, so I can remember a lot of my university experience. I can bounce off on that P3	P3 P5	2
Impact of Topic on Performance or Score (negative / positive) (referring to mock test)	Referring to the mock speaking test, the participant says that the topic of the questions asked in the IELTS speaking test may negatively or positively influence the quality of a test-taker's performance on the test or the score they received	Yeah, I think it was very, like, I think the questions were reasonable, like, were, like, questions that I will face in my real life. So, as an international student all the time, people are asking me- ask me about Brazil. So, it's a kind of question that I answer all the time. Not really about my hometown, but, like, about Brazil and then, yeah, I think they were relevant, yeah. P1	P1	2

ALTERNATIVES - Proposes an Alternative Way of Proving English Abilities	<i>The participant proposes an alternative to IELTS for proving their English abilities for an academic, immigration, or work purpose. The participant may not use words like "instead of", but the described way of proving English is no longer clearly identifiable as an IELTS test.</i>			
Name	Description	Example	Files	Refs
Combine Multiple Skills	The participant proposes merging multiple skills.	And then because he [a hypothetical test-taker] will listen, and he will also comment, so it could be good, good opportunity like to test two or three or I don't know how many of uh.... skills or knowledge or? P4	P4	1
Consider Courses, Workshops, Certifications, or Volunteer Experience as an	The participant proposes a system in which courses, workshops, certifications, or volunteer experience could be considered as an alternative to a standardized test.	[Instead of being required to take a test, prospective immigrants or university applicants could demonstrate English abilities through] Courses or, like, getting	P2	1

Alternative to a Standardized Test.		some certification, or to have some, like, volunteer experiences in and like an organization, who is, like, working in Canada, and you volunteered for an organization. If they could give, like, uh help you- if you are like good in English, then you won't have to go through a test P2		
Consider Test-Taker Backgrounds	The participant proposes a system of proving one's English ability that takes test-taker background into greater consideration than the current does.	if I was a person who had sort of like the chance to recreate what's necessary and or what's not necessary for P.R. applications I would take a look at.... I think it should be unique for each each person, each applicant. P6	P6	1
Find Ways of Reducing Test-Taker Stress or Anxiety	The participant discusses changing current testing system by finding ways of reducing test-taker stress or anxiety.	So, [testing organizations should do] something that puts the interviewee at ease, so they won't have, you know, they won't panic. They won't have anxiety to make them forget that this- you have one shot to do this. Because if you just make them at ease likely, you know, conversation will flow easier. They won't freeze up, or panic, or have anxiety. A lot of people, not just people with anxiety, right. If you feel comfortable, obviously your speech will be smoother and ideas will flow faster. P3	P2, P3, P6	3
Give Prepared Presentation	The participant proposes including a prepared presentation as all or part of proving one's English ability	Also, maybe [ask test-takers to give] an oral presentation about a familiar topic. P1	P1, P3	2
Give Time to Prepare in Advance	The participant discusses changing current testing system by giving test-takers time to prepare in advance.	I think- that standard standardized test is needed for like university entrance exam, but I don't think it was an accurate representation, cause for me personally.... Personally, I tend to have anxiety when it comes to testing situations, especially speaking to people, like- it's like I have lots of ideas- my brain, but my mouth just cannot say it and after the exam, I'm always like, why didn't I say that? Oh, why did I say that? So, I feel like there's a lot more in here that I can't say on the spot. I just make things up on the spot. But if you give me some time to prepare, I can do a lot better for it, so I don't think it was a fair assessment, especially for like English as a second language. Like individuals like- for me, English is my first language, so probably it's a bit easier for me, but the anxiety is still there and always panic on the spot and just go to blank, so.... P3	P3	1
Give Topic in Advance	The participant discusses changing current testing system by giving test-takers a topic to prepare in advance.	For academic purposes, I mean they- [testing organizations] if they want to, they could give you a topic in advance P3	P3	1

Incorporate Video Stimulated Discussion in Speaking Test	The participant discusses changing current testing system by incorporating video stimulated discussion during speaking test.	maybe doing some comments on like kind of videos, short videos or clips, and doing those kind of discussions, it's more involved more- it's like more updated kind of tests... P4	P4	2
Use an Interview to Prove English Abilities	The participant proposes a system of proving one's English ability that makes use of an informal interview rather than a "speaking test" as included in IELTS.	So, I think I don't agree with the standardized test, but maybe an interview would be a better.... assessment then asking you to make up ideas on the spot, like they can just make it a less formal interview P3	P3	2
Let Test-Taker Choose How They are Assessed	The participant proposes a system of proving one's English ability that allows test-takers to choose how they are assessed.	I guess that's also hard to know how to assess that [i.e., English speaking skills] because I feel like everyone's ability to, like, express themselves is different, like for example, if people that enjoy public speaking, they might find themselves to be super confident during the speaking session, whereas for people that are not as confident that might hinder their overall score in the speaking test as well despite how fluent they are. I guess like- I guess maybe, like, a recording of yourself would be good- would be a good option, or if it's up to the examinee to decide, they could choose to write like an essay, or write a letter, or something like that. Yeah, I don't- I don't really know, but I feel like there are better ways, it's just- just I don't know yet. P7	P7	1
Let Test-Taker Choose Topic of Assessment	The participant proposes a system of proving one's English ability that allows test-takers to choose the topic of their assessment.	we can even choose the the field, the subject of the video to give candidate more.... not to be more comfortable, but also like to judge him with like high level because he choose this like.... this this field of this kind of videos, or this subject, so he must showcase some like good understanding. P4	P4	1
Let Test-Takers Provide a Recording	The participant proposes a system of proving one's English ability that allows test-takers to provide a recording.	I guess maybe, like, a recording of yourself would be good- would be a good option P7	P7	1
Observe Test-Takers in a more "Real World" Setting	The participant proposes a system of proving one's English ability in which the test-taker is observed performing tasks in a more natural or "real world" setting.	and maybe stage it as something where you go in the shop and you want to buy something or you need to, you know, talk to someone to discuss, but not make it so rigidly formal <and sensitive> P3	P3	1
Format Assessment as a Less Formal Conversation or Chat	The participant proposes a system of proving one's English ability that the participant characterizes as a less formal conversation or chat.	I think I would benefit more if there was like a program where I could be paired with like a native speaker, and just hang with them. Really, you know, like, just talk about things P6	P1, P3, P5, P6	6
Spread Out Assessment Over a Longer Period of Time	The participant proposes a system of proving one's English ability which spreads the period of assessment out over a longer period of time than the current system used by IELTS.	It might take a bit of a longer time, so let's say like, alright, so this week, when you speak, keep speaking to us, and the next week gonna do writing tests.	P6	1

