Test impact as dynamic process: Individual experiences of the English test requirements for permanent skilled migration in Australia

Kellie Frost

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

July 2017

School of Languages and Linguistics, Faculty of Arts
University of Melbourne
ABSTRACT

The use of language tests in immigration policy contexts raises critical questions concerning the adequacy of current conceptualisations of test purpose, impact, and validity in the field of language testing. Within existing theoretical frameworks, these notions rest on the premise that intended test purpose and intended score meanings exist as coherent, fixed, and stable entities, as defined by test developers, and these intentions must be specified as a precursor to evaluations of consequences and validity. Moreover, within existing frameworks, test takers exist primarily as theoretical abstractions rather than as real persons, deconstructed into the components of knowledge and skills that constitute test constructs. Such underlying assumptions, it is argued in this thesis, have limited our capacity in the field of language testing to understand and explain the role of language tests in immigration and other policy domains, where score meanings, test purposes and test consequences are shaped by multiple and often conflicting interpretations and values, as stakeholders, including test takers, struggle to realise their own intentions.

This thesis explores the dynamics of test impact in the context of Australia’s skilled migration policy by examining how four individuals responded to the English test score requirements that exist within the policy as they sought to transition from temporary to permanent resident status. Temporary visa holders in Australia are able to become permanent residents via the skilled migration program if they possess certain specified skills and attributes, including English language proficiency as demonstrated on an accredited language test, such as IELTS. Since 2011, test scores of IELTS 7 and 8 have been heavily weighted in the overall selection process for skilled migrants, and thus represent a vital transition mechanism in the trajectory of those seeking to move from the status of temporary to permanent resident in Australia.
A series of in-depth, open-ended interviews were conducted over an 18-month period with each of the four individuals during their migration trajectory from temporary to permanent residency, a transition that for each of them involved repeated language test attempts. A grounded theory approach guided analysis of interview data, which focused on identifying how each person thought, felt and acted in response to the test requirements, how their perceptions and feelings shifted over time as they interacted with the test, and how their test experiences shaped their lives and influenced their actions and decisions as they sought to realise their migration intentions in this highly constrained policy space.

Findings show that individual subjectivities and agency intersected with the constraints of the test and score requirements in the context of Australia's skilled migration policy to generate diverse and multidirectional consequences in each of the four migrant cases, highlighting the indeterminable nature of test purpose and test impact in this policy domain, and supporting the argument made throughout the thesis that the perceptions and experiences of those subjected to testing regimes, rather than the intentions of test developers, must constitute the focal point around which to evaluate the appropriateness of testing practices in policy domains.
DECLARATION

This is to certify that

i) the thesis comprises only my original work towards the PhD;

ii) due acknowledgement has been made in the text to all other material used,

iii) the thesis is fewer than 100 000 words in length, exclusive of tables, maps, bibliographies and appendices

.................................................................

Kellie Frost
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professor Tim McNamara and Professor Jill Wigglesworth, for their generous support and guidance, and for the expertise and inspiration they each offered in various ways throughout the writing of this thesis.

I am very grateful to all my friends and colleagues in the LTRC, past and present, and in the School of Languages and Linguistics for their endless support and encouragement. I owe special thanks to Associate Professor Paul Gruba for his wisdom and generous help on so many occasions. Thanks also to my colleagues, Dr Sally O’Hagan and Craig Wattam, for their assistance in the data coding process.

I am especially grateful to the individuals who participated in this study. I would like to thank Ana, Erfan, Sharon, and Yuki for their generosity in sharing their experiences during what was a very stressful period for each of them. I also owe a debt of gratitude to the seven teachers who contributed their time and valuable insights to this study.

Most of all, thanks to my family, Gareth and Harry, for their support in all ways, and to my parents, who have always encouraged and supported me in everything I do.
## CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Language testing in immigration policy domains</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>The current study</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Chapter overview</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>English language testing and skilled migration in Australia</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Background to the skilled migration program</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1</td>
<td>From race-based to points-based selection: the beginning of skilled migration</td>
<td>11</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Skilled migration and English: from implicit criteria to formal test score requirements</td>
<td>13</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Skilled Migration Policy reforms throughout the 1990s</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>The Study-Migration Pathway</td>
<td>17</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Establishing the pathway (1999 to 2005)</td>
<td>17</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Changes to the pathway (2005 to 2008)</td>
<td>19</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Reforms from 2009 onwards - decoupling study and migration</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>The New Points Test</td>
<td>24</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Key changes to selection criteria</td>
<td>25</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Potential impacts of policy changes</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>Conclusion</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Language tests in policy spaces: challenging existing theory</td>
<td>32</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Test purpose in language testing</td>
<td>35</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Intended test construct and test use</td>
<td>35</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Intended test use and test consequences</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>“The Power of Tests” (Shohamy, 2001)</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Test impact in policy contexts</td>
<td>49</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Tests and education policy</td>
<td>49</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Tests and immigration policy</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>Poststructuralist approaches to rethinking test impact and validity in language testing</td>
<td>57</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusion</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Fairness frameworks: social values and the politics of test impact</td>
<td>66</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>66</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>4.2 Validity, values and social consequences – revisiting Messick</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>4.3 The ‘fairness’ debate in Language Testing</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>4.4 Reconceptualising ‘justice’ – Sen (2009)</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>4.5 Implications for language testing</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>4.6 Conclusion</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Chapter 5: Research methods</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>5.2 Methodology</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>5.2.1 A case study approach</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>5.2.2 In-depth qualitative interviewing</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>5.2.3 Grounded Theory methods</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>5.2.4 Summary</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>5.3 Procedures</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>5.3.1 Participants</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>5.3.2 Data collection</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>5.3.3 Data preparation</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>5.3.4 Data analysis</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>5.3.5 Ensuring trustworthiness</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>5.4 Conclusion</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Chapter 6: Individual experiences - Ana, Erfan, Sharon and Yuki</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>6.2 Key themes</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>6.3 Overview of individual migrant cases</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>6.4 Ana</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>6.4.1 Introduction</td>
<td>118</td>
<td></td>
</tr>
<tr>
<td>6.4.2 Experiencing learning and using English</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>6.4.3 Summary of themes: Experiencing learning and using English</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>6.4.4 Before test attempt 1</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>6.4.5 Experiencing repeat test encounters</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>6.4.6 Summary of Ana’s experiences</td>
<td>140</td>
<td></td>
</tr>
<tr>
<td>6.5 Erfan</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>6.5.1 Introduction</td>
<td>143</td>
<td></td>
</tr>
<tr>
<td>6.5.2 Experiencing learning and using English</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>6.5.3 Summary of themes: Experiencing learning and using English</td>
<td>148</td>
<td></td>
</tr>
<tr>
<td>6.5.4 Before test attempt 1</td>
<td>149</td>
<td></td>
</tr>
</tbody>
</table>

vi
6.5.5 Experiencing repeat test encounters ...................................................... 153
6.5.6 Summary of Erfan’s experiences ............................................................ 162
6.6 Sharon ........................................................................................................ 164
    6.6.1 Introduction ....................................................................................... 164
    6.6.2 Experiences learning and using English .............................................. 166
    6.6.3 Experiencing repeat test encounters ................................................. 167
    6.6.4 Summary of Sharon’s experiences ..................................................... 179
6.7 Yuki .......................................................................................................... 181
    6.7.1 Introduction ....................................................................................... 181
    6.7.2 Experiences learning and using English .............................................. 183
    6.7.3 Experiencing repeat test encounters ................................................. 188
    6.7.4 Summary of Yuki’s experiences ........................................................ 199
6.8 Summary of the Four Cases ..................................................................... 200
Chapter 7: Teacher perspectives ................................................................... 202
    7.1 Introduction ........................................................................................... 202
    7.2 The seven English teachers .................................................................... 203
    7.3 Key themes ............................................................................................ 204
        7.3.1 Appraising test takers’ English abilities .......................................... 204
        7.3.2 Understanding score meanings ...................................................... 207
        7.3.3 Characterising the test taker perspective ........................................ 213
        7.3.4 Negotiating score meanings .......................................................... 218
        7.3.5 Losing agency in the classroom ...................................................... 219
    7.4 Summary ................................................................................................. 220
Chapter 8: The dynamics of test impact ........................................................ 222
    8.1 Introduction ........................................................................................... 222
    8.2 A model of the dynamics of test impact ................................................ 224
    8.3 Individual subjectivities, agency and test consequences ....................... 226
    8.4 Validity and fairness issues ..................................................................... 233
        8.4.1 Conventional validity issues .......................................................... 234
        8.4.2 Fairness issues ............................................................................... 240
    8.5 Justice as consequences ......................................................................... 243
    8.6 Conclusion .............................................................................................. 244
Chapter 9: Conclusion ..................................................................................... 245
    9.1 Introduction ........................................................................................... 245
    9.2 Overview of Theoretical Challenges ..................................................... 245

vii
9.3 Summary of main findings ................................................................. 246
9.4 Implications for language testing .................................................... 249
9.5 Limitations and future research ....................................................... 249
REFERENCES ................................................................................................ 251
APPENDIX 1: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (ANA) .............. 264
APPENDIX 2: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (ERFAN) ............. 269
APPENDIX 3: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (SHARON) .......... 272
APPENDIX 4: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (YUKI) ............... 275
List of Tables

Table 2.1. Comparison of key criteria in 2010 and 2011 Points Tests ......................... 27
Table 5.1. Overview of migrant participant details ....................................................... 99
Table 5.2. Interview schedule and duration ................................................................. 102
Table 5.3. Transcription conventions ........................................................................ 104
Table 5.4. Example of initial coding process ............................................................... 107
Table 6.1. Self-appraising and practising agency: Linking Ana’s self-perceptions and actions .......................................................................................................................... 126
Table 6.2. Self-appraising and understanding target scores: Linking Ana’s self-perceptions and test perceptions ........................................................................................................ 126
Table 6.3. Ana’s IELTS scores ....................................................................................... 129
Table 6.4. Ana’s use of feedback from tutor 1 ............................................................. 131
Table 6.5. Ana’s use of feedback from tutor 2 ............................................................. 133
Table 6.6. Ana’s use of feedback from tutor 3 ............................................................. 137
Table 6.7. Summary of Ana’s shifting thoughts, feelings, decisions, and actions over time ............................................................................................................................. 142
Table 6.8. Summary of themes: Erfan’s experiences of learning and using English .... 148
Table 6.9. Erfan’s strategic approach to the test: Linking perceptions, feelings and actions ........................................................................................................................................ 152
Table 6.10. Erfan’s IELTS scores ............................................................................... 153
Table 6.11. Summary of thematic relationships: linking Erfan’s perceptions and actions ................................................................................................................................. 156
Table 6.12. Summary of relationships between themes after Sharon’s first test attempt ................................................................................................................................. 172
Table 6.13. Summary of relationships between themes after Sharon’s second test attempt 2 ............................................................................................................................ 179
Table 6.14. Self-appraising, interpreting scores, and practising agency: Linking Yuki’s self-perceptions, test perceptions and actions ......................................................... 186
Table 6.15. Yuki’s IELTS scores ............................................................................... 189
Table 6.16. Yuki’s OET scores .................................................................................... 189
Table 6.17. Summary of relationships between themes after Yuki’s first three test attempts .......................................................................................................................... 193
Table 7.1. Positive appraisals of migrants’ abilities to communicate in English ........ 204
Table 7.2. Describing IELTS scores in terms of performance attributes .................. 208
Table 7.3. Comparing IELTS 8 to English ‘native’ speaker ability level ..................... 209
Table 7.4. Characterising test taker perceptions of the score requirements ............... 215
Table 7.5. Characterising test taker affective responses ............................................. 215
Table 7.6. Characterising test taker action ................................................................... 216
List of Figures

Figure 6.1: Summary of thematic relationships (Ana) .................................................. 140
Figure 6.2: Summary of thematic relationships (Erfan) ............................................ 163
Figure 6.3. Summary of thematic relationships (Sharon) ........................................ 180
Figure 6.4. Summary of thematic relationships (Yuki) ............................................. 200
Figure 8.1. The Dynamics of Test Impact ............................................................... 225
CHAPTER 1: INTRODUCTION

1.1 Introduction

Language tests are now widely used by governments in various national contexts as screening or selection tools for controls, including restrictions, on immigration and access to citizenship. While this has prompted significant debate within and outside the field of language testing over the appropriateness, validity and fairness of the use of language tests in these policy domains, efforts to incorporate evaluations of the wider social impacts of test use into test validation arguments remain limited.

As argued throughout this thesis, this failure is largely because the sorts of questions of impact, validity and fairness raised by the use of language tests as instruments of immigration policy cannot be adequately addressed using existing theoretical frameworks, centred as they are on the need to specify the intentions of test developers and, by association, test users as a precursor to identifying and evaluating consequences and associated validity and fairness claims. In immigration policy contexts, tests can serve multiple and sometimes conflicting purposes, with consequences likely to extend well beyond those intended or anticipated by test designers.

In locating intended test purpose, as imagined by test developers, as the locus of validation frameworks, the ways tests actually function in policy domains, especially the unanticipated roles tests play in the lives of individuals subjected to score-based judgments, remain hidden from view, thereby limiting our capacity to identify and account for the consequences that emerge as a result of testing practices. It is argued here that a more nuanced understanding of test impact, which foregrounds the experiences and perceptions of individuals as they interact with language testing practices in the context of their migration journeys, is urgently needed to enable the theoretical advances in validity and fairness frameworks required to account for the use of tests in these high-stakes policy domains.
In this chapter, a brief overview of the literature that has emerged in response to the increasingly widespread use of language tests in immigration policies around the world provides a background to the current study and to the theoretical issues it seeks to address. Following this, the research aim and focus of the current study is outlined. The chapter concludes with an overview of the contents of the remaining chapters.

1.2 Language testing in immigration policy domains

As mentioned above, the use of language tests as instruments of immigration and citizenship policies in various countries, including Australia, has pushed the field of language testing to move beyond conventional psychometric investigations of fairness and test validity into the more difficult and complex realm of the social and political dimensions of testing (McNamara & Roever, 2006). Given the high-stakes nature of language testing in these contexts, questions of the broader social impact of the use of language tests, and the values and ideologies that such testing serves to reinforce, are increasingly raised in the language testing literature (e.g. Bachman & Palmer, 2010; Kunnan, 2000, 2004, 2010; McNamara and Ryan, 2011; Shohamy, 2001, 2006, 2009; Shohamy & McNamara, 2009; Xi, 2010).

There has also been an emerging body of research outside the field examining language testing practices within immigration and citizenship policies (see edited volumes by Extra, Spotti, and Van Avermaet, 2009; Hogan-Brun, Mar-Moliner & Stevenson, 2009; Slade and Möllering, 2010). These collections examine various relationships between language tests, language ideologies, and discourses of integration and national belonging in country-specific contexts. The emphasis is typically on the gatekeeping role tests play as political instruments in these policy domains, and the potential for tests to promote or reinforce exclusionary social norms. Researchers have argued, for example, that language tests privilege monolingualism as the natural and preferred state, thereby de-valuing minority languages and perpetuating the exclusion of minority speakers (e.g. Blackledge, 2009; Horner, 2009; Shohamy, 2009; Van Avermaet, 2009). It is also widely argued that language testing reinforces an ideology that places cultural homogeneity as
the norm in nation-states, and represents minority languages and multilingual practices as a threat to social cohesion and security (Blackledge, 2009).

In connection with these and other arguments, issues of justice are raised to varying degrees. Van Avermaet (2009), while suggesting that language test design should be modified to reflect what might be realistically expected of migrants, concludes that integration policies “should be based on human rights principles” (2009, p. 38). Shohamy (2009) also advises that language tests should be better designed so as to reflect more closely the diversity of language (including multilingual) resources that are involved in social participation, although she clearly contests the existence of such tests. She argues that common misconceptions about language, tests, and citizenship have led to discrimination against migrants and the violation of their basic rights, and offers the above suggestions as a way forward given the unlikelihood that testing regimes will be abandoned.

In some cases, traditional test validity arguments have been conflated with moral arguments around issues of justice arising in these contexts. For example, Piller (2001), in a critique of the introduction of language testing for citizenship purposes in Germany, argues that such testing is undemocratic and unjust. Although she is thereby contesting the values implicit in the practice of language testing, her argument is largely supported by evidence that the test lacks linguistic validity. More recently, in the Italian context, similar moral objections to the introduction of language tests for migrants seeking long-term residency have also been supported, for the most part, by evidence against the linguistic validity of the testing materials and procedures (Barni, 2010, 2011). Arguments such as those put forth by Piller and Barni inadvertently imply that improving the validity of testing practices would lead to more just outcomes for would-be migrants and citizens, as has been suggested recently by some in the field of language testing (De Jong, et al., 2009; Saville, 2009). Shohamy and McNamara (2009) challenge the underlying logic of these claims, asserting instead that more linguistically valid testing instruments simply serve as better disguises for an inherently unjust practice.
The conflation of test qualities and issues of values is also evident in current debates among language testers, highlighting the complex relationship between linguistic validity, test consequences and the social and political values inherent in test constructs (Davies, 2010; Kane, 2010; Kunnan, 2010; McNamara & Ryan, 2011; Xi, 2010). McNamara and Ryan propose a distinction between ‘fairness’ and ‘justice’, whereby fairness refers to “all aspects of the empirical validation of test score inferences in the interests of yielding reasonable and defensible judgments about individual test takers” and justice refers to “the consequential basis of test score interpretation and use but also, and particularly, the social and political values implicit in test constructs” (2011, p. 167). Such a distinction opens up a space within validity theory to question the very existence of language tests on the basis of justice arguments, although how such arguments should be evaluated in practice is yet to be established.

As noted above, the validity and fairness issues raised by the use of language tests remains an ongoing concern in the field of language testing. Little is yet known about how individuals perceive, experience and negotiate test score requirements over time as they seek to realize migration intentions, nor about how their lives are affected by testing practices. Further insights into how test takers interact with and respond to tests in these high stakes policy contexts are urgently needed in order to further our understanding of test impact, and to enable the theoretical advances in validity and fairness frameworks needed to account for the use of tests as policy instruments.

1.3 The current study

The aim of the current study is to explore how test impact emerges in the context of Australia’s skilled migration policy, by examining how four individuals already residing in Australia on temporary visas perceive and respond to the English language test requirements that exist as part of the process of transitioning from temporary to permanent resident status. While a limited number of studies have examined the impact of language test use from the perspective of vulnerable migrants (Cooke, 2009; Khan, 2013; Strik, Böcker, Luiten & van Oers, 2010), little attention has yet been paid to problematizing the existence of formal language test requirements within skilled
migrant selection processes in Australia or elsewhere, nor to the impact of language testing on the lives of migrants in this category seeking permanent residency rights.

In particular, the study addresses the following research questions:

1. How do four individuals seeking to transition from temporary to permanent residency in Australia account for their thoughts, feelings, and actions over time in relation to the English language test score requirements within Australia’s skilled migration policy?

2. What are the consequences that emerge as each individual interacts with the test and score requirements in this policy context?

The four individuals involved in the study all completed tertiary studies in Australia, were employed, and had been living in Australia for several years on temporary visas at the time they joined the study. They all began their attempts to transition to permanent resident status in 2012, just after a suite of policy changes had been implemented which included the introduction of highly restrictive English test score criteria. All four individuals had already satisfied the criteria for becoming permanent residents in Australia when they joined the study, apart from the English test score requirements. They each needed scores of at least IELTS 7 (or equivalent) on all four parts of the test (speaking, writing, reading, and listening) to complete the process.