		P6		
Use A Combination of Prepared and Unprepared Elements	The participant proposes a system of proving one's English ability which uses both prepared and unprepared elements as opposed to the current system used by IELTS, in which no pre-prepared elements are included.	It's also, even for IELTS, we can prepare well, we improve English, but we are also prepared, so always like to to check like to give different questions, for example, or new situations that were not prepared. It would be better to evaluate.... too. P4	P4	1

CHANGES 2 IELTS - Proposes Changes to IELTS Itself.	<i>The participant describes a change to the IELTS test itself.</i>			
Name	Description	Example	Files	Refs
Cost of Test	The participant proposes making the IELTS less expensive for test-takers.	if I could change that for that test, definitely, you know, shouldn't worth 300 dollars. It should be way cheaper than that, maybe like 50 dollars tops [laughs]. Yeah. P6	P6	1
Format Of Grades Given to Test-Takers	The participant proposes a change to IELTS in terms of how test results are given to test-takers.	I guess maybe the criteria itself should be assessed as well, like, instead of having a general overall score of that speaking portion, maybe it could be split into the categories that are being marked. So, like, grammar is however much, fluency is however much, and so on and so forth, I think that's probably more helpful. P7	P7	2

COMM SKILLS - Describes Communication Skills Required (or Not Required) for an Academic, Daily Life, or Workplace Setting.	<i>The participant describes a speaking skill (or other communication skill) that they believe is put to use in an English-speaking environment (i.e., in an academic setting, in a workplace setting, or in a daily life setting not directly associated with work or study). OR participant expresses the opinion that a speaking or other communication skill is NOT required for one of these settings.</i>			
Name	Description	Example	Files	Refs
EAP Speaking Skills Required	The participant describes a speaking skill that they believe is put to use in an academic environment such as a university class, working on an assignment, or communicating with teachers or classmates. OR	in my program, speaking is not like, major part because I am in computer science. It's like mostly like coding, or if you were like a good programmer, that's what you need, not like maybe- we do, like, have some like presentations and like, group work, but that comes in like the later part of, like, our undergrad. So yeah. P2	P1, P2, P3, P4, P6	10
EAP Speaking Skills NOT Required	The participant expresses the opinion that a speaking skill is NOT required for an academic setting.	I don't feel that we worked on we work on on speaking and communication so it's more about writing and it's not about presenting. P4	P4	1

Daily Life Speaking Skills NOT Required	The participant expresses an opinion that a speaking skill is NOT put to use in daily life (i.e. general daily needs not better described as studying at a university or working).	you need those skills for university degrees, that presentation skills. So, I think that would be very relevant for school, but for, like, immigration purposes, for, like, PR. Not so much. P3	P3	1
Daily Life Speaking Skills Required	The participant describes a speaking skill that they believe is put to use in daily life (i.e. general daily needs not better described as studying at a university or working).	I used to speak English with the with my, with my son's teacher whenever I come across with her, I mean asking her about my son's attitude and his behavior in the class and all that thing. P5	P5, P6	6
Workplace Speaking Skill NOT Required	The participant expresses the opinion that a speaking skill is NOT required for a working environment.	I don't really need to do anything like public speaking in front of huge crowds or anything like that. It's more like- it's a small office, so I really don't interact with that many people. P3	P3	1
Workplace Speaking Skills Required	The participant describes a speaking skill that they believe is put to use in working environment.	in my line of work I'm required to do scheduling and booking a lot and you need to let people know like, you know, you're booked in and alright, you have to cancel this with that. So, if you're not able to communicate with them properly, they'll just turn up on the spot when you're not booked in and it's gonna mess up the entire schedule. P3	P3, P4	4

ENG LVL - Discusses the English Abilities of Native English Speakers OR Appraises English Abilities or Changes in English Abilities Over Time	The participant discusses native English speakers' English language abilities. OR Participant describes their own English abilities. This could be describing a point in the past, currently, or comparing English at different times, or discussing changes in English abilities over time.			
Name	Definition	Example	Files	Refs
ENG LVL (LI ENG SPKRS) - Discusses the English Abilities of Native English Speakers	The participant discusses native English speakers' English language abilities.	Yeah, this is also very interesting about this question, about English speakers [needing to take IELTS or not] because I'm a TA, and then I was very shocked in a way that sometimes when I'm marketing papers- I was, Oh my God, there's some speaking-English speaking. Sorry, speaking English people who, I don't know, maybe has a worse English than me, like, but at the same time, like, I was very shocked because, like, I thought that everybody who speak English knows English very well, but at the same time, like, I think it's not the end of the world P1	P1, P3, P4, P6	5
ENG LVL (SELF) - Appraises	Participant describes their own English abilities. This could be describing a point in the past,	Interviewer: Do you feel that your speaking skills have changed since you took the IELTS in 2018?	P1, P2, P3,	16

English Abilities or Changes in English Abilities Over Time	currently, or comparing English at different times, or discussing changes in English abilities over time.	Participant: Yes, definitely, it has become like more confident and I have been able to, like, be more expressive in myself and in my opinions. P2	P5, P6, P7	
---	---	--	------------------	--

IMPACT - Describes the Impact of IELTS				
Name	Definition	Example	Files	Refs
<i>IELTS Emotional Impact.</i>				
<i>The participant discusses the impact that taking the test or having to take the test had on them. This could take the form of an emotional impact or a positive or negative consequence they experienced due to taking the test or having to take the test.</i>				
Anxiety, Pressure, Stress, Nervousness, or Intimidation (negative)	The participant describes a feeling of anxiety, pressure, stress, nervousness, or intimidation CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	I think it definitely creates a pressure on my mind when we have to, like, get some standardized tests , but definitely, like, less pressure I know it's like now graduated from school that prepares you well for to take like more difficult and difficult standardized tests. So, definitely, like, the pressure is like still but it's like less and I have like more confidence now than before. P2	P1, P2, P3, P5, P6, P7	11
Boredom (negative)	The participant describes a feeling of boredom CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	And it's also like an anxiety, and the stress, it's mentally draining sometimes when you have to study and repeat the same process for like a longer period of time and reduce like motivation and you don't feel like.. you belong somewhere and you feel like bored enough. P2	P2	1
Confidence (positive)	The participant describes a feeling of confidence CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	Yeah. And also it helps me to speak like confidently because I didn't used to speak in English in in Pakistan in my with with my friends or with my family. So when I spoke over there with another lady, so that helped me to how to I mean, I came to know about me that how much can I understand and how much can I speak or how much confident I am. So yeah, it helps me positively. P5	P2, P5	3
Demotivation (negative)	The participant describes a feeling of demotivation CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	And for the negative sides, sometimes you get like demotivated when you don't feel like you... are getting like a good score when you give like a mock test P2	P2	1
Not Belonging (negative)	The participant describes a feeling of "not belonging" CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	And it's also like an anxiety, and the stress, it's mentally draining sometimes when you have to study and repeat the same process for like a longer period of time and reduce like motivation and you don't feel like.. you belong somewhere and you feel like bored enough. P2	P2	1

Resiliency (positive)	The participant describes a feeling of resiliency CAUSED by taking the test or having to take the test BUT does not discuss these feelings as impacting how well they performed on the test or the score they received.	So it's really more about resilience in general rather than just kind of like English, English, English, English, and you have to be able to speak English perfectly before you're able to stay in Canada forever. You know something along those lines. Yeah. So it's a bit more like, a general resiliency thing than anything else. P6	P6	2
<i>IELTS Other Impacts</i>	<i>The participant describes a positive or negative consequence of taking the test or having to take IELTS. The participant may also discuss a positive or negative consequence they did NOT experience as a result of taking IELTS or having to take IELTS.</i>			
Causing a Delay to Study Plans (negative)	The participant says taking IELTS (or needing to take IELTS) caused a delay in their academic plans (for example, delaying their admission to a university or starting a university program).	I applied at UVic during like uh the early- like the decision date, was like past and I took my IELTS and it took them the IELTS like agency took them quite a while to like get back the results to me and send it back to the university. So, the university got the results, and it was pretty late. I think everyone was gone for like a summer break so, it delayed in like my admission process because I did not get my decision like back in time. So, my classes were started in September. Then I have to get a deferment because I have to apply for a student permit back to like to IELTS and Government of Canada to do that stuff in order to like get into Canada. So, that delayed me for like 1 semester and I could not like start my class at the end of September, and I was registered into classes, then I had to do deferment for a term. P2	P1, P2	5
Contributing to Participant to Consider NOT Pursuing a University Program (negative)	The participant says taking IELTS (or needing to take IELTS) has contributed to them considering NOT pursuing a university program	Uh, yeah, cause I'm also not a very big fan of like [laughs] doing this test again and again and I know if it would be like, uh, like uh computer science skills, they could like test me related to like my study cause it's like, why English? Because I have given like the test once and it won't change like my score. And I have met like a couple of other requirements in the language. I would say it is like a repetitive process. If I would go for a grad school, I won't give like another like ENGLISH standardized test per se, but if I have to give, like, other kind of- like a mix of other tests that I would pursue- I'm ready to give. But, if it is like purely like English, then I won't- I would not say I would go for like further masters or grad school. It is also like I won't say like I feel like too much excited when it comes to like speaking or like writing test because that's not what I'm good at cause I'm good at computers and I like to code and I like learning about like- also good at like in psychology and reading. But that's it. Not good with like English literature or like politics and history stuff. That's not like my main focus.	P1	1

		P1		
Helping Participant Learn or Practice English (positive)	The participant says taking IELTS (or needing to take IELTS) helped them learn (or practice or improve) their English language abilities.	Preparing for taking the test.... it doesn't affected me negatively, but yeah, it helps me positively. It helps me to to know many words, many vocabulary which I can use over here. So yeah. P5	P5, P7	2
Learning About a Different Culture (positive)	The participant says taking IELTS (or needing to take IELTS) provided an opportunity to learn about a different culture.	I think in a positive way, it's definitely like learning about the different culture and practicing like how I would fit into a different culture and to be like- make some decisions about my life with, uh, I would fit into my- into a different culture while learning about a different language and preparing for a test, and how like people study here and what are like their study techniques. P2	P2	1
NOT Causing a Delay to Immigration, Work, or Study Plans (positive)	The participant says taking IELTS (or needing to take IELTS) caused a delay in their NOT causing a delay to immigration, work, or study plans.	I did a lot of research before deciding to pursue studies in Canada, and I knew the timeline for the language test, and I knew how much time I had to get the test back, so I didn't get delayed. I did my test at least one year before I- like submitted any applications for university. So, I got my grades well beyond that, so my application just went on time and nothing was delayed P3	P2, P3, P6	5
Providing Information about English Level (positive)	The participant says taking IELTS (or needing to take IELTS) helped them by providing information about their English level.	Yeah, this test this test shows me that how much good I or bad I am in my English because I have to learn a lot from this survey and this test too, so I just come came to know now that to my English at what level I think my English is. P5	P5	2

MIN SCORE - Offers Opinion about Minimum IELTS Score Accepted by an Institution (i.e., the Government, Universities, etc.)	<i>The participant offers an opinion about the minimum score accepted by a government for immigration (for permanent residency, citizenship, etc) or by an academic institution (for university entrance etc.). This could be appraising the current scores accepted as adequate, or saying they should be higher or lower. Note that participants may not have up to date, accurate knowledge about the current scores accepted. Code based on what the participant THINKS the current score is.</i>			
Name	Definition	Example	Files	Refs
Immigration, Citizenship, or Permanent Residency (should be higher than it currently is)	The participant offers an opinion about the minimum score accepted by a government for immigration (for permanent residency, citizenship, etc).	with regard to citizenship, though I feel like maybe it shouldn't be as high [as the scores required for university entrance].... but you can't really just pass a person if they score like, for instance, 4 overall right? Because, like honestly, man like at the end of the day, Canada speaks English and French. If your English sucks like, how are you going to be able to function over here like, how are you going to be able to land a job? Like a decent	P5, P6	3