A series of in-depth interviews were conducted with each of the four individuals over periods of between 6 and 18 months, during which time they made repeated test attempts in an effort to meet the English test score requirements. A grounded theory based analysis was conducted, which focussed on if and how individuals' perceptions shifted over time as they interacted with the test, how their perceptions of and interactions with the test influenced their actions and decisions, and how their test experiences impacted their lives. Interviews were also conducted with seven teachers of IELTS preparation courses. The accounts provided by the teachers served as a means of situating the experiences of the four individuals in relation to the experiences of a
broader group of test takers seeking to meet the English test score requirements for permanent residency as skilled migrants, and of interrogating the plausibility of the fairness and validity issues that emerged from the migrant accounts. While the perceptions and experiences of migrant participants is the primary focus of the study, the inclusion of teacher interviews was motivated by the fact that the perspective of test takers, particularly their criticism of testing regimes, is vulnerable to being dismissed as simply motivated by a lack of test success.

By investigating the processes through which test consequences emerged in each case, the study aims to contribute to the development of a theoretical framework that might account for the role and impact of language tests situated in immigration policy contexts, and to work towards reconceptualising validity and fairness to account for the use of tests in immigration policy domains.

1.4 Chapter overview

Chapter 2 of this thesis provides an overview of Australia’s skilled migration policy, including an account of historical developments that led to formal English test score requirements entering and becoming embedded in the policy. Attention is given to a policy initiative known as the ‘study-migration’ pathway, introduced in the late 1990s as a means of enabling international students of Australian institutions to transition directly to permanent resident status as skilled migrants upon gaining their qualifications. In 2005, concerns were raised about the efficacy of the skilled migration program as evidence emerged that international graduates transitioning to skilled visas, especially those from non-English speaking backgrounds, were failing to gain employment in their fields, despite skill shortages. These concerns, as well as a subsequent economic downturn in the wake of Global Financial Crisis, led to a series of reforms to effectively dismantle the study-migration pathway. The culmination of these reforms was the introduction of new skilled migrant selection criteria in 2011, which included highly restrictive English test score requirements. The experiences of the four migrant participants in the current study are situated in the midst of these changes, as
they were completing their studies in 2011 and 2012 and commencing their efforts to gain permanent residency around the time these policy changes came into effect.

In Chapters 3 and 4, the theoretical challenges raised by the use of language tests as instruments of immigration policies are discussed. In Chapter 3, a critical analysis of the premises underlying current validation frameworks and existing empirical investigations of test consequences is provided. It is argued in Chapter 3 that attempts by language testers to explain and evaluate the impact and validity of testing practices in immigration policy domains have been limited by what is termed in this thesis a ‘condition of determinability’. This ‘condition of determinability’ refers to the premises within existing theoretical frameworks that intended test purpose and score meanings are fixed and stable, as defined by test developers, and that these intentions must be specified as a precursor to any evaluations of validity or consequences. Furthermore, it is argued that there is no space within these frameworks to consider the role of individual agency in shaping the function and impact of tests in policy domains. Chapter 4 extends these arguments through a discussion of efforts in the field of language testing to reconceptualise ‘fairness’ to account for the broader societal impacts of test use. To conclude the chapter, the concept of justice put forth by economist and philosopher Amartya Sen (2009) is introduced in support of the argument developed throughout Chapters 3 and 4, that to properly understand and evaluate the role of tests in policy contexts, test purposes and consequences, as perceived and experienced by those subjected to testing regimes rather than as intended by test developers, must constitute the focal point around which the appropriateness of testing practices is evaluated.

In Chapter 5, the background to the current study and the research methods are described. The findings of the study are reported in Chapters 6 and 7, and discussed in Chapter 8. In Chapter 6, the experiences of the four individual migrant participants in the study are presented as four separate cases. The focus across the cases is on how individuals integrate their experiences learning and using English with their test experiences to generate self-appraisals of their English abilities and interpretations of score meanings, and how these perceptions interact to shape feelings, decisions and actions over time as they negotiate test and policy constraints in an effort to realise their
migration intentions. In Chapter 7, the perspectives of the seven teacher participants concerning the use of English test scores for migrant selection purposes are described, with a focus on how they characterise migrant test takers seeking permanent residency and the experiences of these test takers. In Chapter 8, a model of the dynamics of test impact, derived from the four migrant cases, is first presented, followed by a discussion of the particular consequences that emerged in each case. Individual cases are compared in this discussion to highlight the central role played by individual subjectivities and agency in influencing the consequences associated with test use in the context of Australia’s skilled migration policy. Following this, migrant and teacher perspectives are compared in a discussion of the threats to validity and fairness that came to light throughout the study. The chapter concludes with a reflection on the implications of the study for the field of language testing.
CHAPTER 2: ENGLISH LANGUAGE TESTING AND SKILLED MIGRATION IN AUSTRALIA

2.1 Introduction

The use of English language testing as a mechanism for regulating the permanent skilled migrant intake has been a consistent feature of immigration policy in Australia for over two decades. As will be shown in this chapter, the English language criteria and associated test score requirements involved in selection processes for skilled migration have become increasingly restrictive in recent years. This is ostensibly because of concerns that migrants are more likely to experience poor employment outcomes if they come from non-English speaking backgrounds. As outlined in Chapter 1, the current thesis is primarily concerned with investigating the ways in which individuals from non-English speaking backgrounds, in particular, those who have completed tertiary qualifications in Australia, negotiate and are affected by language test requirements in their attempt to transition from temporary to permanent residency status within this more restrictive regime.

Skilled migration is one of three broad categories that currently comprise Australia’s overall permanent migration program, which also includes family and humanitarian streams. Since the late 1990s, skilled migration has consistently accounted for the largest proportion of the overall program, with around 67 per cent of Australia’s total migrant intake entering through the skilled stream over recent years (Department of Immigration and Border Protection (DIBP), 2015). Additionally, from the late-1990s until the mid-to-late 2000s, an increasing proportion of skilled visas were granted to individuals who commenced their migration journey on temporary visas as international students of Australian institutions. These individuals, until recently, were able to transition directly from temporary student visas to permanent residency upon completion of their studies via a study-migration pathway, which was introduced into Australia’s skilled migration policy in 1999 (Koleth, 2010). As will be discussed in this chapter, within the study-migration pathway, graduates of Australian institutions were
favoured by skilled migrant selection criteria, and benefited from exemptions from the work experience and language test requirements that applied to other applicants seeking permanent residency. However, a series of policy reversals introduced over the past decade and the implementation of new selection criteria in 2011, which include highly restrictive English test scores of IELTS 7 and 8, have meant that the transition from temporary status to permanent residency is no longer certain for former overseas students in Australia.

The use of English language test scores in Australia’s skilled migrant selection processes reflects the multifaceted nature of language tests as instruments of immigration policy, whereby tests and test scores serve multiple, and potentially competing functions. On the one hand, test scores represent a source of evidence in support of, or against, claims that skilled migrants possess the English language competency deemed necessary to be able to work in their occupations in Australia. On the other hand, policy makers may shift test score requirements arbitrarily, as a means of controlling migrant numbers; cut scores increase or decrease depending on labour market demands, or in response to other political aims or policy objectives, some of which may have little, if anything, to do with the actual level of English language proficiency needed to successfully engage in work in Australia.

Shifting English test score requirements have resulted in uncertain migration outcomes for many temporary residents who have spent many years and committed huge financial resources to fulfilling their intention to remain permanently and build a life in the country (Dickie, 2016). The use of language tests in this policy domain, therefore, poses significant challenges for language testers, particularly in relation to evaluations of test consequences and validity. As will be discussed in detail in the next chapter, these evaluations are currently contingent on a clear and definitive articulation of intended test purpose, typically from the perspective of the test developer and assumed to be in consensus with the views of the test user. When language tests are viewed as embedded within policy spaces, however, the inevitable contestability of accounts of policy intentions will necessarily bear down on questions of test purpose, test consequences, and associated fairness and validity claims.
The aim of the current chapter is to provide an overview of the specific policy context within which the current study is located, in order to provide insights into the role language tests play within Australia’s skilled migration policy. The chapter will begin by providing a brief background to Australia’s skilled migration policy, including the introduction of a skilled category into immigration policy, and an overview of the way in which formal English test score requirements entered and became embedded in policy during the 1980s and 1990s. This is followed by an account of the establishment of the study-migration pathway, and a summary of subsequent policy changes that affected the transition process from student visa to permanent resident status. Details of a study conducted by Birrell, Hawthorne and Richardson (2006) are provided, which was commissioned in 2005 by the Department of Immigration to evaluate the efficacy of the skilled migration program, as recommendations made by these researchers led directly to the introduction of higher English test scores into migrant selection criteria.

Following this, changes introduced in the wake of the Global Financial Crisis, from 2009 onwards, will be outlined. The chapter will conclude by providing details of new skilled migrant selection criteria and processes, introduced in 2011 and 2012, with particular attention given to shifts in the weighting of English language test scores. The experiences of the four migrant participants in the current study are situated in the midst of these changes, as they were completing their studies in 2011 and 2012 and commencing their efforts to gain permanent residency around the time that these new, more restrictive selection criteria were implemented.

2.2 Background to the skilled migration program

2.2.1 From race-based to points-based selection: the beginning of skilled migration

A formal migration program was created in Australia in 1945, with the establishment of the first federal immigration portfolio and specialized Immigration Department. For the next three decades, migrant selection was based on a discretionary approach, characterised by a notorious preference for ‘white’ British and European settlers, known
as the ‘White Australia’ policy (Castles, Hugo & Vasta, 2013). Throughout the 1960s and 70s, there was increasing public and media pressure to abandon race-based criteria in favour of more rational and transparent selection processes (Tavan, 2013). In response, a quantitative, points-based model, known as NUMAS (‘numerical multifactor assessment system’) was implemented in 1979. In the new model, various criteria were weighted with a numerical value, and would-be migrants were accepted upon achieving a specified minimum number of points. While the criteria have been modified and re-weighted over time, a points-based model, in which applicants must attain a specified number of points (the ‘pass mark’) in order to qualify for migration, remains in place as a primary selection mechanism for skilled migrants in Australia (Markus, Jupp & McDonald, 2009).

Tavan (2013) explains that NUMAS was introduced as a means of improving the objectivity and consistency of selection procedures and of more effectively aligning migrant intake figures with public policy objectives, including achieving population growth targets and meeting labour market demands. Within NUMAS, points were allocated according to criteria such as age, education, skills, and family connections, and the weighting of selection criteria prioritised those with qualifications and skills suited to Australia’s labour market conditions. According to Tavan (2013), NUMAS served to “identify economically viable, English-speaking skilled migrants” (p. 54).

Within migrant selection processes since 1979, an ability to use English has always been viewed as an important attribute in terms of facilitating migrant integration into the Australian labour force and society. A new selection process replaced NUMAS in the early 1980s, and the migration program was split from a single stream into the three categories that exist in the current policy: skilled, family and humanitarian migration. Within the skilled migration category, selection criteria have since remained oriented towards selecting those with the potential to deliver benefits to the Australian economy, including an ability to communicate in English.
2.2.2 Skilled migration and English: from implicit criteria to formal test score requirements

English language ability became a particularly salient migrant characteristic during the 1980s as growing numbers of skilled migrants from diverse non-English speaking backgrounds were entering Australia (Markus, et al., 2009). Although English language ability was not mentioned explicitly in the list of points weighted criteria in the selection processes that superseded NUMAS, it remained an important implicit criterion throughout the mid-1980s as immigration officers were required to judge prospective migrants’ settlement prospects on a five-point scale ranging from ‘settlement risk’ to ‘outstanding’ (Hawthorne, 1997, p. 12). Hawthorne (1997) points out that such judgements were based, at least in part, on an informal assessment of the migrants’ English language ability. In 1989, following a review of Australia’s immigration policies, English language competency was reinserted as an explicit criterion into the points system but was not listed as compulsory, and English language points were not critical to selection for tertiary-qualified applicants of prime working age (Hawthorne, 1997).

By the early 1990s, however, as Australia, along with many other OECD economies, was in the midst of a worsening recession, skilled migrants from non-English speaking backgrounds were increasingly identified as an economic burden. Perceptions that many were lacking the English skills needed for successful integration into the labour market fuelled calls for changes to migrant selection criteria (Brindley & Wigglesworth, 1997; Hawthorne, 1997). Consequently, in 1992 changes to the points test system were introduced which made English competency a mandatory requirement for many occupations, listed as ‘Occupations requiring English’. A new test of English communicative skills, access:, was commissioned and developed for skilled migration purposes by the National Centre for English Language Teaching and Research at Macquarie University (Brindley & Wigglesworth, 1997). The test was first administered in 1993 but in 1998, access: was replaced by IELTS, one of the tests currently recognised by the Australian government for skilled migration purposes.
2.2.3 Skilled Migration Policy reforms throughout the 1990s

By the late 1990s, the economic outlook in Australia was improving and the conservative Howard government, newly elected in 1996 and in office until 2007, was in the process of implementing a series of reforms to the migration program, including a significant expansion of the skilled migrant intake at the expense of family migration (Wright, 2014). Throughout the 1980s and early 1990s, family migrants had outnumbered skilled migrants, and accounted for the largest proportion of permanent migration. This trend reversed in 1998 as the number of skilled migrants exceeded those in the family category for the first time since 1945 (Castles et al., 2013; Wright, 2014). Since then, as mentioned in the introduction to this chapter, numbers in the skilled migration category have consistently represented the highest proportion of the overall migrant intake.

As the Howard government set about increasing the skilled migrant intake from the late 1990s onwards, they also began implementing reforms to the selection criteria in the points test for independent skilled migration. Within the skilled migration program, there are different visa categories. Although these categories shift and change over time, broadly speaking, a distinction can be drawn between visas that rely on the applicant being nominated for a skilled visa by, for example, an employer in Australia, and visas for which an applicant can apply without nomination, so-called independent skilled visas. Only applicants applying for independent skilled visas are selected using the points test, but prior to 2008-9, these applicants outnumbered those for nominated and other non-points tested visas (DIBP, 2015).

The reforms implemented by the Howard government included an expansion of English language criteria. By 1998, a number of occupations had been added to the ‘Occupations Requiring English’ list, such that 85 per cent of applicants were subjected to mandatory language requirements. In July 1999, further changes to the points test were implemented based on recommendations from a review of the skilled migration points test, conducted in 1997. These changes were aimed primarily at augmenting the skill base of the Australian labour force, and included the adoption of mandatory language requirements as core migrant selection criteria, which, from then on, applied regardless
of occupation (Birrell, et al., 2006; Chiswick & Miller, 2006). In addition, a list of occupations deemed to be in national shortage was introduced – the Migration Occupations in Demand List - and bonus points were granted to those with qualifications in the occupations listed.

According to Wright (2014), international migration poses a “control dilemma” for democratic states; they face pressure to appease public fears about unwanted migrants by implementing tighter controls on migration, and at the same time they face pressure to satisfy economic and labour market imperatives, which often call for more liberal controls on migrant inflows. Wright argues that a way out of this dilemma for policymakers is to make use of visibly restrictive measures, or what he terms ‘control signals’ to address popular concerns about ‘unwanted migrants’ and to encourage public confidence in a government’s ability to regulate the flow of desirable migrants. Wright suggests that reforms to the selection criteria for skilled migration in Australia, mentioned above, represented an effective use of ‘control signals’ by the Howard government; the entry of migrants perceived to be highly skilled and of significant economic value was encouraged, while the entry of those perceived likely to create a burden on the economy, including family migrants and those with imperfect English skills, was seen to be restricted.

In addition to expanding skilled migration in the late 1990s, the government also began prioritising temporary forms of migration, for work and study purposes (Koleth, 2010). While traditionally, skilled (and other) migration in Australia was conceptualised by governments and policy makers as a permanent phenomenon (Markus et al., 2009), new forms of temporary visas were established, as a means of enhancing the flexibility of the migration program. Temporary migration was thought to be more readily adaptable to changing economic conditions and shifting labour market needs than permanent skilled migration (Cully, 2011; Collins, 2013), and has since remained a significant part of the overall program. Cully (2011) distinguishes permanent and temporary migration as ‘supply-driven’ and ‘demand-driven’, respectively. The effort to increase the flexibility of Australia’s migration system, he points out, led to this prioritising of ‘demand-side’ over ‘supply-side’ migration in immigration policy.
The selection of skilled migrants for permanent settlement, as Cully (2011) explains, has traditionally been characterised as ‘supply-side’ migration: individuals self-nominate and are required to meet selection criteria stipulated in advance by the government, in the form of a points test. Upon meeting the criteria, these individuals qualify for an independent skilled visa, which is not contingent on an existing offer of employment. Permanent residency is granted on the basis of the points test, and the individual is not obliged to gain employment in order to maintain this status. ‘Demand-driven’ migration, by contrast, is the term used to describe the nomination and selection of migrants by employers. Employers are able to recruit workers from overseas according to the labour needs of their enterprises. In the case of temporary skilled migration, the number of visas is uncapped, and employers, rather than immigration officials, are responsible for selecting migrants where needed to fill specific job vacancies that cannot be otherwise filled due to local skill shortages. These migrants are granted an employer-sponsored temporary visa and are obliged to remain with the sponsoring employer for the duration of the visa period (Cully, 2011).

This shift from ‘supply-driven’ to ‘demand-driven’ skilled migration in the late 1990s was mirrored by government efforts to increase Australia’s share of the international education export market. Intense global competition to attract and recruit highly skilled workers was, in part, behind the drive to promote Australia as a study destination; attracting international students was viewed as a means of building an onshore supply of such workers. The introduction of the study-migration pathway in 1999, discussed in the section below, was a key policy measure in this context. Situated as a transition mechanism between temporary and permanent forms of migration, the pathway promised the possibility of permanent residency as an incentive to attract international students to Australian institutions.
2.3 The Study-Migration Pathway

2.3.1 Establishing the pathway (1999 to 2005)

By the late 1990s, tertiary education was becoming an increasingly important and lucrative export industry in Australia, and international competition for both highly skilled migrants and international students was intense (Koleth, 2010). The purpose of the study-migration policy initiative was thus twofold; firstly, to provide an incentive (in the form of permanent residency) to attract international students to Australian institutions, thereby boosting the value of Australia’s education export market. Secondly, and specifically in relation to skilled migration policy, the aim was to create an onshore supply of highly skilled and already acculturated migrant workers, as a means of efficiently augmenting the skill base and addressing skill shortages in the labour market (Gribble & Blackmore, 2012; Hawthorne, 2010; Koleth, 2010). According to Hawthorne (2010), international graduates were widely considered to be ‘ideal’ migrants; they possessed locally recognised qualifications, and were thought likely to have already adapted and assimilated themselves to the local linguistic and cultural context. As a result, in 1999, the Howard government modified existing skilled migration selection criteria by granting additional points to applicants possessing qualifications from Australian institutions. This marked the beginning of the establishment of a link between study and permanent residency via the skilled migration program.

In 2001, the link was reinforced and expanded as international graduates of Australian institutions with qualifications in specified key skill areas, such as information and communication technologies, for example, became immediately eligible for permanent residency as independent skilled migrants upon completion of their studies. Furthermore, whereas graduates wishing to apply for permanent residency had previously been required to leave Australia and apply from overseas, they were now able submit their applications onshore as long as they did so within six months of completing their studies (Koleth, 2010). In addition, the need to meet basic work experience criteria, a compulsory requirement for offshore applicants wishing to gain an independent skilled visa, was waived for those who applied within Australia after
having completed courses of at least 1 year in duration (Birrell, et al., 2006; Birrell & Healy, 2010).