		job, right? A good job. So I think it depends on the person like if they feel like whatever I'm just gonna do it, if I score 4, I don't care. Well, you, I actually have to care about it, because, you know, it's going to determine what's what life's gonna look like for you over here. Yeah. P6		
Immigration, Citizenship, or Permanent Residency (should be lower than it currently is)	The participant offers an opinion about the minimum score accepted by a government for immigration (for permanent residency, citizenship, etc).	It's, yeah, more than six [for permanent residency]. We just mean marks more than six in the English IELTS test General, I guess. So it's six is so much because no one gets in easily if the score 6, so I think it must be like 4 so that people can easily have I mean, they then easily passed the test. P5	P1, P5	2
University Admission	The participant offers an opinion about the minimum score accepted by an academic institution (for university entrance etc.).	if we're talking like linguistics, for instance, or communication, or what else? Like public relations, maybe. I dunno politics, international relations, anything along those lines, I think it's kind of like.... automatically, you know, you're automatically thinking about, like, alright, well, the person who's doing any of these courses any of these subjects they need to speak a lot. So alright, the band, the standard, the benchmark, should be way higher. I don't know, maybe somewhere between 7 or 6.5 minimum. So, yeah, that would be that would be my answer. P6	P2, P6	4

MOTIVE - Gives Motive for Taking IELTS		<i>The motive for taking IELTS is the goal or hope for end result. For example, a participant may take a standardized English test in order to gain entry to a university in an English-speaking environment, gain Canadian citizenship, or as part of their permanent residency application.</i>		
Name	Definition	Example	Files	Refs
Immigration	The participant says that they took IELTS for immigration purposes (i.e., citizenship or permanent residency) This may or may not be the only motive (i.e., it can be one of several motives).	I yeah, I took the IELTS test before coming here because it was a demand for the immigration, we it was a requirement for the immigration document. So yeah, I took the IELTS test and it was general IELTS General. P5	P5, P6	2
Study Permit and Visa	The participant says that they took IELTS for a study permit or visa. This may or may not be the only motive (i.e., it can be one of several motives).	This was for my study permit and for my visitor visa to come over to Canada for my university degree. P3	P3	1
University Admission	The participant says that they took IELTS for admission to a university or other post-secondary program. This may or may not be the only motive (i.e., it can be one of several motives).	Yeah, so one of the major require- English requirements for for UVic. It was like, IELTS score. So, I fulfilled the requirements by sending them my IELTS score official. P2	P2, P3, P6, P7	5
Wanting to Assess Own English Skill	The participant describes wanting to assess their own English skills as a motive for taking the test. This may or may not be the only motive (i.e., it can be one of several motives).	I also just wanted to test my own English skills, P7	P7	1

MULTI - Discusses Taking IELTS Multiple Times (either an Experience OR an Opinion about the Necessity of Taking the Test Multiple Times)	Participant describes personal experience or experience of others taking IELTS test multiple times. Code opinions about the need to take the test multiple times separately.			
Name	Definition	Example	Files	Refs
Discusses Taking IELTS Multiple Times	The participant describes the experience of others taking IELTS multiple times.	So like, if a student don't do like- he's not like, getting it in like first or like second attempt, he has been trying IELTS like, multiple times- I have couple of my friends then back home, they tried IELTS multiple times, but did not- does not get a perfect score , it doesn't mean that they're not like good enough for them. They don't want to go into like college and like study, it's like, they don't get certain things in a different way P2	P1, P2, P3, P7	5
Expresses an Opinion about the Necessity of Taking the Test Multiple Times	Participant expresses an opinion about the necessity of taking the test multiple times.	if you're applying for PR and you're required to do an English speaking test, then when you apply for citizenship, you won't be- you shouldn't meet to do one, because it's probably within, like, two or three years that you're applying for PR and citizenship, so, and I don't think your English skills would, like, would go- would decrease it within that period of time. So yeah, I just don't think it's necessary to take English tests for both PR and citizenship, but definitely PR. For sure P7	P7	1

OPINION - Expresses an Opinion about IELTS	<i>The participant gives an opinion about IELTS or another standardized English test which is not otherwise captured by a different code.</i>			
Name	Definition	Example	Files	Refs
Ability to Demonstrate Language Abilities	The participant expresses an opinion about their ability to fully demonstrate their English language abilities while taking IELTS. This could be in general or more specifically about an aspect of their abilities such as grammar, vocabulary, fluency, pronunciation etc.	I guess that also demonstrated my fluency, especially in Part 2 when I was able to talk for more than like - for one to two minutes. I did kind of stutter at some point because I ran out of things to talk about. But yeah, but I think Part 2 was probably the best method for me to, um, for them to assess my fluency P9	P1, P2, P3, P4, P5, P6, P7	19
Ability to Express Oneself	The participant expresses an opinion about IELTS in terms of giving the ability to be themselves, or the ability to express themselves.	I think the second part, you know, where I had to speak for I mean two minutes was not really a long time, but I could flex a little bit. That would be something that I would say the most useful in terms of how one could actually, you know, like, alright, this is what I've trained for.	P1, P6	4