International graduates of Australian institutions, even if their qualifications were not in key skill areas, were further advantaged compared to overseas applicants for permanent residency by other selection criteria in the points test. In addition to gaining bonus points for Australian qualifications, mentioned above, graduates from non-English speaking backgrounds were able to gain extra points for speaking a community language other than English (Birrell, et al., 2006). Furthermore, these graduates were not obliged to provide evidence of their English language proficiency. While applicants for skilled migration who had completed their qualifications overseas were required to demonstrate at least the minimum required level of English proficiency on a recognised language test, such as IELTS, those who had completed their studies onshore were, from 1999 until September 2007, assumed to possess adequate English skills and were thus exempt from any test requirement (Arkoudis et al., 2009).

The introduction of a study-migration pathway proved to be very effective in expanding Australia’s education export market, with the number of international student enrolments rising from 133,384 in 1999 to 344,815 in 2005 (Australian Education International, 2015). Similarly, the pathway appeared to be effectively producing a supply of onshore skilled migrants in Australia. By 2002, the increase in demand for permanent residency, particularly by overseas students in Australia, was such that the government lifted the points test ‘pass mark’, that is, the number of points needed for skilled migration, from 110 to 115 in order to prevent a potential over-supply of skilled migrants. In 2003, almost half of the independent skilled visas granted were to former student visa holders (Koleth, 2010). As Koleth (2010) reports, a further tightening of selection criteria ensued in July 2003 as the government again sought to limit the likelihood of exceeding annual visa quotas. The minimum period of study in Australia increased from one to two years before former students could access bonus points for Australian qualifications and have the basic work requirements waived. This measure was aimed at tailoring selection criteria to favour highly qualified former students, given the high demand, and was complemented by increases in the points awarded to
applicants with Australian higher degrees, such as Masters and PhDs. In light of continuing high demand, in 2005 the pass mark was again increased, from 115 to 120 (Koleth, 2010).

2.3.2 Changes to the pathway (2005 to 2008)

By 2005, the efficacy of the study-migration pathway in addressing skill shortages in the labour market was beginning to be questioned. Concerns had emerged that international graduates transitioning to skilled visas were failing to gain employment in their fields, despite persistent skill shortages (Koleth, 2010). The immigration minister at the time, Senator Amanda Vanstone, commissioned an evaluation of the general skilled migration program in mid-2005, to be undertaken by three researchers, Bob Birrell, Lesleyanne Hawthorne, and Sue Richardson. The minister had just announced a significant expansion of the skilled migration program for 2005-6, which was aimed at addressing skills shortages and concerns about the future growth of the labour force in light of a continuing below-replacement fertility rate.

Birrell, et al. (2006) examined how well skilled migration policy changes introduced since 1999 were achieving the general skilled migration program’s broad aims, described as skill augmentation, addressing shortages in the Australian labour force, and sustaining Australia’s labour force growth into the future. As with many studies into the efficacy of Australia’s migration program, the investigation was primarily quantitative in nature, based on an examination of data from migrant surveys, including the Longitudinal Survey of Immigrants to Australia (LSIA). As part of their evaluation, migrant attributes, including English language background and self-reported spoken English ability, were correlated with several economic measures of labour market integration. Their analyses revealed that skilled migrants from non-English speaking backgrounds, including international graduates recruited onshore, were experiencing prolonged difficulties accessing the Australian labour market. Furthermore, the researchers found that when these migrants gained employment, they were often underemployed, unable to access job opportunities and pay rates commensurate with their qualifications. Not surprisingly, they also found that migrants’ who reported speaking English ‘very well’
generally achieved rapid integration into the labour market, and were more likely to attain jobs and pay rates consistent with their skill levels.

Although Birrell et al.’s (2006) evaluation provided broad support for the effectiveness of policy changes since 1999 in terms of delivering improved labour market outcomes, they highlighted several problematic issues surrounding the study-migration pathway. Their evaluation of these particular issues is underscored by a premise that poor employment outcomes on the part of migrants from non-English speaking backgrounds are primarily attributable to a lack of English language proficiency. For example, the researchers argued against the efficacy of awarding bonus points to applicants who speak a community language other than English. This, they claimed, served to advantage graduate students who had completed their undergraduate degrees in non-English speaking countries, such as China, over those who had completed previous studies in countries with English as the language of education, such as India, thereby disadvantaging applicants more likely to have high levels of English.

Birrell, et al. (2006) also argued that bonus points attributed to applicants with qualifications listed on the Migration Occupations in Demand List Migrant (MODL) – up to 20 points, combined with bonus points for possessing Australian qualifications – up to 15 points, created perverse incentives for overseas students. As mentioned previously, the Howard government introduced the MODL in 1999 as part a series of reforms to better align the skilled migrant intake with existing labour market shortages. All applicants for points-tested permanent skilled migration, not only former overseas students, benefited from bonus points for possessing qualifications in occupations listed on the MODL, as well as priority processing. The three researchers claimed that the double-benefit available to applicants with qualifications on the MODL and from an Australian institution was likely to encourage those with the primary goal of gaining permanent residency to enrol in relatively undemanding courses in areas that featured on the list, particularly if they possessed limited English language proficiency. They further suggest that the incentive to enrol in courses that would offer MODL points was greatly increased by the raising of the points test pass mark in 2005 from 115 to 120, as many graduates of Australian institutions who had previously been able to meet the
pass mark without accessing MODL bonus points now needed these points. The benefits associated with pursuing qualifications associated with MODL points at an Australian institution, they suggest, was driving poor employment outcomes by producing an over-concentration of graduates in a limited number of occupational areas, with no guarantee that skilled migrants would possess good English communication skills.

As mentioned above, Birrell et al. (2006) supported their evaluation through an analysis of migrant survey data, whereby migrants were asked to provide a self-assessment of their own English proficiency by selecting the best option to describe how well they spoke English from the range: ‘very well’, ‘well’, ‘not well’ and ‘not at all’. Correlations with measures of labour market integration led to the conclusion that “in most dimensions of labour market success, the key is to have a level of English language competence that enables the respondent to report that they speak English at least ‘very well’…” (p. 92). Although Birrell et al. had already noted that there was “no way of aligning the answers of respondents to an objectively measured test of English proficiency such as IELTS” (p. 89), they nonetheless make the connection: “…We equate this to about a 7 on IELTS, although it is a big stretch to do so” (p. 92).

As already noted, many of the issues raised by Birrell et al. (2006) were based on the implicit premise that poor employment outcomes for skilled migrants can be attributed to poor English skills. This premise reflects media discourses emerging from 2001 onwards in which the presence of international students was consistently associated with declining academic standards (Devos, 2003, Birrell, 2006). Birrell (2006) drew on comparisons of IELTS scores achieved by the same test takers pre-entry to university and upon graduation to argue that there was little or no improvement in the English level of international students over the duration of their studies. This, he claimed, reflected declining academic standards as universities, reliant on income from international student fees, admitted international students despite their poor English language skills and accommodated for this ‘deficit’ through lowering the standards of assessment. Such a claim served to further undermine the study-migration pathway, as international graduates were characterised in negative terms, as lacking the language skills for employment.
As has been argued since, there is little empirical evidence to support this deficit perspective of the language skills of international students (see Benzie, 2010; Briguglio, 2011). Moreover, there is an increasing body of evidence suggesting that barriers to the integration of international graduates into the labour force are numerous and complex, and often unrelated to English language proficiency (Arkoudis et al., 2009; Berg, 2011; Gribble, 2014; Hawthorne, 2010; Hawthorne & To, 2014; Koleth, 2010; Robertson, Hoare & Harwood, 2011).

Nonetheless, the premise and the tenuous link Birrell et al. (2006) established between speaking English ‘very well’ and scores of IELTS 7 drove a series of reforms in 2007-8 to the selection criteria for skilled migration. These reforms included an increase in the English language requirements within the points test. In September 2007, the minimum language level for non-trade occupations was lifted from ‘vocational English’, defined then as IELTS 5, to the existing minimum level required - ‘competent English’, or IELTS 6. Scores of IELTS 6 attracted 15 points towards a pass mark of 120. In addition, significant bonus points were attributed to above-minimum scores of IELTS 7 – ‘proficient English’, 25 points. Furthermore, international graduates of Australian institutions were no longer assumed to possess the required level of English, and have since been subjected to the same language test requirements as offshore applicants (Hawthorne, 2011). As Hawthorne (2011) points out, these reforms meant that IELTS scores became the key determinant of selection for independent skilled visas, rather than qualifications in an occupation on the MODL.

Relevant work experience was also reintroduced in this period as a requirement for international graduates of Australian institutions, to address concerns that former overseas students transitioning to permanent skilled visas lacked work-readiness and employability. An 18-month temporary post-study work visa was introduced as a means of allowing recent graduates the opportunity to acquire the skilled work experience and English language skills needed to meet the new requirements (Phillips and Spinks, 2012).
2.3.3 Reforms from 2009 onwards - decoupling study and migration

From 2009 onwards, in the context of the Global Financial Crisis and amidst persisting concerns that the skilled migration program was failing to adequately address skill shortages, the Rudd Labour Government (elected in 2007) set about reforming the skilled migration program in an attempt to ensure that those selected were readily employable and able to fill existing demand (Gribble & Blackmore, 2012). In January 2009, higher priority processing was introduced for skilled migrants with pre-arranged and confirmed employment opportunities (employer nominated applicants) over applicants for independent skilled visas (Koleth, 2010), as a means of promoting rapid and efficient labour market integration outcomes. Previously, as noted in section 2.2.3 above, permanent skilled migration had been characterised as ‘supply-driven’, with the independent skilled visa category comprising the largest proportion of the permanent intake. This shift to prioritise selection of employer nominated migrants for permanent skilled migration thereby signalled a significant shift within the permanent migration program away from a ‘supply-driven’ towards a ‘demand-driven’ system (Phillips & Spinks, 2012).

A new Critical Skills List (CSL) was also introduced in January 2009. Applicants with skills featured on the new CSL, which included medical and IT professions, engineers and construction trades, received higher priority processing than those with skills on the MODL, in order to reflect shifting labour market priorities (Phillips & Spinks, 2012). In March, the government announced a reduction in the skilled migrant intake, and the CSL was amended to remove trade occupations. In May, a further reduction in the skilled migrant intake was announced, and in July 2009, a further tightening of English language test score requirements was implemented. Vocational English (IELTS 5) had remained sufficient for those in trade-related occupations until this point, when the requirement for these applicants was also increased to ‘competent English’ (IELTS 6).

During this period, some of the issues that had been raised by Birrell et al. (2006) were being compounded by wider concerns over the integrity of Australia’s export education industry and the student visa system, as well as the welfare of international students in
Australia. These concerns stemmed from increasing evidence of unethical and exploitative practices by some English language and international education providers, particularly in the context of rapidly increasing demand for courses that provided access to permanent residency, as well as the potential, due to inadequate regulation and quality assurance measures, for such providers to fraudulently align their courses with migration requirements (Hawthorne, 2011). Alarming reports of violent attacks against Indian international students in Australia provided further exacerbated welfare concerns. These factors, taken together, threatened to undermine Australia’s reputation as a study destination for international students (Gribble & Blackmore, 2012).

It had become increasingly apparent that the study-migration pathway was producing unforeseen and undesirable consequences. Thus, the government moved to assert a clear distinction between the purpose of the student visa program and the skilled migration program, which would effectively mean the decoupling of study and permanent residency (Koleth, 2010). The process of separating study and permanent migration included the announcement of significant reforms to the skilled migration selection criteria in 2010, including the introduction of a new, more targeted, Skilled Occupations List (SOL), and the implementation of a new points test in July 2011.

### 2.4 The New Points Test

This overhaul of selection processes reflected a marked preference for employer-sponsored migrants, with the bar significantly raised for independent skilled visa applicants; from mid-2011 onwards, eligibility for independent skilled visas was restricted to those with high-level education qualifications, proven work experience and advanced English ability (Gribble & Blackmore, 2012; Hawthorne & To, 2014). The key changes to selection criteria that characterised the shift from the 2010 points test to the new points test are first summarised below, and then to conclude this section and the chapter, the potential impact of the inclusion of highly restrictive IELTS scores within migrant selection criteria is discussed.
2.4.1 Key changes to selection criteria

At the introduction of the SOL in 2010, and prior to changes to the points test introduced in 2011, the pass mark for permanent skilled migration was set at 120 points across a range of criteria including occupational category, age, English language, Australian education, relevant employment, Australian employment, professional experience, partner skills, regional study, regional or state sponsorship and recognised skills in a community language. The SOL prioritised occupations according to labour market needs, with the top priority category attracting 60 points out of the 120 needed, and other high-demand occupations attracting 50 or 40 points, according to priority. In relation to English language ability, 15 points were attributed to scores of IELTS 6, the minimum requirement, and 25 points were attributed to scores of IELTS 7. Three years or more work experience in an occupation on the SOL offered 10 points for a 60-point occupation, and 5 points for other SOL occupations. The number of points awarded for this three years of work experience was the same for that gained overseas and in Australia. Australian work experience of at least one year was rewarded separately with a bonus of 10 points, regardless of occupation.

The new points test, which came into effect in July 2011, consisted of fewer criteria and, as mentioned above, privileged those with significant work experience, particularly experience gained in Australia, and included an increased emphasis on high level qualifications and English skills. The pass mark for the new points test was set at 65 when it was introduced, and remained at 65 for the period throughout which the data for this thesis were collected in 2012 and 2013. The key changes between the points test in effect in 2010 and the new points test introduced in 2011 are outlined below and summarised in table 2.1, further below.

In the new points test, points were no longer awarded according to occupational category, but possessing qualifications in a listed occupation remained as a prerequisite. In relation to work experience, one, three, five and eight years of experience gained in Australia attracted 5, 10, 15 and 20 points respectively out of 65, thus representing a significantly higher weighting than in the previous points test, and demonstrating a new
preference for extensive work experience on the part of skilled migrants. As a result, international students were unlikely to be able to access permanent skilled visas immediately upon completion of their qualifications in Australia. These graduates were able to apply for a temporary post-study work visa of up to 3 years, during which time they needed to obtain relevant employment experience and fulfil the English language test score requirement if they wished to become eligible to apply for permanent residency.

In terms of educational criteria, in the previous points test, 25 points towards a pass mark of 120 were awarded for Doctorate level qualifications from Australian institutions, with 15 for Bachelor and Masters level degrees and 5 for diplomas and trade qualifications. In the new test this changed significantly; points were still awarded according to qualification level but were more heavily weighted - 20 points towards a pass mark of 65 were awarded for Doctorate qualifications, 15 for a Bachelor or above, and 10 for trade-level qualifications. More importantly, though, in relation to the study-migration pathway, in the new points test qualifications gained overseas attracted the same points as those gained at Australian institutions. In the previous test, only Australian qualifications had attracted points. Consequently, international graduates of Australian institutions were no longer favoured by educational selection criteria for permanent skilled migration.

In terms of English language, the pre-2011 minimum requirement of IELTS 6, ‘competent’ English, remained in place in the new points test, but points were no longer awarded for meeting this requirement. Out of the 65 points needed, scores of IELTS 7 attracted 10 points, a slightly lower weighting than in the previous points test, but an extra discrete point item - ‘superior’ English - defined as IELTS 8, was added for 20 points, accounting for almost a third of the overall requirement. Achieving points for English language skills through scores of IELTS 7 or IELTS 8 became a necessity for many international graduates wishing to obtain permanent residency in an increasingly competitive selection process, especially recent graduates yet to accumulate extensive work experience in their professions.
## Table 2.1. Comparison of key criteria in 2010 and 2011 Points Tests

<table>
<thead>
<tr>
<th>Points test 2010 – pass mark 120</th>
<th>New points test – pass mark 65</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criteria</strong></td>
<td><strong>Criteria</strong></td>
</tr>
<tr>
<td><strong>Points</strong></td>
<td><strong>Points</strong></td>
</tr>
<tr>
<td>Occupational (SOL)</td>
<td>N/A</td>
</tr>
<tr>
<td>60, 50, 40</td>
<td></td>
</tr>
<tr>
<td>English language:</td>
<td></td>
</tr>
<tr>
<td>IELTS 7</td>
<td>IELTS 8</td>
</tr>
<tr>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>IELTS 6*</td>
<td>IELTS 7</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>IELTS 6*</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Education:</td>
<td></td>
</tr>
<tr>
<td>Australian Doctorate</td>
<td>Doctorate (Australia or overseas)</td>
</tr>
<tr>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Australian Bachelor or above</td>
<td>Bachelor or above (Australia or overseas)</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Australian Trade Qualification</td>
<td>Trade Qualification (Australia or overseas)</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Employment:</td>
<td>Employment#</td>
</tr>
<tr>
<td>60-point occupation + 3 years</td>
<td>1 year in Australia</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>other SOL + 3 years</td>
<td>3 years in Australia</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Closely related in Australia (1 year)</td>
<td>5 years in Australia</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Employment in Australia:</td>
<td></td>
</tr>
<tr>
<td>One year relevant work experience</td>
<td>8 years in Australia</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>3 years overseas</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5 years overseas</td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>8 years overseas</td>
</tr>
<tr>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

*minimum English test score required for eligibility

#Employment in nominated occupation or closely related undertaken in past 10 years (maximum points that can be awarded for any combination of Australian or overseas skilled employment is 20 points).

Since mid-2012, scores of IELTS 7 and 8 have become even more critical in the transition from temporary to permanent resident status via the skilled migration program, regardless of work experience. At this time, priority processing was introduced for employer and state sponsored visa applicants. This means that in a given selection round
(there are six per year), sponsored applicants are processed first and if quotas are not filled, independent skilled visa (non-sponsored) applicants are ranked in point order, with priority given to those with the highest points score (Department of Immigration & Citizenship, 2012). An implication of this is that an applicant's visa status may remain unresolved for prolonged periods, potentially indefinitely, depending on the number of sponsored applicants and on the points totals attained by other independent applicants.

2.4.2 Potential impacts of policy changes

In this unstable policy domain, the relative importance of achieving a certain number of points for English language shifts dynamically and arbitrarily, depending on the number and profile of other applicants. Furthermore, the English language test scores required to access a permanent skilled visa may be different for applicants within a single profession, even if they have similar levels of education and work experience; those with employer sponsorship only need IELTS 6, whereas independent applicants may need to achieve IELTS scores of 7 or even 8 to achieve the points they need for selection, depending on how much of their work experience was gained in Australia. Similarly, the IELTS score required by an individual applicant may shift over time, as the government makes regular adjustments to the minimum points total needed to become eligible to apply for permanent residency.

Shifting score requirements thus represent a source of uncertainty for migrants seeking to transition from temporary to permanent resident status in Australia. Reporting on a study that explored what ‘being temporary’ meant to 20 individual migrants on temporary visas in Australia, Robertson (2016) argues that in creating uncertainty over access to permanent residency in this way, the policy itself “hinders meaningful engagement with study and work, and often forces migrants to renegotiate their goals in complex ways” (p. 64), thereby undermining government intentions to promote skill augmentation and labour market integration. Robertson supports her argument by highlighting a case of a former Master of Accounting graduate from China, Liu, who had made repeated attempts at gaining the IELTS scores she needed for permanent residency, without success. With her temporary visa soon to expire, Liu felt forced to
resign from a position at an accounting firm she had obtained after graduating and to instead take up evening hospitality work, so that she could dedicate her time during the day to preparing for IELTS.

Dickie (2016) describes the Australian immigration policy space as “crowded with unintended consequences” (p. 2), due to its highly politicised nature and the influence of public opinion and social media commentary on policy directions. She points out that “this often leads to outcomes that have the opposite outcome of the original intention, prompting government to implement corrections that, in turn, have unexpected impacts” (p. 2). Berg (2011) also argues that the use of high English test scores for skilled migration selection undermines the purpose of the policy, which is ostensibly aimed at ensuring labour market integration. She claims that the use of scores of IELTS 7 and 8 for migrant selection purposes delivers an “implicit homogenising message that Australia wants migrants who sound like us and speak our language” (p. 110). She thereby points out the contradictory nature of the recent policy changes, asserting that progressively raising English requirements potentially reinforces “an underlying nativism in Australia” (p. 110), supporting racist attitudes and behaviours, and further marginalising migrants who are linguistically unassimilated.

Neither the potential for the use of language test scores to give rise to contradictory effects, nor the inherent instability of test scores, are specific to this particular policy domain. Rather, as will be argued in depth in the next chapter, these characteristics of test consequences and scores are inherent to testing practices in policy contexts, and thus present a fundamental challenge to current conceptualisations of test validity and impact in the field of language testing.
2.5 Conclusion

This chapter provided an overview of Australia’s skilled migration policy, the specific policy context within which the current study is situated. As demonstrated, increasingly restrictive English test score requirements have accompanied a suite of other changes to migrant selection criteria over recent years, in response to concerns that international graduates from non-English speaking backgrounds experience poor labour market outcomes, among other things. The culmination of these changes was the introduction in 2011 of a new points test for skilled migrant selection, within which scores of IELTS 7 and 8 are heavily weighted, and the introduction of priority processing, effectively limiting the number of independent permanent visas allocated. As mentioned in the introductory chapter, the four migrants involved in the current study were all affected by these changes.

Even though there is little controversy surrounding the idea that migrants need English skills to access skilled employment opportunities in Australia, the appropriateness of using English language test scores as a means of selecting migrants requires substantial interrogation. Existing research, as shown in this chapter, undermines the assumption that poor labour market outcomes can be attributed to a lack of English skills on the part of migrants, and suggests that increasing English test score requirements is an inadequate policy response, at best. Furthermore, as noted in relation to the new points test, the actual English test score requirement faced by individual migrants is contingent on how they fare across other, non-language criteria, as well as on the profile and status (sponsored or not) of other applicants, which undermines any meaningful relationship between test scores and the level of English language required to effectively work in Australia, as it might be imagined by language testers.

The next two chapters examine the theoretical challenges raised by these and other issues related to the use of language tests as instruments of immigration policies, particularly concerning how questions of test impact (chapter 3) and fairness (chapter 4) might be addressed. As will be discussed in detail in the next chapter, evaluations of
validity and impact in the field of language testing are currently contingent on a clear and definitive articulation of intended test purpose, typically from the perspective of the test developer, which precludes consideration of the contestability of intentions and the arbitrariness of score uses that inevitably occur in policy spaces. Furthermore, the view of test takers implicit in theoretical and empirical work ignores the role of individual agency, however constrained, in shaping the function and impact of tests in policy domains. Across chapters 3 and 4, a focus on the perspectives of individual stakeholders, particularly the often absent voices of test takers, is presented as offering a way forward in meeting the challenge of accounting for the social and political dimensions of tests.
CHAPTER 3: LANGUAGE TESTS IN POLICY SPACES:
CHALLENGING EXISTING THEORY

3.1 Introduction

This chapter examines the theoretical challenges associated with identifying and evaluating test consequences when tests are inserted into contested policy spaces. It is argued here that the use of language tests to regulate access to immigration rights raises critical questions concerning the adequacy and appropriateness of existing theoretical paradigms in language testing, particularly in relation to the way in which test purpose is conceptualised. The location of language tests within policy spaces, it will be claimed, necessitates the reconceptualization of tests and testing practices as politically situated and contingent, as sites of struggle in which score meanings and the values underpinning them are multiple and dynamic, and outcomes and consequences are fluid, unpredictable and multidirectional.

An overview of current validation frameworks will show that efforts in the field of language testing to account for the role and impact of tests in policy contexts have been hindered by what is termed in this thesis a ‘condition of determinability’. This ‘condition of determinability’ refers to the premises that intended test purpose and score meanings are fixed and stable, as defined by test developers, and that these intentions must be specified as a precursor to any evaluations of validity or consequences. Furthermore, within existing frameworks, test takers exist primarily as theoretical abstractions rather than as real persons, deconstructed into the components of knowledge and skills that constitute test constructs. Individual attributes, perceptions, feelings or behaviours that fall outside what is defined as ‘language ability’ in theoretical models are considered to represent potential sources of construct irrelevant variance, and efforts are made to reduce the impact of these factors in test design (Bachman & Palmer, 2010; O’Sullivan & Green, 2011).

By contrast, in areas of research closely related to the interests of language testers, including second language acquisition and second language teaching and learning, there
have been moves away from a focus on abstracted and decontextualized individual
difference factors towards a more holistic view of learners (Ushioda, 2009). This has
been accompanied by a focus on the complex interrelationships between learner
subjectivities and language learning outcomes, both of which are viewed as situated,
dynamic and emergent, shaped in and through interactions between individuals and the
social worlds they both inhabit and constitute (Dornyei & Ryan, 2015; Kalaja, Barcelos,
Aro, & Ruohotie-Lyhty, 2015; Larsen-Freeman & Cameron, 2008; Mercer 2011a, 2011b,
2014).

As will be demonstrated in this chapter, existing research into the use of tests as policy
instruments similarly reveals that individual subjectivities play a significant part in
shaping the outcomes of testing practices. In educational contexts, several studies have
highlighted the role of individual stakeholders in negotiating score meanings and
influencing the consequences that arise once tests enter social and political domains
(Hardy, 2013; Evans & Hornberger, 2005; Menken, 2008; 2009; McNamara, 2011;
Shohamy, Inbar, & Solomon. 2010; Van Avermaet & Pulinx, 2010). Recognising that
language tests, even when they are not explicitly incorporated into policy, function as
de-facto policy instruments (Shohamy, 2006; Spolsky, 2004), McNamara and others
(McNamara, 2005; McNamara & Roever, 2006; McNamara & Ryan, 2011) have
emphasized the need to recognise the social and political values implicit in test
constructs, as well as to develop a more nuanced understanding of the social
consequences of test use. Shohamy (2001, 2006, 2009) has also repeatedly pointed out
that tests are used for purposes far broader than simply measuring knowledge, and that
the widespread acceptance of the legitimacy of tests means that potentially
discriminatory and illiberal policy agendas remain hidden.

Although such critical perspectives have created space in the field of language testing
for discussion of the social and political contexts within which tests function, and
foregrounded the importance of accounting for the broader societal consequences of
test use, existing theoretical frameworks, it is argued here, still preclude consideration
of tests and testing practices as multidirectional and contested political spaces. Further,
a close examination of the basis of Shohamy’s argument will reveal that although her
problematizing of intended test use suggests a way of understanding the role of language tests as social practice, her view suggests that the intentions of the ‘powerful’ in society will inevitably prevail, and the consequences for test takers will be felt directly as enacted by policy. As a result, test purpose and consequences remain singular and determinate, only this time conceptualised from the perspective of the impact on test takers of the intentions of powerful institutions rather than from the perspective of test developers.

In a similar way, as discussed in the introductory chapter, recent criticisms of the use of language tests in immigration and citizenship policy contexts have emphasised the gatekeeping role tests play in these contexts, and tend to focus on exclusion as the primary, if not singular purpose of such testing (Blackledge, 2009; Horner, 2009; Shohamy, 2009; Van Avermaet, 2009). However, while empirical research on the impact of language tests on migrants seeking permanent residency or citizenship is as yet limited, investigations thus far have shown that the perceptions and experiences of individuals negotiating immigration policy in different countries are complex and varied, with tests functioning in some ways to exclude and marginalise, but in other ways to facilitate a sense of inclusion and belonging (Cooke, 2009; Khan, 2013; Strik et al., 2010).

The location of language tests within policy, it is thus claimed in this chapter, necessitates a reconceptualization of test purpose and a recognition of the role of test takers in negotiating score meanings, the values underpinning them, and the consequences that emerge from test uses. An attempt will be made to draw on recent efforts to reconceptualise ‘language’ (Blackledge & Creese, 2013; Blommaert, 2010; Hall, Vitanova and Marchenkova, 2005; Malinowski & Kramsch, 2013; Pennycook, 2001) and to bring poststructuralist perspectives to bear on language testing in particular (McNamara, 2012a, 2012b), in order to shed the ‘condition of determinability’ that currently underscores theoretical developments in the field of language testing. It will be argued that a poststructuralist approach offers a means of reconceptualising test purpose and consequences as dynamic, negotiable, and indeterminable in a way that is more consistent with how tests function in policy domains.
The chapter will be organised as follows: firstly, an overview is provided of the ways in which ‘test purpose’, that is, intended test construct and use as well as intended test consequences, is generally articulated within current test validity theory in the field of language testing. Secondly, consideration will be given to Shohamy’s (2001) seminal work *The Power of Tests* and the insights her critical perspective offer for furthering understanding of the social and political dimensions of test use within policy contexts. Following on from this, an overview of literature on the use and social impact of high-stakes educational testing (not language testing *per se*) will be offered, as this work illustrates Shohamy’s (2001) claim that such high-stakes testing practices tend to ‘create’ policy and evidences the need for a more fluid and multidirectional conceptualisation of test consequences. Then, a review of Cooke’s (2009) investigation of the impact of language and citizenship testing on migrants in the United Kingdom will be presented, as this study centres on the policy domain and some of the issues most relevant to the current thesis. Finally, a review of recent literature promoting poststructuralist approaches to current dilemmas in language testing and applied linguistics will be offered, as a means of offering a way forward in theorizing the complex relationship between language testing practices and policy domains.

3.2 *Test purpose* in language testing

3.2.1 Intended test construct and test use

As noted in the introduction, in the field of language testing, and arguably educational measurement more broadly, test ‘purpose’ is typically assumed to be uncontested, conceptualised solely from the perspective of the test developers. As shown below, this is clearly evident in current models of test validation, including the argument-based approach (Chapelle, Enright, & Jamieson, 2010; Kane, 2006, 2012, 2013; Xi, 2008) and the assessment use argument (Bachman & Palmer, 2010), in which test purpose is necessarily defined *a priori* by test developers, and test consequences are imagined, investigated and evaluated in relation to the intended test purpose.
The argument-based framework of test validation was developed by Kane (1992, 2006) and associates (Kane, Crooks & Cohen, 1999). It was subsequently adapted by Xi (2008) for language testing, and again by Chapelle et al. (2010) specifically in relation to the Test of English as a Foreign Language (TOEFL). Broadly, it involves two types of argument, the interpretive argument, and the validity argument. The interpretive argument consists of the explicit articulation of the way in which test scores are to be interpreted, in terms of a chain of inferences and underlying assumptions that link actual test performances to interpretations of score meaning and score-based decision making. The validity argument then involves gathering evidence in support of (and against) the chain of inferences and underlying assumptions that constitute the interpretive argument. Xi (2008) outlines the five key inferences involved in the interpretive argument as follows: an evaluation inference, linking test performance to observed score; a generalization inference, linking observed score to universe score; explanation and extrapolation inferences, linking universe score to interpretation of score meaning; and a utilization inference, where the interpretation of score meaning is linked to a score-based decision.

According to Kane (2012), the argument-based framework “invites assessment developers and users to specify their proposed interpretations and uses” (p. 4). In this way, the intentions of the test developers or test users define test purpose, in terms of the construct the test purports to measure and the way in which scores should be understood and used (the interpretive argument), as well as the sort of evidence required to evaluate the plausibility of the interpretive argument (the validity argument). In a more recent articulation of the argument-based approach to validation, Kane (2013) reiterates the centrality of a defined, bounded and singular notion of intended test purpose: “Test scores can have multiple possible interpretations/uses, and it is the proposed interpretation/use that is validated, not the test itself or the test scores” (p. 21).

Although Kane (2013) acknowledges that there is often disagreement concerning the value (positive or negative) that should be attributed to particular consequences arising from assessment, he argues that decisions based on score interpretations should be
evaluated by weighing negative and unintended consequences against the intended consequences, from the point of view of the test developer or user. While this appears to be a reasonable and ethical stance for language testers to take, it takes no account of the socially and politically situated nature of test development, and of the contingent and multidimensional nature of interpretations of score meanings and perceptions of test uses and consequences.

Reflecting on the issue of test uses and consequences that Kane’s approach raises, Chapelle (2011) summarises the implications for language testing in unproblematic terms: “test developers typically need to specify the intended decisions and consequences of test score use” (p. 26). This is not to say that the perspectives of test users and associated stakeholders are ignored. In relation to TOEFL, for example, Chapelle et al. (2010) explicitly acknowledge the importance of the views of relevant stakeholders through the ‘utilization’ inference, in which any argument for validity depends on evidence that test scores are useful for decision-making, in this case decision-making about university admissions. The authors explain that the inference is underscored by assumptions that “the meaning of test scores is clearly interpretable by admissions officers, test takers, and teachers” (p. 7). The meaning of test scores, and test purpose more broadly, is thereby assumed to pre-exist actual test use, and the way in which test scores are interpreted and used by stakeholders is subsequently evaluated in relation to the intentions specified by the test developers. Any mismatch between stakeholder perceptions and test developers’ intentions is assumed to represent either a validity issue with the test, in which case test developers make adjustments and improvements to the instrument, or a problem with the way stakeholders are interpreting and using scores, typically attributed to a lack of ‘assessment literacy’ on the part of test users (Pill & Harding, 2013; Taylor, 2009, 2013).

In policy domains, where language tests arguably serve multiple purposes, many of which may be unrelated or very distantly related to the purpose for which the test was originally designed, it is difficult to imagine how the test validation argument might look. IELTS, for example, was originally designed to assess the English language abilities of test takers wishing to enter university (IELTS Academic) or technical or business colleges
It is now used for a variety of other purposes, including professional registration and, of course, immigration purposes (Ahern, 2009). Fulcher and Davison (2009) propose the notion of ‘change retrofit’, drawing on architectural concepts as metaphors, to suggest how test developers might revise test design to ensure their instruments remain fit for purpose when put to new and previously unanticipated uses. However, this notion similarly rests on maintaining a single, unified, and unproblematic notion of test purpose. In any case, as Fulcher and Davidson point out, rather than making such efforts, testing companies tend to actively promote the acceptance of their tests for immigration and other such ‘unintended’ uses, deferring responsibility by not explicitly endorsing them.

In evaluating the use of language assessments, Bachman and Palmer (1996) emphasise the need to consider not only the technical qualities of testing instruments, but also the real world practicalities and constraints associated with test use, as well as the associated consequences. Their recent ‘assessment use argument’ (AUA) (2010) expands on the earlier ‘test usefulness’ framework (1996), by making justification for the use of language assessments a central component of test evaluation. Both frameworks demand that a language test must effectively serve the purpose for which it is designed (test purpose is here also thus conceptualised as unproblematic and singular), but the AUA makes the additional demand for a value-oriented evaluation of test use, in the form of a rationale which demonstrates that the use of an assessment will deliver beneficial effects to stakeholders. Bachman and Palmer (2010) attribute responsibility for test consequences to both test developers and decision makers, arguing that both groups are accountable to stakeholders and that both must therefore be able to appropriately justify test use, through reasoned argument and empirical evidence. The authors carefully acknowledge that consequences cannot be beneficial for all stakeholders; in particular, they highlight test takers who do not perform well on tests, clarifying that any negative consequences should be weighed against the intended beneficial effects of assessment use.

and Palmer’s (2010) model privileges test use, whereas the argument-based validation framework privileges score interpretation over test use. Bachman and Palmer’s treatment of consequences as a part of the overall justification for test use is also closely aligned with Messick’s (1989) framework, in which consequences are situated as an aspect of overall validity, as will be discussed in the next chapter. The underlying premise for Messick (1989), Kane (2006, 2012, 2013) and Bachman and Palmer (2010), as well as those who have adopted their frameworks in practical validation efforts in the field of Language Testing, is that tests are intended to bring about societal ‘benefits’.

This is, in and of itself, an inherently political position; equally so, language testing is an inherently political practice, although this is rarely acknowledged - test developers design tests with a particular, albeit often implicit, understanding of ‘beneficial’ consequences in mind, which rests on associated notions of ‘fairness’ and ‘social justice’ that are value-laden and often contested. As will be discussed further below under the heading ‘test impact in policy contexts’, even in educational settings where it might appear reasonable to evaluate consequences as beneficial (or not) in terms of effects on teaching and learning, and to assume consensus over what might constitute ‘beneficial’, perceptions can vary markedly, not just between policy makers and teachers, but between individual policy makers and between individual teachers.

Although Bachman and Palmer (2010) attempt to move beyond considering only the intended use of tests in order to address issues of fairness and social justice arising from actual tests uses, their AUA offers no space for considering ‘intentions’ as multiple and competing. Rather, the assumption is that test developers design assessment tools for a particular use, in collaboration with the test user. Further, little space is available for problematizing the social and political values underlying broad notions of societal ‘benefit’, ‘fairness’, and ‘social justice’. These notions have assumed an all-pervasive ‘naturalness’ in language testing theory. The AUA rests on the idea that there will be consensus (or at least that consensus is possible) among stakeholders about intended test use and intended consequences, which in turn rests on consensus over the meaning of notions such as ‘beneficial’, ‘fairness’ and ‘social justice’. In policy contexts, however, consensus seems to be far from the norm and a conceptualisation of test consequences
in terms of a beneficial versus harmful, positive versus negative dichotomy is increasingly difficult to sustain. In addition, the indeterminacy of consequences that characterises policy contexts necessarily and fundamentally challenges methods of evaluating validity and/or fairness claims that rest on determining whether or not a test delivers overall societal benefit.

In the case of immigration policy, for example, the complexity involved in identifying and evaluating ‘beneficial’ test consequences is compounded by the dynamic and shifting policy space within which tests function. As discussed in the previous chapter, on the one hand, English proficiency is a prerequisite for skilled migration in Australia because research has demonstrated strong links between speaking English ‘very well’ and positive employment outcomes (Birrell et al., 2006); thus, the test requirement. On the other hand, test scores are set by immigration officials, not by language testers, in a seemingly arbitrary way, shifting regularly in order to meet other, non-language related policy objectives, such as regulating migration numbers, or countering political criticism, or simply to appear to be ‘doing something’ in response to public concerns about immigration, for example (Dickie, 2016). As Shohamy (2001) highlights, there is a disconnect between test constructs as understood by language testers, conceived in terms of the particular aspects of language knowledge that a particular test measures, and the multiple, shifting and competing understandings of intended test use and score meanings on the part of test users and stakeholders.

3.2.2 Intended test use and test consequences

The underlying assumption that test purpose is unitary and definitive is further evident in the focus of most empirical investigations of test consequences, which have centred on the effects of language tests on teaching practices and learning outcomes. These investigations, commonly known as ‘washback’ studies, are premised on the belief that well-designed tests, high-stakes tests in particular, are instruments that are designed and introduced to facilitate positive improvements in educational practices (Cheng, 2008; Wall, 2013). Numerous ‘washback’ studies have been conducted into the effects of language testing on classroom and teaching practices, teaching materials and
textbooks, teacher perceptions and reactions to tests as well as effects on curriculum design and learning (e.g., Alderson & Wall, 1993; Cheng, 2005; Cheng, Watanabe, & Curtis, 2004; Hamp-Lyons, 1997; Wall & Anderson, 1993; see Cheng, 2008 for an overview). Investigations are typically retrospective, and serve as a post-hoc means of evaluating tests and testing processes according to the extent to which they lead to improvements in language education. ‘Test purpose’, therefore, is implicitly understood in general terms as delivering benefits to learners, teachers and the education system more broadly, and as a result, ‘consequences’ are typically conceptualised in dichotomous terms as either positive effects on teaching and learning (positive washback), or negative effects on teaching and learning (negative washback) (Wall, 2013).