		This is what I've studied for. So this is time for me to shine right? That will be the best part of all three parts. The third part might come in second? You know? But for the most part, I truly enjoyed the second part, because I felt it was, you know, the part where I could be myself without being interrupted. P6		
Being a Barrier or Burden to International Students or Immigrants	The participant explicitly offers the opinion that the the test or the need to take the test is or isn't a barrier to international students.	I've seen many international students who could probably speak better than locals here too. So, why do they have to prove it just because they're born in a different country, even though their official language in that country is English? What makes them any different from people who were born here? I don't think it's fair. I think it's an additional barrier for international students. P3	P1, P3	2
Being Subjective	The participant describes IELTS as "subjective"	Because subjects are it's it's it's subjective. Quest- those questions, many questions are subjective so. P4	P4	1
Comparing IELTS to DuoLingo English Test	The participant compares IELTS to DuoLingo English Test. This is a comparison of the tests' qualities rather than given as a personal reason for choosing one test or the other.	Yes. Yeah. Like honestly, depending on on the field. And on the score too. As a test, yes, it's good. But then, if they feel like needs more writing, then there must be like something, not not standard for all for all programs. This need more writing then I don't know, maybe IELTS will be better. If it's more about communication then Duolingo is better. P4	P4	1
Cost	The participant gives an opinion about IELTS respect to the cost of taking the test.	So, I think it was just kind of like again, just the system that sort of like made people do that just because they knew they would pay. They would have to pay for that. And it wasn't cheap. It's still not, right? So, it's just like such an unfair approach to the whole thing here I feel like. P6	P3, P6	4
Easiness or Difficulty: IELTS is Easier than TOEFL	The participant expresses the opinion that IELTS is easier than TOEFL.	I find TOEFL is a bit difficult than IELTS. I find IELTS is pretty relaxed, then you have like, longer time span than TOEFL. In TOEFL you get like, the limited amount of time, and it's not like as much as spread out as IELTS is. P2	P2	1
Easiness or Difficulty: Preparing for IELTS is Difficult	The participant expresses the opinion that preparing for IELTS is difficult.	I was working. I don't have enough time.... And it's hard. I I I I didn't give a try, so I don't know. I'm not like going to say it's too hard for me or no, but I'm sure that it needs... long time, MUCH time to review, to prepare. So and then I don't, then I don't have enough time to do that. Okay, because things went quickly and shortly. I have time to prepare, like I don't have the six months or one year to prepare. I work, so I don't have time.	P6	1

		P6		
Easiness or Difficulty: Comparing Difficulty of Mock IELTS Speaking Test and Real IELTS Speaking Test	The participant compares the difficulty of the mock IELTS speaking test and a real IELTS speaking test in terms of easiness or difficulty	Yeah, the the difference [between the mock test and the real test I took in 2014] is in Pakistan, I was, I was not speaking English at that time even in, even with the teachers or with my son, because we were speaking native language over there. So when I came here so it's it's a different because I can easily understand you I can easily understand those persons who are speaking in English. So yeah, it's it's a difference because when I gave the IELTS test to go to to, to that, to lady, so when she was talking in English or she was asking me questions, it was very hard for me to understand her. So yeah, it's it's a difference. Way much better, like kind of [laughs] P5	P2, P3, P5	6
Fairness or Unfairness: Testing Requirement for Immigration (unfair)	The participant expresses the opinion that the testing requirement for immigration is unfair	Yes! Because, like, all these people who are, who immigrate to Canada, they, like, they have many factors. They are struggling through many factors, and I know they, like, they put effort, like, in the language, but they also- they're doing a lot of other things to the country, like they're pay their taxes, they are, like, working hard, sometimes in more than one job. So, like, it's not fair, I think. Yeah. And also I think thinking about, like, how these immigrants, they came here to work and, like, how much effort and money they put in this country, like, sometimes also I wonder, like, what time they will have time to study English, or, you know, so it's very hard. I think it's a lot to ask from them, they're already giving so much. You know? P1	P1	1
Fairness or Unfairness: Testing Requirement for Immigration (fair)	The participant expresses the opinion that the testing requirement for immigration is fair	Participant: Yeah. Would be so much better that way. [I.e., considering test-taker backgrounds rather than making all P.R. applicants prove their English ability with a standardized test.] Interviewer: Do you think there are cases where it is appropriate to give a test like that? Participant: Oh, yeah, absolutely. I feel like, so let's say, if you're like from, I don't know, let me just, well from Indonesia, where I'm from, right? And you did your degree in Indonesia, you have no experience living like in an English-speaking country. So basically, there's no like English really experience in your background, then, of course, it's gonna be case by case. But that person needs to test. I think that's fair. Yeah. P6	P6	1

<p>Fairness or Unfairness: IELTS as an Assessment of English Abilities (unfair)</p>	<p>The participant expresses the opinion that IELTS is unfair as an assessment of English abilities.</p>	<p>: I think- that standard standardized test is needed for like university entrance exam, but I don't think it was an accurate representation, cause for me personally.... Personally, I tend to have anxiety when it comes to testing situations, especially speaking to people, like- it's like I have lots of ideas- my brain, but my mouth just cannot say it and after the exam, I'm always like, why didn't I say that? Oh, why did I say that? So, I feel like there's a lot more in here that I can't say on the spot. I just make things up on the spot. But if you give me some time to prepare, I can do a lot better for it, so I don't think it was a fair assessment, especially for like English as a second language. Like individuals like- for me, English is my first language, so probably it's a bit easier for me, but the anxiety is still there and always panic on the spot and just go to blank, so.... P3</p>	<p>P3</p>	<p>2</p>
<p>Fairness or Unfairness: Comparing Fairness of the Mock IELTS Speaking Test and a Real IELTS Speaking Test</p>		<p>It's definitely- in a stimulated test, I'm in more in like a relaxed and kind of like- I'm free to like speak also like not in like a pressure while in like a real test. I would be like- pressurized and also, I am under pressure to like perform good and I wanted to get a good score so it was not like a thing in a simulated test whether I was given like actual reality check to just not to be pressurized and be just assessed on my truly based on my skills. Removing all that like stress, pressure and anxiety. So, it was kind of like a more fair- both are like fair in their own different ways but would feel the like the stimulated test- if they would remove like an anxiety or like a stress kind of, but it is also like a more of a personal choice. Some people feel like anxiety/stress when they go for a test. More people are like, just relax. So, it's on a personal level, but I would say that's the like the difference between those tests that could have impacted the score.</p>	<p>P2</p>	<p>1</p>
<p>Fairness or Unfairness: L1 English Speakers Needing Test (PR) or NOT Needing Test (EAP)</p>	<p>The participant expresses the opinion about the fairness or unfairness of L1 English speakers needing to take the test for certain purposes.</p>	<p>I just don't think it's fair. It's- I think it's quite biased, in particular, like why is it that only international students have to prove that they can, you know, enter a university program, that they can speak English? If it's a matter of English proficiency- like proficiency, shouldn't like locals have to do it too? Because, I've seen many international students who could probably speak better than locals here too. So, why do they have to prove it just because they're born in a different country, even though their official language in that country is English? What makes them any</p>	<p>P2, P3, P4, P5, P6, P7</p>	<p>9</p>