While the educational effects of language tests have been widely investigated, very few studies have examined the broader social consequences, or ‘impact’, of language test use (Cheng, 2008; Shohamy, 2001; Wall, 2013). In very general terms, which encompass but extend beyond ‘washback’, Wall (1997) describes test impact as “any of the effects a test may have on individuals, policies or practices, within the classroom, the school, the education system, or society as a whole” (p. 291; cited again in Wall, 2013, p. 1; see also Hamp-Lyons, 1997 for a similar definition). Shohamy (2001), by contrast, focuses on the values underlying test use, thereby emphasising the socially and politically-situated nature of testing practices in her definition of societal consequences: “the effects of tests on aspects such as gate-keeping, ideology, ethicality, morality and fairness” (p. 46).

A recent investigation into the effects of the College English Test (CET) in China by Garner and Huang (2014) represents a substantial step forward in understanding, theorising and evidencing impact in language testing. Their view of impact is consistent with the definition suggested by Wall (1997, 2013), above, and also encompasses the political and social dimensions of test consequences articulated by Shohamy (2001). Garner and Huang (2014) adopt a problematized view of test purpose, noting from the outset that anecdotal evidence about the CET suggests multiple and diverse interpretations and uses of test scores, as well as conflicting claims about test
consequences. The authors begin by situating the introduction and use of the CET in the context of political debate in China concerning the quality of education and educational outcomes, particularly in relation to English language teaching and learning. Over 18 million people sat the CET in 2011, compared to just over 100,000 when it was first introduced in 1987. The dramatic increase is attributed by Garner and Huang to two main factors: “the enormous expansion of higher education and...the growing influence of the test as the test results are increasingly widely used as criteria for awarding degrees, jobs, and other social benefits” (p. 21).

Garner and Huang (2014) point out that the CET Committee specified the intended purpose of the test as furthering English language teaching and learning in tertiary education in China. They note, however, that there have been intense debates in China surrounding the extent to which the introduction of the test has achieved its intended purpose, and also concerning the wider social and educational consequences of the CET. Further questions have been raised in academic circles, regarding the legitimacy of the test results for decision making about access to employment and other social opportunities. Garner and Huang empirically investigate various competing perceptions of the CET and its consequences through an examination of how test results are used by various stakeholders and why they used in such ways, as a precursor to identifying test impact. They found that a range of different stakeholders were using CET results for various purposes, and that uses were sometimes at odds with the intended purpose of the test. For example, the education ministry made use of results to monitor and evaluate school practices and to implement educational policies. CET scores formed part of the job selection criteria used by many employers, and local government departments were found to use test scores for administrative, rather than educational purposes, requiring particular CET scores as a precondition for granting work or residency permits in some cities.

Garner and Huang (2014) found that the use of CET scores impacted test takers in various ways, with some students viewing the test explicitly and solely in terms of the use of scores by employers in job selection, while others saw it as a source of motivation for spending time studying English. In terms of perceptions of fairness, Garner and
Huang’s findings illustrate the significance of broader social values in conceptions of beneficial consequences. They found that some test takers felt that CET results represented a fair basis for job selection, in that allocation of opportunities according to test scores was perceived to be more equitable than selection on the basis of social connections and wealth (‘guanxi’). The authors explain that guanxi is a double-edged sword:

“while exploiting their own guanxi, people are afraid that others may use guanxi against them. This is one reason why Chinese people tend to believe that a public examination is the fairest means of selection and competition” (p. 142).

Claims of unfairness, Garner and Huang found, were based on the argument that inequalities exist in relation to access to job opportunities and English learning opportunities, thus some test takers felt disadvantaged by the universal use of CET scores.

Garner and Huang (2014) emphasise the connection between stakes, or the significance attached to test scores, and the impact the use of test scores is likely to have. They conceptualise impact as a complex set of interrelated factors, summarised according to four main categories: “The impact introducers – test users; the impact medium – the test; the impact receiver – e.g. students, teachers; and the impact context – e.g. economic, educational, political, social and cultural contexts” (p. 162). In the case of the CET, test users are identified as powerful institutions, such as governments, universities and employers. The test is thus conceptualised as an instrument of control, in similar terms to those described by Shohamy (2001), discussed further below. The role of impact receivers, particular test takers, is considered according to Bourdieu’s (1989) notions of social and cultural capital. According to Garner and Huang (2014), the CET represents a site of struggle over capital, and the impact on individual receivers will be contingent on their existing capital, as well as their desire for gains in capital, which test scores can potentially satisfy, depending on the stakes attached by particular test users. While this represents a useful and significant advance in theorizing impact, the notion
of ‘impact medium’ needs to be further problematized to reflect the potential for multiple interpretations of scores meanings, even within a single instance of test use; test users, test takers and other associated stakeholders act and react according to their perceptions of what scores mean and how they are likely to be used. In this way, as Garner and Huang’s own findings suggest, stakeholders negotiate impact in dynamic and unpredictable ways. In turn, a more nuanced understanding of ‘impact medium’ calls for further problematizing of the category of ‘impact receivers’, as test takers, teachers and associated stakeholders ‘create’ as much as ‘receive’ impact through their understandings of and reactions to different interpretations and uses of scores.

The difficulty of imagining and accounting for the intentions and perceptions of diverse stakeholders once a test enters the public space is reflected in discussions of responsibility among testing professionals. Issues of responsibility for test use and test consequences have centred on the extent to which test developers are able to exercise control and influence over testing processes. Codes of ethics and of practice have been developed by a number of language testing associations as a means of both ensuring that language testers meet professional standards, and protecting the profession from the misuse of their tests by external stakeholders (see Davies, 1997 & 2004). Misuse of tests is associated with ‘unintended consequences’, or consequences arising from the use of tests for purposes or in contexts beyond those for which they were designed and validated (Kane, 2006, 2012, 2013). Unintended consequences are typically characterised as negative, beyond the control of the test developer, and contrast with the positive consequences that are assumed to arise from a well-developed and validated test put to its intended use. As McNamara and Ryan (2011) point out, test developers have little, if any, influence on the policy domains within which language tests play a central role. It is therefore not surprising that discussions about test use and consequences have focussed predominately on aspects of testing that fall unambiguously within the domain and control of language testers.

The distinction between the ‘intended’ and ‘unintended’ uses and consequences of tests becomes almost impossible to uphold, however, in cases where large and influential companies such as IELTS, ETS and Pearson, among others, actively bid to have their
English language tests accepted and used as instruments of immigration policy, as is the case in Australia and in other English-dominant countries, such as the United Kingdom, Canada and New Zealand, for example. Recent efforts to develop theoretical tools with which to account for the social and political dimensions of language test use in these controversial policy contexts, and also to evaluate the broader societal consequences of such test use, such as Kunnan’s (2004) ‘fairness framework’, Bachman and Palmer’s (2010) ‘assessment use argument’, and Xi’s (2010) ‘fairness argument’, perhaps reflect an emerging crisis of awareness within the field of language testing, as language testers are increasingly identified, within and outside of the field, as playing a complicit role in furthering contested political agendas. These contributions will be discussed in detail in the next chapter. Before moving to a review of research on test impact in policy contexts, Shohamy’s notion of ‘the power of tests’, through which she highlights the political agendas that underlie all test use, will be examined below.

3.3 “The Power of Tests” (Shohamy, 2001)

Shohamy (1997, 2001, 2006) first drew attention to the significance of the political agendas that underlie the uses of tests two decades ago, highlighting the fact that test developers and test users often have very different agendas, and therefore very different conceptions of test purpose. She makes a distinction between the overt and covert intentions behind test use, whereby ‘overt’ intentions are those that are openly declared by test users and made publicly available through, for example, formal documents, and policy statements. ‘Covert’ intentions, on the other hand, refer to hidden agendas, which are not made publicly available and “can only be inferred and deduced from other sources” (Shohamy, 2001, p. 45).

Shohamy (2001) emphasises the importance of bringing the covert intentions that underscore the uses of tests to light, in order to properly evaluate the appropriateness of testing practices and consequences. She argues that tests offer policymakers a powerful and flexible means of pushing political agendas that are often far removed from the purpose intended by test developers, such as controlling quotas, shaping education agendas, and defining and controlling knowledge. For example, Shohamy
claims, 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. She draws on Foucault (1979) to illustrate that tests are imbued with a punitive authority, in that failure to properly conform to the norms prescribed by test constructs can lead to sanctions and exclusion. In addition, she claims that those who are excluded by tests accept and support their power and authority, because tests are seen to offer objective and factual accounts of the knowledge or skills possessed by test takers. Shohamy further attributes the acceptance and defence of testing practices on the part of those disadvantaged by tests to a sort of ideological manipulation by the power holders in society, through which they are able to reinforce existing hierarchies and perpetuate existing inequalities:

“...even low-status groups, minorities and immigrants who are not part of the elite, and who are constantly excluded and rejected by tests, also have an overwhelming respect for tests and often fight against their abandonment. Their behaviour is similar to that of victims believing they are unworthy, mostly as a result of the effective propaganda that turns tests from symbols to ideologies” (p. 120).

She links this to Bourdieu’s (1991) notion of ‘symbolic power’:

“The public becomes dependent and hooked on tests, developing unchallenged trust in their results and in their symbolic power. From an early age people are socialized to believe that tests are the prime indicators of their societal values” (p. 122).

Tests, according to this view, are clearly not neutral and objective measurement devices. Rather, they are extremely powerful tools that serve particular interests and reinforce existing values, ideologies and hierarchies. According to Shohamy (2001), the power of tests derives from two main components. The first component is a source of power asymmetry between test takers and those who introduce tests, because the test embodies the threat of negative sanctions, both material and symbolic, and test takers,
educators and other affected stakeholders inevitably fear such sanctions. Consequently, they change their behaviour in order to align their ‘knowledge’ or the knowledge they are teaching with the institutionalised ‘knowledge’ embodied in the test. The second component relates to the ways in which the legitimacy of test and test scores has been established as all pervasive and unquestionable, thereby silencing dissenting voices.

Shohamy (2001) suggests that the effects of testing on individual test takers and other associated stakeholders, such as parents and teachers, raise moral and ethical questions that reach well beyond the sorts of validity issues upon which investigations in the field of language testing have typically focused. She argues that the coercive power of tests, or in her words, the ability of tests “to cause a change in behaviour” (p. 105), is utilised by policymakers, bureaucrats and other decision makers to control and redefine knowledge and demand conformity in line with their own agendas. Such de-facto policy making, she claims, is unethical and undemocratic; unethical because tests do not lead to improved educational outcomes. Rather, she argues, tests create “parallel systems in which tests become the de facto knowledge – often in contradiction to more sophisticated systems” (p. 116). Undemocratic because tests serve the interests of those in power; test takers and teachers bear all the costs and little, if any benefit. They are deprived of choice and agency, and have no real means of resisting the imposition of a test, however much they may oppose the intentions behind its use or challenge its quality.

As shown in the previous section, Shohamy’s insights into test use, particularly the notion that there may be several competing definitions of test constructs, intended test uses and intended consequences, have yet to be effectively translated into theoretical advances in language testing. Furthermore, our understanding of language testing as social practice, particularly the ways in which social beings are categorised and organised in and through tests, especially salient in policy contexts, or how consequences might be reconceptualised to allow for evaluations of fairness and justice claims that inevitably arise in contested policy spaces, remains extremely limited.
Reluctance to engage deeply with Shohamy’s argument may stem from the fact that language testing as a discipline has traditionally been oriented towards issues of procedural fairness, with an associated focus on the technical, rather than ideological, moral and philosophical aspects of testing (McNamara, 2013). Furthermore, Shohamy’s view is inconsistent with the view that language tests are instruments that promote societal benefit, which, as already discussed, underscores validity theory in the field. Tests, according to Shohamy, reinforce existing power hierarchies and inequalities through both coercive means - test scores carry punitive sanctions; and consensual means – tests are instruments of ideological control, defining knowledge and creating and reinforcing categories of belonging and exclusion. She urges resistance on the part of language testers to what she sees as an inherently unjust practice, advocating overturning existing testing procedures in favour of what she describes as democratic assessment practices, whereby power is redistributed and shared between educational authorities and learners, rather than resting solely with testers. Tests as they currently exist, according to this view, are fundamentally malevolent instruments, rather than benevolent instruments.

While Shohamy’s critical perspective effectively problematizes many of the key underlying assumptions in language testing theory, her view, as mentioned above in the introduction, leaves little space for test takers, educators and other less powerful stakeholders to play a role in shaping the meaning of test constructs or the consequences and policy outcomes associated with the use of test scores, nor of escaping the construction of impact according to a beneficial versus harmful dichotomy. By contrast, test (and policy) purpose is seen to be determined exclusively by powerful test users, such as governments in the case of testing for immigration purposes, with individual stakeholders seen to experience the test (and policy) in the way intended by test users. As discussed in the previous section, Garner and Huang’s recent model of test impact is consistent with this view, as test takers and other related stakeholders are conceptualised as ‘impact receivers’. By contrast, it is argued in this thesis that impact is not solely a one-way, top down process that can be categorised definitively as either positive or negative. Although institutional powers inevitably play a significant role in determining the range of possibilities available to individual stakeholders, particularly
test takers, the ways in which these individuals interact with the test and the policy within which it is embedded will shape, albeit subtly, score meanings, test consequences, and policy enactments in dynamic, multidirectional and unpredictable ways.

In the next section, ‘Test impact in policy contexts’ a brief overview of investigations of the broader social consequences of high stakes educational (not necessarily language) testing regimes will first be offered, followed by a review of Cooke’s (2009) investigation of the impact of language and citizenship testing on migrants in the United Kingdom. As will be shown, these studies illustrate that tests often function in unintended ways that are not easily foreseen by test developers or policy makers, with test ‘purpose’ serving various different and competing agendas. Existing research also supports the view that test takers, teachers and other associated stakeholders, in addition to governments and policy makers, play a role in constructing the meaning of test scores and the consequences associated with the uses of tests in multifaceted and sometimes conflicting ways. As argued above, these stakeholders ‘create’ as much as ‘receive’ impact through their perceptions of and orientation to different interpretations and uses of scores, as well as their beliefs and values about the appropriateness and the fairness of the test and the policy.

3.4 Test impact in policy contexts

3.4.1 Tests and education policy

Notwithstanding Garner and Huang’s (2014) investigation, discussed above in section 3.2.2, the number of investigations into the broader social consequences of language testing per se remains, as yet, limited. Existing literature on the use and impact of high-stakes tests in policy contexts has mainly focussed on the impact of educational testing. In particular, attention has centred on the ways in which high-stakes educational testing has produced de-facto language policy, as in the case of No Child Left Behind (NCLB) in the United States (for example, Chalhoub-Deville, 2009; Evans & Hornberger, 2005; Menken, 2008; 2009). There is also an emerging literature examining how international
testing regimes involved in the Programme for International Student Assessment (PISA) have influenced domestic education policy formulation and led to unforeseen consequences for migrants and other minority language speakers in participating nations (McNamara, 2011; Shohamy et al., 2010; Van Avermaet & Pulinx, 2010). In the Australian context, an investigation into the impact of the National Assessment Program – Literacy and Numeracy (NAPLAN) on teachers and teaching practices has shown that the meaning of test scores can be interpreted in diverse ways, leading to varied and sometimes conflicting policy enactments at the micro level by different stakeholders (Hardy, 2013).

In the context of educational policy, the use of high-stakes standardized tests have often been associated with the creation and perpetuation of what have been called ‘discourses of deficit’ (Candlin & Crichton, 2011), whereby non-dominant languages and minority speakers are situated in opposition to achieving national goals of improved educational outcomes, as well as social and economic advancement more broadly (McNamara, 2011). As will be discussed below, researchers investigating the effects of NCLB, PISA and NAPLAN have linked such discourses to testing practices in various national contexts, and have emphasized the negative consequences that have emerged in these policy contexts, particularly for linguistic minority groups. Investigations of policy effects at the micro level, specifically on educators and teaching practices, suggest, however, that although high stakes testing regimes have led to negative and undesirable effects in many respects, the diverse ways in which individual stakeholders appropriate, modify, and resist policy directions and test purpose as they perceive them can lead to complex and sometimes conflicting test consequences and policy outcomes.

NCLB, for example, was implemented as federal education policy in the United States in 2001. It has since been superseded by the Race to the Top policy, introduced by the Obama government in 2009, which maintains an emphasis on standardized testing and similarly links test outcomes to school funding1. One of the policy intentions of NCLB (and Race to the Top), was to close the achievement gap between minority and

1 https://www.ed.gov/category/program/race-top-fund
disadvantaged students, and high performing students from non-minority backgrounds. Menken (2009) notes that the initiative emerged in the context of increasing public and political opinion that teaching and learning quality was poor, with schools failing to deliver on government investment in education. Under NCLB, standardized tests, administered in English, were introduced as a means of measuring and comparing learning outcomes across schools; thereby delivering increased accountability. Schools faced heavy sanctions, such as loss of funding or closure, if their students did not demonstrate adequate progress and achievement as measured by these tests (Menken, 2009; Shohamy, 2008).

Until recently, a significant aspect of NCLB was that students from non-English language backgrounds were required to take and pass the same tests as native-English speaking students. As a consequence, several writers have drawn attention to the ways in which NCLB created de facto language policy, by undermining bilingual education practices in the United States and encouraging instead English-only education as a means of facilitating English acquisition on the part of minority language speakers (Evans & Hornberger, 2005; Wiley & Wright, 2004; Menken, 2008; Menken, 2009). Evans and Hornberger (2005), for example, argue that despite the viability of a bilingual approach to education, the shift to NCLB reflects a shift away from a view of minority languages as resources towards a “language as problem orientation” (p. 92). In other words, towards a deficit view of the development of languages other than English, in which such development is seen to hinder rather than facilitate the acquisition of English as well as learning in general, in turn leading to low levels of achievement among students for whom English is not the first language. Furthermore, the deficit view was reinforced and perpetuated by test outcomes, as national studies repeatedly showed that students from non-English speaking backgrounds were achieving significantly lower test scores than native-English speaking students, even on subjects that would be considered less language-oriented, such as math (see Menken, 2009, for an overview).

As an example of empirical research into the impact of NCLB on minority students, Menken (2008) conducted a qualitative study of the effects of NCLB testing across ten high schools in New York City and found that minority students felt pressure to learn
English quickly, and suffered from low self-esteem when they were unable to do so. She also found that schools and teachers typically interpreted test outcomes in line with a linguistic deficit perspective, viewing poor results as evidence that minority students were ‘problematic’, leading to a push for some students to undertake extended hours of test preparation or to be prevented from progressing through school grades. In some cases, she reports, students deemed unlikely to pass were encouraged by schools to leave the education system and return to their home country. Despite such overwhelmingly negative impacts, and the pervasiveness of the deficit perspective, Menken also found that individual educators differed in their approach to minority students, with some resisting the policy push towards English-only teaching, instead increasing teaching in students’ native language. She concluded that teachers’ own beliefs concerning language diversity and the effectiveness of bilingual education influenced the way in which the policy, which may have been oriented around monolingual ideals, was realised and manipulated.

In a similar way to NCLB, except at a transnational policy level, the development of PISA was also linked to broad goals of improving educational outcomes by improving accountability through standardized testing practices. Commissioned by the OECD and introduced in 2000, PISA offers participating countries a means of comparing and evaluating the impact of their own educational policies against those of other participating countries, through comparisons of student performances on the tests. In turn, these comparative evaluations should lead to policy improvements at national levels as poor performing countries look to other, more successful nations for policy initiatives. PISA tests are administered every three years to 15-year-old students in selected schools across several countries, with each triennial administration focussed on three subject areas, defined as key areas by the OECD: reading, mathematics, and science. The tests are not linked to educational curricula; rather, they are designed to measure a more generic set of skills and knowledge, specifically: “to assess to what extent students at the end of compulsory education, can apply their knowledge to real-life situations and be equipped for full participation in society” (www.oecd.org/pisa/aboutpisa/).
McNamara (2011) highlights the way in which the simultaneous release of PISA results for all countries has led to the creation of ‘league tables’, in which ‘winner’ and ‘loser’ nations are defined on the basis of test outcomes. This, he suggests, inevitably gives rise to “a ‘discourse of deficit about the ‘losers’” (p. 311), which in turn leads to the creation or reinforcement of other discourses of deficit, as poor national performances are often attributed to a lack of competence on the part of bilingual or minority language speaking test takers, particularly immigrant children. Through an examination of the impact of PISA assessments of reading in Austria, McNamara reveals that while test results could be linked to discourses of deficit, particularly in relation to migrant groups, the existence of PISA also had positive impacts at the school level, in terms of rendering the role of language in education visible and salient, for example. In this way, he argues, “measurement, even where it has the potential to reinforce discourses of deficit, does not play a single role, but is in a sense inherently indeterminate in its effects, which will be both negative and positive” (p. 312).

The reasons behind the introduction of NAPLAN in Australia in 2008 mirror in some ways the impetus for NCLB and subsequent policy in the United States, as does the controversy surrounding it. The testing regime was implemented to promote improvements in educational outcomes by ensuring greater accountability and a more centralised approach to education through national literacy and numeracy benchmarks. National reporting in the form of comparative tables of test score outcomes across schools was intended to increase pressure on schools and teachers to maintain quality and ensure their practices led to student success (Cumming, Kimber & Wyatt-Smith, 2011). NAPLAN has been widely criticised as counterproductive, however, with claims that the testing regime has in fact led to a narrowing of curricula and pedagogy through ‘teaching to the test’, as well as increased workloads for educators and negative effects on the personal wellbeing of students (Dulfer, Polesel & Rice, 2012).

Hardy (2013) investigated the impact of NAPLAN on teachers’ understandings of their practice and student learning in the context of one Australian school. He found that teachers varied in their response to the testing regime. While some oriented their teaching to the test requirements, many took an ambivalent stance towards the test.
On the one hand, Hardy explains, teachers viewed NAPLAN score data as evidence of student learning, particularly where aspects of the test were considered useful for informing teaching practices. On the other hand, they simultaneously critiqued the conflation of statistical data and learning practices and sought what they saw as more appropriate forms of evidence, collected within school, such as samples of student work or expert opinions from other teachers. He situates this tension within a framework proposed by Saetnan, Lomell and Hammer (2011), in which conflicting forces, described as ‘centripetal’ and ‘centrifugal’, shape teaching and learning practices in complex ways. Centripetal forces are those pushing educational practices towards standardisation, into fixed and uniform categories, whereas centrifugal forces are those “which problematize such categorisation and standardisation processes” and involve efforts to “reorient particular practices which give greater value to the local, the situated” (Hardy, 2013, p. 69).

Rather than focusing on categorising the impact of NAPLAN in terms of positive or negative effects on educational outcomes, Hardy (2013) emphasises the “co-constructed nature of statistics and students’ learning” which emerges from the tension between these opposing centripetal and centrifugal forces (p. 67). He argues for the value of statistical data as providing a means of simplifying the world, “thereby enabling a greater degree of control than would otherwise be the case” (p. 69), while at the same time acknowledging that such simplification necessarily abstracts from the local, the specific. Drawing on Saetnan, Lomell and Hammer (2011), he asserts that standardisation processes, or pressures towards simplification, were both supported and simultaneously challenged by teachers involved in his study. The teachers, he claims, were influenced by centring forces and the social value of statistics, but were at the same time “agents for centrifugal forces which challenged and contested this neat simplification process, and which sought simultaneously to dismantle such statist projects and processes” (Hardy, 2013, p. 75).

Hardy’s (2013) investigation suggests that, in the case of NAPLAN, the existence of the testing regime served to foreground questions about what constitutes learning and effective teaching in teachers’ minds. Furthermore, his results show that teachers’
interpretations of score meanings, while influenced by the social value they perceived to be attributed to the measures, were not fixed, static, or predetermined. Rather, interpretations of score meaning shifted as they sought to reconcile conflicts between their own world views and perceptions of learning and those they saw embodied in test constructs and score outcomes. Their shifting perceptions, in turn, influenced their teaching practices. Such findings illustrate that score meanings and test consequences evolve in complex and dynamic ways that are shaped by interactions between testing practices and individuals.

As shown above, investigations so far into the intersection between tests and policy spaces reveal that the beliefs and actions of various individual stakeholders extend beyond and sometimes in contrary directions to the objectives set out in official policy documents, leading to complex and unpredictable test and policy consequences. The tension between the general and the local that Hardy (2013) brings to light in the context of NAPLAN, and the indeterminacy of test score meanings and consequences that emerges from this tension, provides a useful point of departure for situating and understanding the impact of language testing practices within policy domains more broadly, as will be discussed under the heading ‘poststructuralist approaches to rethinking test impact and validity in Language Testing’. First, a review of Cooke’s (2009) investigation is presented below.

### 3.4.2 Tests and immigration policy

Cooke (2009) investigates the background to and the impact of the introduction of language test requirements for citizenship and permanent residency in the United Kingdom (UK) on stakeholders, including migrants subjected to the requirements. She highlights the multiple and contingent nature of test purpose in these contexts, pointing out that language tests for immigration and citizenship take many different forms and demand different levels of language proficiency across different countries, and that these differences “are mirrored by differences in how the purpose and aim of the tests themselves are seen by analysts and by stakeholders; there is no agreement as to what,
or indeed who, these tests are for, nor as to reasons why they have come about at this particular moment in time” (2009, p. 71). She reflects on the diverse and competing interests involved in setting the citizenship agenda in the UK, finally comparing the political discourses surrounding the introduction of formal language criteria with perceptions individual migrants have of the test requirements. In particular, Cooke contrasts the liberal framing of the test in political discourses, in which the test was presented as an entitlement, or tool of empowerment for migrants, with individual migrants’ sense of indignation and exclusion at the imposition of a link between the dominant language and citizenship.

Cooke begins by situating her study within the context of the wider debate over the purposes of tests for immigration and citizenship. In addition to the widely mentioned gatekeeping agenda, discussed in the introductory chapter of this thesis, she notes that some see the purpose of such testing as a means of ensuring transparency and minimising the bias involved if decisions are left to individual civil servants. This argument for imposing such tests, based on fairness claims, is often employed to dismantle opposition to the introduction of formal testing in immigration and citizenship contexts (see De Jong, Lennig, Kerkhoff, & Poelmans, 2009, for example). Others, she points out, see the introduction of formal testing as way of empowering new citizens, thereby avoiding their segregation or “ghettoization”. Cooke and others (e.g. Blackledge 2009), identify the latter as a core argument in British political rhetoric concerning the test requirements. Cooke draws on previous work by Blackledge (2006, 2007) to present an argument consistent with much of the literature on the use of tests in these contexts, that is to say, underlying these discourses are ideologies privileging monolingualism over multilingualism, in which a common language is linked to integration and social cohesion.

Cooke (2009) specifies the intended aims of the UK legislation as twofold, to limit immigration from outside the European Union as a means of addressing security concerns, and “to appease the government’s conservative opponents” (p. 73). As discussed in the previous chapter, immigration control measures in the Australian context, such as increasing English language requirements, have been similarly linked to
the need for governments to alleviate public fears about “unwanted forms of migration” (Wright, 2014, p. 397). Cooke emphasises, however, that many diverse stakeholders, including “left-leaning political theorists” and English as a second language (ESOL) teachers, were involved in the development of the test, thereby problematizing (albeit implicitly) the notion of intended test purpose, and challenging assertions such as those made by Blackledge (2009) that the purpose of the test is predominately to exclude.

Through interviews, she examines the response to ESOL and citizenship classes on the part of the teachers and students, reporting that the response from both groups was positive. Based on interview data from migrants taking the citizenship classes, Cooke (2009) reports that would-be citizens experienced problems accessing English classes, ranging from finding a place in a course to begin with, to getting time off work from low-paid, insecure jobs in order to attend classes. She also found that while would-be citizens are committed to learning English, they resent the citizenship stakes attached to their achievement of language proficiency, perceiving the requirements as exclusionary and racist. On the other hand, Cooke points out that students as well as many involved in the provision of ESOL education saw the new citizenship agenda as a positive means of improving migrant access to English classes. Such a perspective further problematizes the notion that test purpose, or the ideologies underlying the implementation of language criteria, can be characterised in homogenous terms.

3.5 Poststructuralist approaches to rethinking test impact and validity in language testing

As discussed so far in this chapter, research examining the relationship between testing practices and public policy shows that the beliefs, actions and reactions of various individual stakeholders interact with macro-level institutional requirements and constraints to produce, influence and shape the meanings of test scores as well as test and policy consequences in complex and multidirectional ways. It has been argued here, therefore, that a more nuanced and problematized view of test purpose is needed in the field of language testing, which accounts for the role of language tests as policy instruments on the one hand, and the influence of individual beliefs and actions in
generating competing score meanings, policy enactments and consequences, on the other. A review of recent literature promoting poststructuralist approaches will now be offered, as a means of conceptualising the inherent tensions and conflicts that exist within and between individual stakeholders as they negotiate test and policy spaces.

Poststructuralism is the name attributed to a series of intellectual discussions which first emerged in France in the 1960s, largely in response to the dominant influence of Saussurean structuralism at the time, across various intellectual domains in France and elsewhere (McNamara, 2012a). Very broadly speaking, poststructuralism is characterised by the questioning of all fixed categories and a view that all relations of power are inherently unstable (Pennycook, 2001). Foucault (1977), one of the major thinkers associated with poststructuralism, emphasises that social norms are established and that subject positions, or categories of ‘being’, are constructed in and through ‘discourses’, which are historically contingent and thereby transient and mutable, rather than fixed and stable. Hall defines Foucault’s notion of ‘discourses’ as a “group of statements which provide a language for talking about—a way of representing the knowledge about—a particular topic at a particular historical moment” (1992, p. 291, cited in McNamara, 2012a, p. 475). McNamara further explains that “Foucault sees all systems of knowledge as discourses, not as ‘truth’ but as ‘regimes of truth’, that is, widely shared and accepted systems of representation” (2012a, p. 475).

Shohamy (2001, 2006) and McNamara (2012a, 2012b) have argued strongly for the relevance of Foucault’s work, in particular *Discipline and Punish* (Foucault, 1977), to understanding the social and political dimensions of language testing practices. Both have referred to Foucault’s discussion of public examinations as instruments of surveillance through which individuals are rendered visible for the purposes of classification and potential punishment. McNamara (2012b), for example, cites the following well-known passage from Foucault, in which the significance of the examination as an organising and disciplining mechanism in modern societies is highlighted:
“The examination combines the techniques of an observing hierarchy and those of a normalizing judgement. It is a normalizing gaze, a surveillance that makes it possible to qualify, to classify and to punish. It establishes over individuals a visibility trough which one differentiates and judges them. That is why, in all the mechanisms of discipline, the examination is highly ritualized. In it are combined the ceremony of power and the form of the experiment, the deployment of force and the establishment of truth. At the heart of the procedures of discipline, it manifests the subjection of those who are perceived as objects and the objectification of those who are subjected” (Foucault 1977, p. 184-185; cited in McNamara, 2012b, p. 564-565).

Shohamy (2001, 2006), discussed earlier, draws on this aspect of Foucault’s work to argue that tests embody, create and shape historically situated and politically contingent views of ‘knowledge’, which necessarily privilege and reinforce existing hierarchies of power. Language tests, according to her argument, serve as a means of assigning social categories (and social value or punitive measures) to individuals according to their ability to assimilate to language behaviours that align with socially and politically dominant, typically monolingual, ‘norms’. McNamara (2012a, 2012b), while acknowledging the importance of recognising the exclusionary and punitive role of tests, as identified by Foucault, extends Shohamy’s argument by giving greater emphasis to an insistence on the inherent and inevitable instability and unfinalizability of any act of classification, which he identifies as a central feature of the works of Foucault and also Derrida.

McNamara (2012b) draws on this theme of indeterminacy, as manifest in the work of Derrida in particular, to present a poststructuralist account of the social and political role of language tests. He identifies a parallel between the widespread use of the biblical story of the shibboleth as a metaphor for the social and political dimensions of testing practices, and Derrida’s discussion of language, and language as shibboleth. Put simply, a shibboleth test is a language-based identity test typically associated with contested
Some textual content extracted from the image.
which different stakeholders associated with high stakes testing regimes, such as PISA, for example, have been shown to interpret test scores in ways that simultaneously reflect their own views and beliefs as well as the broader ideologies they perceive as underlying testing practices, scores and the policy contexts in which tests are embedded. Hardy’s (2013) examination of the impact of NAPLAN, discussed in section 3.4.1 above, foregrounded the tension between the individual/local and the centring forces of social, cultural, and political norms embodied in tests and policies in a way that is consistent with the framing that McNamara (2012b) derives from his examination of Derrida’s use of the shibboleth metaphor.

Recent work in applied linguistics and sociolinguistics has also sought to redefine ‘language’ in ways that allow for the inherent tensions and indeterminacy of language and language practices suggested in poststructuralist thought. Blommaert (2010), for example, argues that new metaphors and new analytic tools are needed to conceptualise and understand the ways in which the effects of power are produced and negotiated in and through the diverse and hybrid array of language practices that characterise an era of globalisation and mass migration. He claims it is necessary “to think about phenomena as located in and across different ‘scales’, from the global to the local, and to examine the connections between these various levels in ways that do not reduce phenomena and events to their strict context of occurrence” (p. 1). His notion of scales refers to a continuum of spatiotemporal frames, ranging from the micro (local) to the macro (global) at the extremes, which are “filled with codes, norms and expectations” (p. 32). Scales, according to Blommaert, thus involve vertical (indexical, stratified, normative) spaces, which are historically contingent and saturated with social, cultural and political ‘discourses’ in the Foucauldian sense, in addition to the horizontal (community, region, country) spaces typically considered as ‘context’ in sociolinguistic studies.

Blommaert (2010) introduces two useful concepts for the purpose of conceptualising how relations of power are manifest in language behaviours: ‘orders of indexicality’ and ‘polycentricity’. The former, he explains, derives from Foucault’s ‘order of discourse’ and refers to the normative organisation of what he calls ‘forms of semiosis’, whereby “some
forms of semiosis are systematically perceived as valuable, others are less valuable and some are not taken into account at all” (p. 38). In other words, ‘orders of indexicality’ relate to what counts as language and who has access to it, in any given moment. According to Blommaert, they represent “systemic patterns of authority, of control and evaluation, and hence of inclusion or exclusion by real or perceived others” (p. 38). A parallel can thus be drawn between this notion and language testing practices, as described by Shohamy (2001) and McNamara (2012b), above, also derived from Foucault’s work, in which tests are viewed as instruments that organise societies through the role they play in surveillance, evaluation and punishing deviation from dominant norms. In line with Blommaert’s (2010) discussion, tests might be viewed as embodying ‘orders of indexicality’.

Blommaert (2010) derives the latter concept, ‘polycentricity’, from Bakhtin’s (1986) notion of the ‘super-addressee’, in which individuals orient to ‘authoritative voices’, or, in Blommaert’s words, “norms and perceived appropriateness criteria, in effect the larger social and cultural body of authority into which we insert our immediate practices vis-à-vis our immediate addressees” (p. 39). This idea is consistent with the analytic framework adopted by Hardy (2013) in relation to the impact of NAPLAN tests on teachers, particular his notion of centripetal forces, which he describes as centring pressures towards standardisation, in line with prescribed policy directions (the authoritative ‘voice’ of political institutions, in Bakhtin’s (1986) terms). Blommaert’s (2010) ‘polycentricity’ extends this idea to account for a multiplicity of competing and sometimes conflicting ‘centres’, or as he describes them ‘evaluating authorities’ (p. 39), to which individuals orient their practices.

Bakhtin’s conceptualisation of language, particularly his conceptualisation of language as dynamic, dialogic, and contingent, has also been drawn upon in relation to second language teaching and learning. Hall, et al. (2005) emphasise the need to look beyond traditional, formalist perspectives on language, in which languages are viewed as fixed, stable, self-contained systems, in order to further understand the experiences of language learners and the nature of language learning, particularly in the face of increasingly diversified and globalized learning contexts. Malinowski and Kramsch
(2013) make use of Bakhtin’s (1981) notion of ‘ideological becoming’ to situate language learning in terms of the tensions and struggle between what Bakhtin describes as ‘authoritative’ discourses or ideologies, such as the discourses of schools, the media and political institutions, and ‘internally persuasive’ discourses, those “of individuals based on their own experiences and convictions” (Malinowski & Kramsch, 2013, p. 156). Similarly, Bakhtinian concepts have been central to expanding theoretical conceptualisations of linguistic diversity in an increasingly globalised world, characterised by mass migration and the destabilisation of traditional norms and boundaries (Blackledge & Creese, 2013).

In a way that is also consistent with Foucault’s notion of ‘discourses’, discussed above, and the poststructuralist emphasis on indeterminacy, as highlighted by McNamara (2012b), Bakhtin (1981) emphasises the mutual interdependence of ‘voices’ (‘discourses’/’ideologies’) that exist simultaneously in every utterance; every utterance holds multiple and competing traces of the past, present and future. Meaning is necessarily co-constructed between two or more ‘voices’ that merge and interact in unpredictable ways to produce something new which then becomes part of the system of discourses/ideologies, thus having an effect on, or altering the course of, ‘language’. Thus ‘language’, according to Bakhtin, (and the social semiotic system within which it exists) is viewed as a dynamic, multiple and open-ended phenomenon that can never be complete, bounded or finalised, except in the abstract. The blending of discourses/ideologies creates new forms of ideological positioning and subjectivities, which then act on and react against subsequent encounters with other discourses/ideologies.

In relation to the current thesis, the work outlined briefly in this section offers a way of conceptualising the multidimensional, layered and contingent nature of values, ‘discourses’ and ‘language’ that are necessarily embedded in language testing practices, particularly as they are utilised in the contexts of public policy. A view of language tests as spaces saturated with competing voices, both within and between individual stakeholders, highlights the need for micro-level investigations of the point of intersection between individuals, with their own shifting and multidirectional set of
beliefs, values and ‘evaluative authorities’, and the ‘authoritative voices’ recognised as embodied by language tests and policy.

3.6 Conclusion

This chapter has sought to demonstrate that the ‘condition of determinability’, a central premise of existing language testing theory, precludes consideration of the values and normative frameworks that underlie testing practices within contested policy spaces. Validity claims, it was shown, are based on an assumption of consensus between test developers and test users, and generally take no account of the potential for multiple and competing interpretations of test constructs, test uses and scores to emerge when tests operate in the domain of policy.

As discussed here, and argued in further detail in the next chapter, within the field of language testing, efforts thus far to incorporate the social and political dimensions of tests have mainly centred on articulating what is meant by ‘fairness’, which has in turn focused on the notion of societal benefit; tests are designed to bring about beneficial consequences for society as a whole. The notion of ‘fairness’, however, is itself value-laden and disputed. As shown in the current chapter, perceptions of test purpose and test consequences can vary markedly between stakeholders, and any assumption of consensus over what might constitute ‘beneficial’ has been undermined by investigations that have explored the beliefs and actions of individuals in response to testing regimes.