		different from people who were born here? I don't think it's fair. P3		
<i>Taking IELTS is Currently Necessary</i>	<i>The participant offers the opinion THAT the test is necessary for a purpose.</i>			
Immigration, Citizenship, or Permanent Residency (necessary)	The participant offers the opinion that the test is or isn't necessary for immigration (i.e., applying for citizenship or PR).	In probably the next year or two, yeah. It's not may. It's probably a need, because that is a requirement for that application too. P3	P2, P3, P6, P7	8
Citizenship (not necessary)	The participant offers the opinion that the test isn't necessary for citizenship.	Yeah, I haven't really looked into, like, the specific details of it, but I just know that an English skill test is probably not required anymore. But yeah, I'm- I'm just applying- I'm just hoping to apply for it myself online because that's what I did for um PR, so yeah. P7	P7	1
Work (not necessary)	The participant offers the opinion that the test is or isn't necessary for work purposes.	I don't think I've done anything, because I didn't list it in my resume and no one asked for it. So, and I don't need an IELTS test to know that I can communicate with people properly. It's not representative. People talk to me and know me in real life know that I can speak English, so there's no need to do that. P5	P5	1
<i>Taking IELTS Should be Necessary</i>	<i>The participant explicitly offers an opinion about whether IELTS should be necessary for a purpose.</i>			
Citizenship (should be necessary)	The participant offers the opinion about that the test should be necessary for Canadian Citizenship.	No, they they should there should be a minimum, at least a minimum, because you have to live the area where most of the people speak English, so they should they should know they, yeah, they should be, they should be, I mean, have a minimum of understanding of the language. P5	P5, P7	4
Citizenship (should not be necessary)	The participant offers the opinion about that the test should NOT be necessary for Canadian Citizenship.	but I guess if you've already- if you're applying for PR and you're required to do an English speaking test, then when you apply for citizenship, you won't be- you shouldn't meet to do one, because it's probably within, like, two or three years that you're applying for PR and citizenship, so, and I don't think your English skills would, like, would go- would decrease it within that period of time. So yeah, I just don't think it's necessary to take English tests for both PR and citizenship, but definitely PR. For sure. P7	P7	1
Permanent Residency (should be necessary)	The participant offers the opinion about that the test should be necessary for Permanent Residency in Canada.	I do see why an English test would be required [for PR], particularly for citizens of countries that- where English isn't the first language P7	P5, P7	5
Permanent Residency	The participant offers the opinion about that the test should NOT be	18 years later, you know, like having two degrees from two English-speaking countries, I was actually asked to do it again for my P.R.	P2, P6	6

(should not be necessary)	necessary for Permanent Residency in Canada.	application. Which I thought was pretty dumb, because, you know, I was actually like having a conversation with with someone from the Canadian immigration authority, I just asked them like, hey, like last time I checked like Canada, U.K. spoke English, I don't know if it's changed, P6		
University Admission (should be necessary)	The participant offers the opinion that the test should be necessary for university admission.	I think- that standard standardized test is needed for like university entrance exam P5	P3, P4, P7	6
University Admission (should not be necessary)	The participant offers the opinion that the test should not be necessary for university admission.	And also I would feel like- is there like a really a need for a standardized test even though we like study English? And they were like something, as I got a good score in like since my like high school, also they asked for my academic transcript- I had some really good score in grade 11 and grade 12. P2	P2	1
<i>Relevance of IELTS to Communication</i>	<i>The participant offers an opinion about the relevance of the test (or part of the test) to the communication required in one of these settings.</i>			
Academic Setting	The participant offers an opinion about the relevance of IELTS (or part of the test) of the communication required in an academic setting.	That it depends on what kind of topics are given to you during the IELTS test because we do- we have some assignments in class that require you to- to talk and discuss on certain topics, but that can be relevant to IELTS not at all, so I put somewhat agreed because I feel like there's some topics that could apply to this, but not others so.... P3	P1, P2, P3, P4, P6, P7	29
Life (Daily English Use)	The participant offers an opinion about the relevance of IELTS (or part of the test) to the communication required in daily English use (or "real life").	Yeah, I feel like for familiar topics, I feel like it's only relevant probably for the first few days when you're just meeting like- the ice breaker period, where you can talk and make new friends, but other than that, once we start getting into like academic topics, no one's gonna ask you like you know, what do you like about home or like, what's your favorite subject, like that is not relevant anymore. So, yeah. P3	P1, P3, P5, P6, P7	23
Representativeness of Score Received to English Ability of Performance	The participant gives an opinion about IELTS with respect to the representativeness of the score received to their actual perceived performance on the test.	I wish I got better scores, but I think it was pretty accurate considering like how I did on that day itself. Like I said, anxiety and stuff, I was pretty okay with that score. Whether it's representative of my academic degree, I would say, yeah, I suppose so. P3	P3	1
The Fact that IELTS is a Money-Making Venture	The participant offers an opinion about the FACT that IELTS is a money making venture for the owners of the test. Code discussion of HOW MUCH the test costs separately under "Cost".	the whole system gradually makes so much money out of this. Because, regardless of like what type of IELTS you take, you would still have to pay, I believe 320 dollars for a single test. So just imagine, if, like.... for like, when the same day you have like 10 people taking the test, you would make like 3,200 dollars. So	P6	3

		that's like a lot of money. So I feel like the system was made hust to kind of like to facilitate that. And they're like, well, if we can make more dollars out of this, why not just make them do that regardless of their nationality? P6		
Not Understanding Grading Criteria	The participant reports that they were confused about the IELTS marks/grading criteria.	I was confused about the marks P5	P5	1
Usefulness or Helpfulness	The participant gives an opinion about IELTS with respect to its usefulness or helpfulness to the participant.	And then I think for like degree requirement they say for certain degrees if you did like an English-speaking test, you won't have to do like first year English classes in university. But for my degree, I had to do it anyway. It was a required course. So, it didn't really bring me much benefit, other than, you know, giving the chance to apply for university. P3	P2, P3, P5, P6	8
Variability of Topic	The participant offers an opinion about the variety, variability, or perceived randomness of IELTS test topics.	Um, yeah. I dunno, I mean like I think.... I get-[exhales] the whole thing here why I feel sort of like it's not really the most reliable test ever, it's because.... the topic can be so random and sometimes it's something that you don't really like, what's that? I've never even heard of that term before. And you just end up like making things up just because, you know, like you don't know anything about <beers>, for instance, like black bears, whatever it is right? Or like, if you grew up in Sweden, and then, you know, you gotta talk about snakes in Southeast Asia, I'm like, oh, wow, like I know nothing about them, but let's just make things up here. P3	P3	1