As discussed, for example, in relation to the impact of NCLB, PISA and NAPLAN testing regimes, investigations of policy effects at the micro level, specifically on educators and teaching practices, reveal that individual stakeholders appropriate, modify, and resist policy directions and test purpose as they perceive it in diverse and unpredictable ways, leading to complex and conflicting test consequences and policy outcomes. Furthermore, Cooke’s investigation of test impact in the context of immigration and citizenship policy in the United Kingdom showed similar findings, thus problematizing the argument that the use of such tests is necessarily exclusionary (rather than
‘beneficial’), and undermining the notion that test purpose, or the ideologies underlying the implementation of language criteria, are one dimensional and homogenous. Existing research thus supports the argument put forth here that a more nuanced and problematized view of test purpose is needed, which accounts for the various ways tests function as policy instruments and which foregrounds the role of individual stakeholders’ perceptions, feelings and actions in negotiating score meanings and producing impacts through their responses to their test experiences.

In the next chapter, recent efforts in the field of language testing to incorporate issues of fairness and justice into theoretical frameworks will be examined in detail. These efforts, prompted by the increasingly widespread use of language test scores in immigration and citizenship policy contexts, represent a response to Messick’s demand that evaluations of the underlying values and societal consequences associated with score interpretations and uses be incorporated into validity arguments. As will be argued, an insistence on the a priori specification of intended test purpose and score meanings, evident in Messick’s consequential basis of validity and efforts to create practical applications of it, constrain the nature of the questions asked concerning test fairness. This, in turn, limits the field’s capacity both to identify the unanticipated roles tests play in the lives of individuals subjected to score-based judgments, and to account for the unanticipated consequences associated with test use.
CHAPTER 4: FAIRNESS FRAMEWORKS: SOCIAL VALUES AND THE POLITICS OF TEST IMPACT

4.1 Introduction

In the previous chapter, it was argued that current conceptualisations of test purpose, consequences and validity in the field of language testing, premised as they are on what is called here a ‘condition of determinability’, preclude consideration of the multiple and competing values that underscore interpretations of test score meanings, and shape evaluations of test consequences and validity claims. Given the now widespread use of language tests in immigration policy spaces, attention both within and outside the field has increasingly focused on the political role of language tests (and language testers) in furthering contested agendas. The issues raised by the use of language tests as policy tools in these contexts can be situated in an ongoing discussion in the field of language testing concerning fairness and social justice, which is the subject of the current chapter (Davies, 2010; Kane, 2010; Kunnan, 2000, 2004, 2010; McNamara and Ryan, 2011; Xi, 2010).

Language tests as instruments of immigration policy function as gatekeeping mechanisms, determining who will have access to important rights and opportunities and who will be excluded. Test consequences thus inevitably involve issues of fairness that are beyond the internal workings of the test, as questions of what constitutes positive or negative consequences necessarily involve questions of values, which are contingent, multiple and open to conflicting interpretations – who should or should not have access and why or why not? While Messick’s (1989) call to recognise that social and political values underlie test constructs and uses has been taken up in recent discussions of fairness and justice in language testing, how to identify and evaluate such values remains an unresolved issue in the field.
As argued in the previous chapter, although attempts have been made to reconceptualise test fairness to encompass broader issues of social justice, the assumption underlying current validation frameworks is that tests are intended to deliver ‘beneficial’ social consequences, thus validity and fairness claims are evaluated according to the extent to which these intended outcomes are achieved. As will be further argued in this chapter, however, definitions of intended societal ‘benefits’ have been, at best, vaguely articulated in current efforts to account for test consequences. Furthermore, there has been little recognition that multiple and competing notions of societal ‘benefits’ exist, nor that such notions rest on value-laden and contested notions of what constitutes ‘fairness’. As discussed in the previous chapter, interpretations of societal ‘benefits’, ‘fairness’, and ‘social justice’ are all open to dispute.

The aim of this chapter is to examine the ways conceptions of fairness and justice have emerged in the language testing literature as the field struggles to find new ways of imagining the role of tests in high stakes immigration policy contexts. Consistent with the argument made in the previous chapter in relation to validation frameworks in language testing, it is argued throughout this chapter that the need to specify intended test purpose and score meanings as a precursor to evaluations of fairness, a premise also underlying Messick’s (1989) theory of validity, is not only impractical but also misguided, especially when these intentions are necessarily intertwined with policy intentions. Furthermore, adherence to the notion that intended test purpose represents the locus for identifying and evaluating consequences means that the ways tests actually function in policy domains, especially the unanticipated roles tests play in the lives of individuals subjected to score-based judgments, remain hidden from view and unaccounted for.

In this chapter, Messick’s (1989) explanation of validity is revisited briefly, particularly aspects of his work concerning the relationship between test constructs, values and social consequences, as a way of contextualising and shedding light on the as yet unresolved ‘fairness’ debate in language testing. Following this, discussions in the field concerning how to define ‘fairness’ as well as recently proposed fairness frameworks, are considered in detail. This is followed by a brief discussion of how issues of fairness
and justice have been reconceptualised by Amartya Sen (2009) in his ground-breaking critique of Rawls’ (1971) ‘justice as fairness’. The chapter concludes by considering the implications of Sen’s (2009) theory of justice for language testing. A parallel is drawn between Sen’s call for a focus on social realizations rather than on ideals of ‘just’ institutions, and the need to shift the focus of test validation frameworks away from intended test purpose - an unrealisable ideal in immigration policy contexts, towards an emphasis on the lived experiences of individuals subjected to testing regimes, as a means of shedding the condition of determinability and creating space to evaluate the appropriateness of test uses in these contested policy spaces.

4.2 Validity, values and social consequences – revisiting Messick

As is well known, Messick’s (1989) theory of validity consists of two interrelated components, the evidential basis, and the consequential basis of validity, the latter of which relates to questions of values and the societal consequences of test use. Messick’s inclusion of values in his concept of validity highlights the potential influence of implicit ideologies inherent in both the decision to use a test and test scores, but also in the way in which test constructs and consequences are conceptualised, examined and evaluated. Messick suggests the potential for an all-pervasive presence of bias by acknowledging that the influences of perspectives and value positions cannot be eliminated but simply disappear from view when they are shared and uncontested. By specifying a place for values in validity, Messick allows for the possibility that bias exists beyond the internal properties of the test itself, affecting ‘fairness’ and all other aspects of the testing system as it is a product of the shared values and ideological framework within which the testing system was created and operates. Testing thus becomes conceptualised as a socially, culturally and politically-situated practice.

Messick’s (1989) theory suggests a complex and interactive relationship between construct validity, values, and social consequences. Firstly, the construct that a test is designed to measure, however it is conceptualised (as knowledge, abilities, traits or behaviours, for example), is necessarily privileged or valued over other possible constructs that are not identified, defined and measured in the sense that the specified
construct becomes institutionalised. As Shohamy (2001) has pointed out, this creates the potential for tests to play a role in defining and enforcing national language ideologies and norms, privileging certain varieties and uses of a language and delegitimising others (see Chapter 3.3).

Secondly, Messick suggests that the way in which the construct is labelled for testing purposes is likely to have normative and evaluative implications, that is, it comes with an existing ‘value’, which in turn influences the types of consequences likely to be investigated. As an example, Messick argues that if the concept ‘stress’ was relabelled ‘challenge’, the nature of the concept would not change (‘stress’ would not change from a bad thing to a good thing through its renaming), rather, he explains “by not presuming stress to be a bad thing, we would be more likely to investigate broader consequences, facilitative as well as debilitative” (p. 60). This can be applied to the ongoing debates concerning the relationship between ‘validity’ and ‘fairness’ in language testing, including discussion over whether or not ‘fairness’ needs to be separately specified and investigated, as detailed in the next section. ‘Fairness’ is very much a social value, inscribed in the legal systems of democratic societies. Along Messick’s lines, the labels we use to describe desirable test qualities have normative and evaluative implications, which will in turn influence the types of investigations likely to be conducted. Drawing terms such as ‘fairness’ from the vocabulary of political philosophy to label the social and political dimensions of language test use and consequences may therefore direct the field towards the sort of economic, political and philosophical reasoning needed to address the validity questions raised by the use of tests within immigration policies.

Thirdly, Messick specifies that score meanings, interpreted in relation to the already value-laden construct, depend on the theoretical perspectives adopted, which in turn are necessarily contingent on existing worldviews, shaped by existing ideologies. Messick describes an ideology as “a complex configuration of shared values, affects and beliefs that provides, among other things, an existential framework for interpreting the world” (p. 62). Messick emphasises that any given theory, as situated within a given ideology, “carries its own construct system and influences in subtle or substantial ways the kinds of data collected, the types of analyses conducted, and ultimately the range
and nature of conclusions drawn” (p. 61). In terms of consequences, the actual and potential outcomes associated with interpretations of score meaning will also be identified within a given worldview or ideological framework - consequences are only ‘actual’ or ‘potential’ if they are considered possible or imaginable - and consequences will be evaluated as justified, potentially justifiable, or not depending on the relative weighting of existing social values within an existing ideological framework.

In the field of language testing, as set out in Chapter 3, theoretical frameworks are premised on assumptions about test purpose and about test takers that, it was argued, are incommensurate with the way tests function in immigration and other policy spaces. It is assumed, for example, that definitions of score meanings and intended consequences are fixed and stable, as specified by test developers, which, as demonstrated in the previous chapter, is inconsistent with findings from existing studies of test use in policy contexts (see section 3.4). Furthermore, a view of test takers in terms of abstracted and decontextualized components of knowledge and skills precludes any examination of the complex interrelationships between test takers’ thoughts, feelings and actions, which are likely to influence score meanings and consequences. Moreover, societal ‘benefit’ is considered in unproblematic terms, when in fact the notion has been highly contested (for example, Blackledge, 2009; Shohamy, 2006). As argued in chapter 3, these assumptions, which represent fundamental tenets of existing theoretical frameworks in language testing, limit the field’s capacity to describe how tests operate as instruments of policy, constrain the nature of investigations into consequences, which, for the most part, remain narrowly focused on intended test purpose, and hinder efforts to incorporate broader social impacts into evaluations of test validity.

Following Messick’s (1989) line of argument, a certain interpretation and use of test scores never produces an unequivocally positive or negative social impact, as value judgments are implicit in interpretations of consequences as much as they are involved in understandings of score meanings. Although Messick affirms that construct validity encompasses social consequences, he is not claiming that validity is contingent on these consequences being positive, just that the intended ends of the test are met. Bachman
and Palmer (2010) maintain a similar position in their ‘assessment use argument’ framework. An interpretation and use of test scores that delivers negative consequences does not necessarily lack validity. Rather, adverse consequences represent a potential threat to validity and signal the need for further investigation to determine whether or not such consequences stem from a measurement problem, in which case test design is in need of modification, or an acceptable and justifiable outcome of test use. As mentioned in the previous chapter, Bachman and Palmer note that for test takers unable to achieve score requirements in a particular context of use, consequences may be negative since they be unable to access opportunities they seek. These sorts of negative effects are, however, expected and acceptable if test use is justifiably intended to regulate access on the basis of language proficiency.

Importantly, in policy contexts where tests are an instrument used to pursue both measurement and policy ends, Messick (1989) argues that the test maker’s responsibility is to consider and evaluate all intended and unintended consequences arising from test use, including any broad social and systemic effects. In other words, when tests are used as policy instruments, evaluations of test consequences and therefore test validity must include an evaluation of “the outcomes and side effects of policy implementation” (p. 85). Messick thus highlights that claims of validity are inevitably on shaky ground, contingent on supporting evidence but permanently threatened by potential evidence not yet considered or taken into account. His demand to account for the intended and unintended consequences of test use, rests on the premise, critiqued in the previous chapter, that ‘intentions’ are determinable from a single perspective (that of the test developer) or if multiple perspectives are involved, they are implicitly assumed to be congruous. Judgments of validity depend on the determinability of test purpose, even within Messick’s framework, as it is on this basis that evaluations of the appropriateness of test score interpretations, uses and consequences, inevitably rest.

Furthermore, while Messick emphasises the role of competing values, particularly in policy contexts where questions of equality, fairness and distributive justice inevitably arise, any conflict is assumed to be resolvable in the sense that individual values can be
determined, categorised and weighted against each other to evaluate the extent to which they accord with societal values. The broader social values that Messick often refers to, though, are neither specified nor problematized. The underlying assumption is that there are a set of views and beliefs to which members of society broadly subscribe, and that social institutions should embody and reflect such views. In cases where testing is used for gatekeeping purposes, for example, there will clearly be adverse consequences for those excluded on the basis of the interpretation and use of test scores. In such instances, according to Messick’s theory, claims of validity will depend, in part, on a singular intended purpose of the policy (is the purpose served by the test?) and in part on the extent to which adverse and unintended consequences (also determined from a single or set of congruous perspectives) are considered and evaluated as justifiable, on the one hand a question of social values and justice, but on the other a question of the test’s “functional worth”- are the costs (social and economic) of the test use justified by the social benefits delivered? The construct in question (determined/specified/agreed in advance) must be theoretically and empirically defensible as a legitimate basis for such decision-making (Messick’s conditions of relevance and utility) as a prerequisite to considerations of the functional worth of score meaning and use in terms of weighted social consequences.

However, as noted in the introductory chapter, intended test purpose is highly contentious and debated in the context of language testing for immigration and citizenship (Blackledge, 2009; Horner, 2009; Shohamy, 2009; Van Avermaet, 2009). Language tests have been identified by many as unjust tools of exclusion, reinforcing and perpetuating the symbolic and actual marginalisation of particular ethnic and religious groups. In relation to Australia’s skilled migration policy, as noted in chapter 2, Berg (2011) highlights the possibility that, as part of selection criteria for migrants seeking permanent residency, the use of high English test scores is functioning to reinforce discriminatory and racist agendas, sending the message that those who wish to ‘belong’ are required to demonstrate their willingness and capacity to assimilate to dominant English norms.
Whether or not this is the purpose that language tests in these contexts are intended to serve is open to question, as such assertions are inevitably contingent on perspective and political orientation. De Jong et al. (2009), for example, hold a less critical view. The researchers make a link between ‘integration’, described as social inclusion, and the national language. They imply that the use of a Dutch language test as a prerequisite to immigration is intended to protect potential migrants, by ensuring that they have the language skills needed to participate in the new society, as well as to protect existing members of Dutch society in that a lack of participation on the part of migrants can be linked to various problems, both economic and social. Test scores that function as a basis for decisions about the allocation of rights will inevitably result in both inclusion and exclusion; some will pass and some will fail. The intended purpose of a language test in such contexts, following De Jong et al.’s logic, would ideally be to include those who possess the linguistic skills and attributes necessary to succeed as a migrant in the new society (assuming the link between language, integration and tests might be sustained by empirical evidence and reasoned argument) and to exclude those who do not. Deciding what these skills and attributes might be, however, let alone agreeing on what is meant by ‘succeed’, are likely to be hugely problematic endeavours, and were not attempted by the researchers. Similarly, the link between language, integration, and tests was accepted without question.

In the context of test use within Australia’s skilled migration policy, as detailed in Chapter 2, Birrell et al. (2006) made a similarly tenuous link between language, integration, and tests that drove a series of reforms to the English test score requirements for permanent skilled migration (see section 2.3.2 ). As noted in Chapter 2, the English language criteria that have long since existed as part of skilled migrant selection have always been ostensibly aimed at ensuring migrants possess a level of English adequate for accessing skilled employment. However, any previous link that might have existed between test scores and the English level needed by migrants was undermined by the policy changes introduced in 2011 and 2012, and since then, intended test purpose has become difficult, if not impossible, to specify in terms consistent with the requirements of Messick’s (1989) validity theory. As discussed in Chapter 2, the score requirements for permanent residency have become increasingly
contingent; whether IELTS 6, 7 or 8 is needed depends on how an applicant fares across other, non-language related criteria as well as on the number and profile of other applicants. Any link between scores and the English demands faced by migrants thereby becomes almost impossible to define (see section 2.4.2 ). Testing companies, such as IELTS, maintain their distance from the questions raised by such score uses by asserting simply that their tests provide a measure of test takers’ English proficiency, similar to De Jong et al.’s defence, above, of the use of a Dutch language test for immigration purposes in the Netherlands. However, as Fulcher and Davidson (2009) assert, “a test that does not have explicitly defined purpose also creates validity chaos” (p. 125, emphasis in original). This is because test validation, as it is conceived by Messick (1989) and in the field of language testing, demands an explicitly defined and specific test purpose around which to orient decisions about what validity evidence to collect and how such evidence should be prioritized.

While Messick’s (1989) emphasis on accounting for the social and political values implicit in testing practices highlights the contestability of construct definitions and underlying notions of ‘success’ and ‘failure’, for example, the assumption remains that consensus can (and should) be achieved. He advocates the inclusion of multiple and diverse perspectives as a means of promoting reasoned arguments capable of shedding light on hidden values and invisible ideological positions, with the purpose of arriving at a particular position. Tests play a role in positioning individuals and organising societies, supposedly in a way that is fair and just. Messick recognised that beliefs about the world are ever-present in all of our empirical investigations and reasoned arguments, and are therefore ever-present in our conceptualisations of ‘fairness’, ‘justice’ and in the case of testing, ‘validity’. While his framework, through its emphasis on implicit values, provides an important basis for considering the role of tests in policy contexts, the underlying demand for determinability does not allow for a view of score meanings and values as multiple and dynamic, and thus does not allow for the indeterminable nature of consequences and fairness outcomes. As shown in the next section, the circular debate over issues of ‘fairness’ and ‘justice’ in the field of language testing highlight the impossibility of maintaining this condition of determinability while trying to
reconceptualise tests as socially and politically-situated policy instruments with pressing social justice implications.

4.3 The ‘fairness’ debate in Language Testing

As mentioned at the outset, the challenge of conceptualising language testing as a socially and politically situated practice has led to the emergence of several attempts to broaden definitions of ‘fairness’ as well as different views on the relationship between fairness and validity. Stoynoff (2013) distinguishes two broad categories of fairness investigations evident in much of the existing language testing and applied linguistics literature: research which takes a psychometric perspective on fairness, and research focussing on issues of test impact. In the former, fairness is treated as a property of the test itself, viewed in terms of equity and absence of measurement bias, while the latter is concerned with social consequences and the social and political values associated with test use. Stoynoff further points out that notions of fairness that have emerged more recently in the language assessment and applied linguistics literature are derived from and shaped by views on fairness articulated by specialists in the field of educational measurement, such as Messick (1989), Willingham and Cole (1997), Willingham (1999) and Kane (2006).

Kunnan (2000, 2004) and Xi (2010) in the field of language testing, for example, have sought to integrate the two ‘fairness’ categories mentioned above in order to more closely reflect Messick’s (1989) expanded notion of validity, as well as to mirror current understandings of fairness as articulated in the *Standards for Educational and Psychological Testing*. Both Kunnan and Xi, despite differences in the way they define ‘fairness’, argue for the inclusion of ‘fairness’ criteria in evaluations of test quality that encompass both the psychometric properties of tests and the actual social consequences of test use and score interpretations, viewed primarily in terms of social equity outcomes. Given the important influence of the *Standards* and the educational measurement literature on thinking on ‘fairness’ in applied linguistics, ‘fairness’ as it has been conceived in these contexts is briefly introduced below, before turning to an
examination of the notions of fairness articulated by Kunnan, Xi and others in the field of language testing.

‘Fairness’ in the educational measurement literature is recognised as a social value that is defined and understood within particular social and political contexts (Camilli, 2006). Camilli (2006) notes that the term ‘fairness’ has a broad usage in the 1999 *Standards for Educational and Psychological Testing* as well as in the educational measurement literature. The *Standards* specify that:

“A full consideration of fairness would explore the many functions of testing in relation to its many goals, including the broad goal of achieving equality of opportunity in our society. It would consider the technical properties of tests, the ways test results are reported, and the factors that are validly or erroneously thought to account for patterns of test performance for groups and individuals” (cited by Camilli, 2006, p. 225).