PREP - Describes Preparing for IELTS	<i>The participant describes things they did or did not do before the test to prepare.</i>			
Name	Definition	Example	Files	Refs
Describes Preparing for IELTS	The participant describes preparing for IELTS. They may go into detail about how they prepared or they describe another aspect of preparation such as how long they prepared for.	And I didn't prepare for this test, like, formally like I know some people who would buy exams, formal exams and stuff like that to prepare and go to like classes to prepare for IELTS. I did not, because my sibling took the IELTS like multiple years ago before me. So, she had an old book. I just kind of flipped through it to see what the format was like, and I just went and did it. P3	P1, P2, P3, P4, P6	12

REASON - Gives Reasons for Choosing or not Choosing IELTS		<i>The participant gives their reason for choosing or not choosing to take the IELTS test.</i>		
Name	Definition	Example	Files	Refs
Accepted for PR in Canada	Describes choosing IELTS because it is accepted as part of Canada's permanent residency application process.	I think I would prefer- what I found out about IELTS that I think was very interesting is, like, the possibility to talk to someone in the speaking section because I think compared to the TOEFL- to the TOEFL, like, experience, I think it'll be better, yeah, to myself. I think now that I, like, I have more knowledge about IELTS, I think – even, like, IELTS has, like, some- how can I say, formats or things that I need to follow, I think it's more, like, flexible than TOEFL, yeah. And then also if- I don't, I don't know if Canada considers TOEFL to the residence- permanent residence, yeah. P1	P1	1
Didn't Get a Good Score in TOEFL	Participant describes choosing to take IELTS because they previously took TOEFL and did not get a high enough score for their needs.	So I use my IELTS score. I did not get any good score in TOEFL, so that's why- one of the other reason I have to get- Uh sorry, I keep mixing between TOEFL and IELTS. Sorry, it was (???). I did not get a good score, so I wanted to improve and I wanted to try something different. So, I made sure IELTS. P2	P2	1
Familiarity with the Test	Participant describes choosing IELTS because they were familiar with the test as a whole or some aspect of the test.	So, I'm familiar with like more IELTS kind of, and I performed good in the past so I would think I would do like good in IELTS for the future. P2	P1, P2, P6	3
Preferred Test Format	Participant describes choosing to take IELTS because of the format of the test.	what I found out about IELTS that I think was very interesting is, like, the possibility to talk to someone in the speaking section [vs TOEFL, that relies on responding to a pre-recorded prompt] P1	P1, P2	2
TOEFL AND IELTS: Accepted for University Admission in USA	The participant describes choosing to take TOEFL AND IELTS because of the perception that TOEFL is more widely accepted by universities in the United States and IELTS is more widely accepted in Canada.	Interviewer: And why did you take both IELTS and TOEFL? Participant: So for TOEFL, I was going into the United States, initially, but then I changed my mind and came to Canada P2	P2	1
TOEFL NOT IELTS: Cost	The participant describes choosing to take TOEFL rather than IELTS because TOEFL is less expensive.	I think it [i.e., TOEFL] was cheaper back in time than IELTS, for example. P1	P1	1
TOEFL NOT IELTS: Preferred Variety of English Used	The participant describes choosing to take TOEFL rather than IELTS because they believed that TOEFL used North American English and IELTS used British English.	Also, I think at that time I thought that IELTS will require me, like, British English, instead of only like English, so I was, like, Okay, I will choose TOEFL because, yeah, I want to go to Canada and I think it'll be like the same English. Yeah, kind of the same English- North American English. P1	P1	1

CELPIP NOT IELTS: Convenience	The participant describes choosing to take CELPIP rather than IELTS because of how convenient taking CELPIP is in terms of accessing the test.	for CELPIP the dates- there were more dates to choose from and the process just seemed easier cause all four parts of the test were in one sitting as opposed to IELTS - you have to, like, book a specific time for the speaking portion. So yeah, so I just, I prefer the convenience over, yeah, over anything so. P7	P7	1
CELPIP NOT IELTS: Cost	The participant describes choosing to take CELPIP rather than IELTS because CELPIP was less expensive.	I think it [i.e., CELPIP] was because it was cheaper [than IELTS], I think. I think cost-wise, it was cheaper P7	P7	1
CELPIP NOT IELTS: Preferred Test Format	The participant describes choosing to take CELPIP rather than IELTS because they prefer the format of CELPIP.	I likely won't take IELTS again. I was thinking maybe taking CELPIP, so maybe try a different one because I know that for that speaking component for CELPIP, that's an electronic component, it's not a face to face. So, I'm hoping that maybe because of that, anxiety will be lower, but we'll see. P3	P3, P7	3
DET NOT IELTS: Cost	The participant describes choosing to take DET because of rather than IELTS because DET was less expensive.	First time. Yeah, [the DET] cheaper is or personally, it wasn't the main reason [for taking the DET rather than IELTS] P4	P4	1
DET NOT IELTS: Preferred Test Content	The participant describes choosing to take DET rather than IELTS because they prefer the content of DET.	And and then even like.... for Duolingo, it's I think it's, it's a good test. Because, like I think IELTS it's more academic. Duolingo is a kind of.... it's like you meet someone on the street and if, like you can speak or you can't. You you don't have time to prepare for anything. So it's it's good test. P4	P4	1
DET NOT IELTS: Less Time Consuming / Difficult to Prepare for.	The participant describes choosing to take DET rather than IELTS because they perceived the preparation required for IELTS to be more time consuming and difficult.	So I I was preparing to the.... doing preparation, doing IELTS preparation at first, but it's too heavy. I was working. I don't have enough time.... And it's hard. I I I didn't give a try, so I don't know. I'm not like going to say it's too hard for me or no, but I'm sure that it needs.... long time, MUCH time to review, to prepare. So and then I don't, then I don't have enough time to do that. Okay, because things went quickly and shortly. I have time to prepare, like I don't have the six months or one year to prepare. I work, so I don't have time. So yeah, I just like try the Duolingo. I prepared.... I was preparing for three months before, like before doing the Duolingo and then I said okay, I'll do the Duolingo, I prepared for one week. Just one week.... /<and I did it>/ P4	P4	1

STRAT - Describes Test-Taking Strategies Used (or Didn't Use) for IELTS		<i>The participant describes strategies they used or did not use during IELTS</i>		
Name	Definition	Example	Files	Refs
Making Notes	Participant describes making notes during the speaking test.	I think- Also like the longer response preparing I was going to like write about my thoughts, it also like- note taking also helped me to like keep organized my thoughts P2		
Being Concise	Participant describes consciously trying to be concise while taking IELTS.	I think- Also like the longer response preparing I was going to like write about my thoughts, it also like- note taking also helped me to like keep organized my thoughts and like be under like, when I was given like a time limit to convey like my amount of thoughts I want to say like, concise and like a specific, a specific manner and be like detailed and not to create some like- [coughs] Sorry- not to create some long sentences, and carry forward the conversation and convey my point in less amount of time. So, yeah, it definitely helped me. P2	P2	1
Holding Back	Participant describes consciously holding themselves back during IELTS. Holding back here can be understood as limiting or restricting themselves during the test in a way they may not have done in a non-test act of communication.	Let's put it that way. I could flow so much better, because I knew like, well, you know, I'm gonna have two minutes here, so let's just do it. Although it was not like, you know exactly two minutes, but at least it gave me that sort of like window for me to be able to express myself a bit more, and I felt like it was more natural because I didn't really kind of like, it's basically like what I'm doing with you now, you know. It was really more about conversational English rather than like, alright, what do you think about this? One sentence full, stop, bam! Alright, next, you know. Whereas with the second part it was more like, alright. let me talk about this. Although, you know I really really tried to held myself back from like saying something like oh, damn! I forgot about that one, you know P6	P6	1
Modifying Speech in Response to the Variety of English Spoken by the Examiner	Participant describes consciously thinking about and modifying how they speak during the test in response to the variety of English spoken by the examiner.	Interviewer: When you took I guess, IELTS in the different countries did you modify your speech according to the examiner that you had, in any way? Participant: I had to do it, I guess. I think there is some kind of a like pattern that everyone has to follow.	P6	1

		P6		
Not Consciously Thinking about Criteria or Strategies	Participant describes NOT consciously thinking about testing criteria during the test. OR participant describes NOT consciously using a strategy during the test.	For example, uh, when I was answering, I don't know if I answered in the format that I don't know, an IELTS will require, you know, like, because it was just like "blah blah blah" talking, but, like, if I was, like, if I were in an official test, I think I would be -I'll fail because yeah, I was just talking, blah blah blah. I was not thinking about formats, yeah. Yeah, I didn't give, like, a structured answer, I think. P1	P1, P2, P3, P4, P5, P6, P7	9
Reflecting Back on Earlier Parts of The Test in Part 3	Participant describes consciously thinking about Parts 1 and or 2 during Part 3.	Yeah, the Q&A discussion [in part 3] was like the amount of notes I made, and I was able to follow like the conversations from the notes and added on my- some notes on my own at like, the same time. So, I think it was a good balance, that kind of information that I have to speak. I was able to reflect back on the earlier parts and added some on my own. P2	P2	1
Reflecting or Thinking Nonstop	Participant describes thinking nonstop as a strategy as opposed to trying not to think too much during the test.	For the second part- or even for the part when they give you like 2 minutes to think and you can write down ideas, but then you're constantly thinking about what to say next. Does that make sense? Like I say, just thinking on the spot, because sometimes ideas will come, but they're not exactly fully formed. So, you have to expand on them in your brain before speaking and when you have no time to do so, things just come out awkwardly, or disjointedly. So, definitely I've been thinking about nonstop. P3	P1, P2, P3	2
Thinking about Vocabulary	Participant describes consciously thinking about word choice or another aspect of vocabulary during the test.	So, I kind of wanted to exaggerate how skilled I am by using, like, unnecessary vocabulary [laughs]. P7	P3, P4, P7	4
Thinking about Grammar	Participant describes consciously thinking about some aspect of grammar during the test.	Interviewer: So when you were doing the [IELTS] test, were you actively thinking about those different [criteria] components as you were answering the questions in that Duolingo test? Participant: Okay. Not that much. I always have like this in my mind... but I can't focus too much because I have to focus in many things and, like, understand the question, providing the right answer, and then I can't also like check if my answer will satisfy vocabulary standards and grammar standards. But however, I'll always keep that in mind like trying to put some good vocabularies, or paying attention to my grammar. P2	P2, P4, P6	3

Thinking about the Test Format	Participant describes consciously thinking about “format” during the test.	Oh yeah, for sure. I'll be very scary, I think, [laughs] if I were in a real test. I'll be very, like, “Oh my God” I need to say then, the format that they want, using the vocabulary that way from me. Yes, I was more relaxed [during the simulated test]. Which is sad, I think. P1	P1	1
Thinking about Time Limit	Participant describes consciously thinking about the time limit of the test (or a part of the test) while they took the test.	Yeah I sort of l- I sort of, I mean, I don't think it was like precisely 1:30 but it was more like, alright, I can't really function like what I would do if I was still like when I was <working as a T.A.>, for instance, right, when I was working as a T.A., I would just wanna like ramble on, you know, like, this is basically what we're gonna do today, guys. Blah, blah, blah blah, and I could basically speak for like one hour, non-stop, right? So with that, I was like, alright, gonna change my mind here [i.e., during the speaking test], and the mindset had to be changed had to be switched, so I think to some extent it helped because the mentality changed. P6	P6	1
Thinking of the Test as a Casual Conversation	Participant describes thinking of the test as a casual conversation during the test.	during the mock speaking test, I wasn't briefed on the criteria, I guess. I just knew it was to assess my fluency, but I didn't know, like, grammar was involved as well. And I guess in the actual test that I took in 2014, I probably did have those criteria in mind, but to me, it was a little bit overwhelming to just have the criteria in mind. So, what I did was just thinking of it as like a casual conversation I'm having with someone- that made it easier for me. So, yeah. P7	P7	1
Translating in My Head	Participant describes “translating in my head” as a strategy used during the test.	It's an exam, so you're not, like, spontaneous that I think for- I feel like as having English as a second language, like, I'm always, you know, in my mind translating and then thinking about everything, you know, and then when you got more pressure in a real exam, like, yeah, it's a total mess! P1	P1	1