Camilla goes on to define fairness in testing in a way that reflects the broad usage of the term in the field of educational measurement:

“Fairness in testing refers to perspectives on the ways that scores from tests or items are interpreted in the process of evaluating test takers for a selection or classification decisions. Fairness in testing is closely related to test validity, and the evaluation of fairness requires a broad range of evidence that includes empirical data, but may also involve legal, ethical, political, philosophical, and economic reasoning” (2006, p. 225).

The broad and situated approach to fairness illustrated in the above quote is embodied in the work of Messick (1989), as discussed previously, as well as others, such as Kane (2006), and is clearly evident in the views of Kunnan (2000, 2004) and Xi (2010), as mentioned above. Kunnan (2000, 2004), for example, emphasises the primacy of fairness over concepts such as test validity and reliability. He conceptualises fairness as an all-encompassing quality that includes both tests and testing practices, and which
centres around issues of equity and social justice that extend beyond test validity as it is typically represented in the language testing research.

Kunnan (2004) claims that despite the fact that the Standards have reflected Messick’s expanded view of validity since 1985, which incorporates values and the broader social consequences of tests and testing practices, validity-related literature in the field of language testing, with the exception of Bachman (1990), remains primarily focussed on the internal and psychometric properties of tests, with little attention given to questions of social justice. Kunnan (2004) argues that this lack of attention is further evidenced by the absence of any reference to fairness in much of the research conducted as part of practical language test evaluation efforts. In response, he proposes a ‘test fairness framework’, derived from Rawls’ (1971) ‘justice as fairness’. Rawls (1971) specifies an imagined ‘original position’ of primordial equality - public deliberation under a ‘veil of ignorance’ whereby members of society have no knowledge of their personal identities, their positions in society or their respective vested interests - a hypothetical position that leads to unanimously chosen principles of justice. Rawls thus viewed fairness as foundational to justice, a necessary precondition to his principles of justice. By contrast, Kunnan’s (2004) test fairness framework tends to conflate fairness and justice, as illustrated by its two broad ethical principles, below:

“Principle 1: The Principle of Justice: A test ought to be fair to all test takers; that is, there is a presumption of treating every person with equal respect.
Principle 2: The Principle of Beneficence: A test ought to bring about good in society; that is, it should not be harmful or detrimental to society” (p. 33).

Kunnan goes on to present his concept of fairness as encompassing five broad qualities: (1) Validity (by which he seems to mean ‘validity’ as it has been narrowly conceived in much of the language testing literature); (2) Absence of bias (broadly conceived as absence of prejudice and discriminatory characteristics of test materials, as well as equality of testing outcomes across different groups); (3) Access (defined as access to education, as well as financial, geographical and personal access, and familiarity with test equipment, procedures and conditions); (4) Administration (appropriateness of
physical conditions and uniformity across administrations); and (5) Social consequences, viewed as both ‘washback’ and the broader impact of testing on society. In terms of the broader social impact of testing, Kunnan asks questions that are essentially the same as questions concerning social consequences raised by Messick (1989): if “the social consequences of a test and/or the testing practices are able to contribute to societal equity or not and whether there are any pernicious effects due to a particular test or testing programme” (Kunnan, 2004, p. 39).

Xi (2010) is also concerned with incorporating broader social consequences, particularly social equity outcomes, into language test validation investigations. She proposes the inclusion of a “fairness argument” into the argument-based validation framework originally proposed by Kane (1992) and associates (Kane, Crooks & Cohen, 1999). The original framework includes an interpretive argument (a clear statement of how test scores are to be interpreted), and a validity argument which addresses the extent to which the assumptions underlying the interpretive argument are supported or undermined by empirical evidence. More recent work by Kane (2001, 2002, 2004, 2006) and Bachman (2005) sought to extend the framework beyond the interpretation of scores to include issues of test use and consequences, but provide little in terms of how test use and consequences might be conceptualised and evaluated in practice.

Xi (2010) argues that the fairness argument provides a means to guide practical evaluation efforts concerning the broader social consequences of test use and score interpretations. Adopting Willingham and Cole’s view that fairness means comparable validity across all relevant groups (Willingham & Cole, 1997; Willingham, 1999), or, in other words, equitable treatment and equality in terms of testing outcomes across groups, Xi (2010) uses the TOEFL iBT test as an example to outline how evidence related to the question of comparable validity or fairness might be systematically organised into an overall fairness argument within an argument-based validation framework.

Kunnan (2010) criticises the framework proposed by Xi (2010), raising doubts as to the feasibility of constructing a coherent fairness argument when competing and conflicting claims about test purpose exist. Using the example of the US Naturalisation Test, Kunnan
(2010) lists seven different claims that have been made about the test by different government officials, competing political commentators, testing experts, educators and test takers. Some claim that the test assesses English language proficiency and knowledge of American society, others make no reference to language. Some believe the test assesses integration and nationalism, others claim that it assesses whether applicants will make a positive contribution to the USA, and some argue that the test is not assessing anything, functioning instead as an arbitrary gatekeeping mechanism. Kunnan (2010) argues that Xi’s (2010) fairness argument must take account of these competing claims, deciding which to investigate and then incorporating the relevant “claimants’ challenges of the warrants and backing” (p. 187). He also suggests that beyond these explicit claims, there are potentially implicit claims that further confound the construction of a coherent and integrated fairness argument.

Davies (2010), by contrast, questions the need and the usefulness of making any distinction between fairness and validity in language assessment. Firstly, he argues, all of the fairness questions identified by Xi (2010) are already subsumed under the heading ‘validity’. Secondly, he points out that ‘fairness’, understood as contingent on equity, is not always necessarily a desirable test quality. Equitable treatment, he claims, requires individualising test materials and testing procedures, which can undermine comparability of results. Davies further argues that fairness should be understood in terms of the ‘rules of the game’- if the rules of the game are fair (if the test is valid), then any inequalities or negative consequences for individuals or groups are simply unlucky, or accidental, rather than unfair. He is not thereby suggesting that such inequalities and consequences should be thought of as fair outcomes; simply that it makes no sense to label such things as ‘fair’ or ‘unfair’ if there is no agency involved in causing them. He draws on Rawls’ thinking to suggest that “life is unfair: it is the responsibility of society through its system of justice to make it marginally less unfair” (p. 175). Validity, he goes on to argue, like justice, “guarantees that an ability is being appropriately tested for a relevant population: this population will be made up of various groups but there is sufficient homogeneity across groups for them to be treated as comparable” (p. 175). It should be noted, however, that neither Xi (2010), nor Kunnan (2000, 2004) are suggesting that the fairness questions and qualities they propose are outside of
Messick’s (1989) expanded notion of validity. Davies (2010) argues that there is, therefore, no point re-labeling some aspects of validity as fairness. Davies’ argument, however, seems to at least imply that validity promotes fairness, even if only in terms of reducing unfairness.

Kane (2010) highlights a distinction in law and political philosophy between two broad conceptions of ‘fairness’, which he argues are embodied in the distinction between procedural and substantive due process in the US legal system. From this distinction, Kane derives two types of ‘fairness’ that he deems relevant to assessment: ‘procedural fairness’ and ‘substantive fairness’. According to Kane, procedural fairness demands equal treatment for all test takers. That is, the same test materials, conditions, and scoring procedures apply for all test takers, unless adjustments are deemed necessary to ensure equity. Substantive fairness, he explains, requires that “score interpretation and any test-based decision rule be reasonable and appropriate” (Kane, 2010, p. 178). In other words, procedural fairness requires that everyone is treated in the same way; test materials and conditions of test administration are the same for everyone, and that rules for scoring are consistently applied across all test takers. Substantive fairness, on the other hand, requires that the scoring and decision-making rules, the rules which are applied ‘fairly’ or equally across all individuals, are themselves fair.

Kane (2010) acknowledges that ‘procedural fairness’ is likely to be understood differently in different testing contexts, depending on the purpose of the test. He claims that ‘fairness’ might be conceived of as equitable treatment in some situations, whereas equality would be privileged in other situations. In an achievement test, for example, Kane argues that ‘fairness’ demands that all test takers have had equitable opportunity to learn the materials covered in the test, and where this has not been the case, allowances should be made to level the playing field for disadvantaged test takers. In the case of certification or licensure, however, whereby a particular set of language skills or level of language proficiency is deemed necessary in order to function in a specific domain, then ‘fairness’ might be understood in terms of equality rather than equity.
Kane (2010) explains that ‘substantive fairness’ concerns the function of testing programs, in particular, how the test works for particular groups, but he does not explicitly acknowledge that any evaluation of substantive fairness is contingent on how ‘fairness’ as a social value is conceptualised. The terms he uses, ‘reasonable and appropriate’, represent subjective judgments that rest on underlying moral principles, and the complexity of this is understated by Kane (2010). In terms of Messick’s (1989) validity framework, Kane’s (2010) procedural fairness is consistent with the evidential basis of validity, and substantive fairness is consistent with the consequential basis of validity. Kane (2010) emphasises that procedural fairness is a precondition for substantive fairness in testing, but that procedural fairness does not guarantee substantive fairness. Unlike Messick (1989), however, he fails to make explicit the role of values in test constructs, testing practices and any evaluation of test consequences.

In a more recent paper, McNamara and Ryan (2011) make a distinction between ‘fairness’ and ‘justice’ as a means of highlighting the role of social values in testing. ‘Fairness’, the authors suggest, refers to “not only a concern with equal treatment of groups and absence of psychometric bias but all aspects of the empirical validation of test score inferences in the interests of yielding reasonable and defensible judgments about individual test takers”, whereas ‘justice’ refers to “the consequential basis of test score interpretation and use but also, and particularly, the social and political values implicit in test constructs” (p. 167). ‘Justice’ is thereby connected to both the evidential (test constructs) and consequential bases of validity. As McNamara and Ryan (2011) emphasise, the field of language testing is well equipped with the statistical means to address potential threats to ‘fairness’ caused by internal inconsistencies, unequal treatment for individuals or disproportionate outcomes across different groups, but lacks the means to uncover and explicate the social values inherent in test constructs and testing practices, an implicit source of bias in even the most well defined tests, according to Messick (1989).

The distinction proposed by McNamara and Ryan (2011) is similar to Kane’s (2010) procedural and substantive fairness distinction, discussed above, although Kane (2010) places no explicit emphasis on ‘values’. Procedural fairness, described by Kane,
becomes simply ‘fairness’ in McNamara and Ryan’s terms, and substantive fairness resembles McNamara and Ryan’s notion of justice. Like Kunnan (2004), McNamara and Ryan (2011) highlight an emphasis in existing language testing validation work on the evidential basis of validity, and the relative lack of research concerning Messick’s (1989) consequential basis of validity, particularly aspects which call for social values to be made explicit and evaluated.

While, as Messick (1989) emphasises, there is a clear need to explicate the values implicit in test constructs and in score-based decisions and evaluations of consequences, it is unclear how the distinction proposed by McNamara and Ryan (2011) serves this purpose. Firstly, it is in some ways inconsistent with broader understandings of ‘fairness’ inscribed in law and political philosophy, whereby, as noted by Kane (2010), ‘fairness’ is conceptualised as encompassing both procedural and substantive aspects: individuals should have equal access to opportunities and wealth, which may mean they are all subjected to the same rules in the same way, but might also require equitable adjustments (procedural aspect of ‘fairness’). Furthermore, the rules of the game must make sense before ‘fairness’ can be asserted (substantive aspect). Justice, following on from this, is interrelated with fairness, not a distinct concept: In Rawls’ (1971) terms, the rules of social organisation must be derived from principles or values that are accepted and endorsed by members of the social group. Justice, according to Rawls (1971), is a matter of both procedural and substantive fairness. Overall fairness (procedural + substantive) is embodied by just social institutions, and justice is conceptualised as institutionally guaranteed procedural and substantive fairness (‘justice as fairness’). Questions of justice arise when institutions are thought to fail to deliver ‘fairness’ in overall procedural and substantive terms.

Secondly, McNamara and Ryan (2011)’s distinction between ‘fairness’ and ‘justice’ implies that ‘justice’ is a question of social values, whereas ‘fairness’ is not. This is inconsistent with the socially and politically situated understanding of ‘fairness’ evident in the educational measurement literature, particularly in the work of Messick (1989). The contingent nature of ‘fairness’ is also acknowledged in the 1999 Standards, which
state that fairness is “subject to different definitions and interpretations in different social and political circumstances” (cited by Camilla, 2006, p. 225).

McNamara and Ryan (2011) provide the example of the Australian Dictation Test, an instrument of Australia’s explicitly racist ‘White Australia Policy’ in the early 20th century. The test was administered to undesirable would-be migrants in a language that immigration officials knew the applicant was unlikely to speak or understand. The authors argue that while such a test met the conditions for ‘fairness’ - applicants were treated in the same way and the ‘rules’ were consistently applied – the values represented by the test construct and use of the test, as well as the policy within which the test was used, “were objectionable on all sorts of moral and political grounds” (McNamara & Ryan, 2011: 168). In other words, the test was ‘fair’ but not ‘just’. The implication is that ‘fairness’ is a function of consistency and equal treatment, not values, whereas ‘justice’ is a question of morality – it is wrong to use a language test to promote a racist policy agenda.

The justice issue above seems, of course, self-evident. However, the claim that the Dictation Test satisfies conditions for ‘fairness’ is not altogether convincing. In order for a test to be deemed ‘fair’ (or valid), even in the most narrow sense with no explicit reference to values, the link between score meaning and any score-based decision must be substantiated both theoretically and empirically (Kane, 2006, 2010; Messick, 1989). In other words, values aside, conditions for procedural and substantive fairness must be met. As Kane (2010) explains, and as McNamara and Ryan’s (2011) own definition of ‘fairness’ implies, it is not enough that criteria are consistently and equally applied, the decision-making rules linked to the criteria also need to be shown to be reasonable and defensible in order for ‘fairness’ to be asserted.

McNamara and Ryan (2011) conclude with an evaluation of the ‘fairness’ and ‘justice’ aspects of the Australian citizenship test. The citizenship test is a formal measure of knowledge of society rather than language proficiency, but as the authors point out, it also functions as a measure of proficiency in the national language, albeit implicitly, given that all of the test materials are in English. The authors argue that demands for
test ‘fairness’ led to a refinement of the linguistic and content-related properties of the test in an attempt to more closely reflect legislation and to eliminate prejudiced content, but questions of justice, particularly concerning whether or not the test should exist at all, were largely ignored.

Messick (1989) argues that it is the responsibility of test developers to reveal and justify the values test constructs and test use embody. His notion of validity rests on the premise that properly constructed and validated tests should embody positive social values of equality of opportunity and equitable treatment. McNamara and Ryan argue that this is a naïve view, given that “test constructs are increasingly dictated as a function of policy” (2011, p. 175). In policy contexts, the authors suggest, the purpose and impact of test use (issues of justice) can be covert and difficult to uncover, and a preoccupation with scientific evaluations of ‘fairness’ obscures issues of ‘justice’. Messick’s (1989), like Rawls’ (1971), assumes that although competing perspectives are inevitable, underlying the multitude of diverse positions there are identifiable core values, which are universal and which are, or at least should be, embodied in valid tests and testing practices. McNamara and Ryan’s (2011) discussion of the evolution and associated debate concerning the Australian Citizenship test suggests, by contrast, that tests are sites of struggle where social values are negotiated and contested. Such a view denies the existence of universal values, let alone the possibility that such values might be embodied in tests and testing constructs.

As shown in the next section, recent work by economist and philosopher Amartya Sen (2009) to reconceptualise the notion of justice to account for the existence of diverse and potentially irreconcilable perspectives and values, offers a way forward in addressing the justice questions raised by McNamara and Ryan (2011).

4.4 Reconceptualising ‘justice’ – Sen (2009)

In a similar vein to McNamara and Ryan’s (2011) emphasis on conflicting rather than universal values in framing an approach to justice in testing, Sen (2009) calls for the abandonment of the ideal, fixed formulations of ‘justice’ that derive from Rawl’s (1971)
seminal work ‘justice as fairness’. Sen (2009) asserts that while Rawls’s notion of a veil of ignorance represents an awareness of the threat to fairness posed by a lack of impartiality, his ideal of social cooperation between citizens who deliberate without prejudice and agree to the principles of justice is nonetheless susceptible to local bias. Without a procedural means of examining local values, Sen argues, there is no guarantee that the principles derived are not based on “preconceptions and biases that are common in the focal group” (2009, p. 28).

In identifying the potential influence of local values and prejudices in deriving principles of justice, Sen (2009) highlights an implicit threat to ‘justice as fairness’ inherent in the very process by which ‘justice’ is derived. Hidden bias is potentially built-in to the position of primordial equality and sustained throughout discussions because members of a common group have internalised particular values and moral positions. These values and moral positions appear to be universal, and are not recognised as open to contestation or even question. With no method to expose the ideologies underlying the ‘original position’, there is no guarantee of ‘primordial equality’ or ‘fairness’.

This seems similar to arguments raised by Messick (1989) concerning the role of values in testing practices, detailed in section 4.2, in that he also identifies local prejudice or hidden ideologies as a potential threat to fairness and therefore validity. Sen (2009), however, is demanding more than Messick. In fact, in terms of fairness, broadly conceived, and social justice, Messick’s (1989) theory of validity is consistent with Rawls’ (1971) theory of ‘justice as fairness’. Rawls’ ‘original position’ is derived from an imagined debate in which participants are unaware of their own attributes or of their position in society and so are not prejudiced by their own individual interests. Viewed another way, the ideal principles of justice are derived from all possible perspectives, which are all equally weighted. Implicit in both Rawls’ theory of justice and Messick’s theory of validity is the assumption that, while competing perspectives of course exist, underlying the multitude of diverse positions there are identifiable core values, which are universal. Rawls articulates what he believes these fundamental values to be (his two principles of justice), whereas Messick refers generally to ‘social values’, but the premise is that a society exists in which the members share a set of beliefs and values.
Sen (2009), by contrast, is demanding a more global view of society and of justice in which ideologies are not simply competing but are potentially irreconcilable, which aligns well with considerations of test use in immigration policy contexts, where diverse perspectives, conflicting interests, and disputes over values are inevitable. Both Rawls (1971) and Messick (1989) assume that if bias can be made explicit and removed, then agreement or consensus can somehow be reached and differences resolved, whereas Sen calls for the abandonment of ideals of justice, which depend on consensus over values, in favour of an approach that recognises and accepts diversity and competing beliefs without requiring any sort of universality. Sen implies that values and beliefs, and therefore justice, are inevitably sites of conflict and struggle. He argues instead for an approach that aims to identify and reduce actual injustices in the world, rather than a transcendental approach that aims to define and institutionalise principles of perfect justice.

Sen (2009) similarly criticises the “prevailing concentration in mainstream political philosophy on transcendental institutionalism” (p. 24), which he argues stems from Rawls’s concept of ‘justice’ as institutionally embodied fairness. Transcendental institutionalism, he explains, assumes that sovereign states apply principles of justice through the choice of a perfect set of institutions. This notion is highly problematic, Sen argues, because it takes no account of actual social realizations. So-called ‘perfect’ institutions, he claims, are abstracted from the social, political, economic, and cultural contexts within which they function. There is no recognition that these institutions are heavily situated, deriving meaning and facilitating real social processes and consequences in interaction with various other societal factors. The assumption is simply that ‘just’ institutions deliver social justice; the institutions themselves rarely receive any critical scrutiny. Sen argues for a theory of justice that allows for ongoing critical examination and evaluation of actual social processes and consequences: “to ask how things are going and whether they can be improved is a constant and inescapable part of the pursuit of justice” (p. 86).
4.5 Implications for language testing

A parallel can be drawn between Sen’s notion of ‘just’ institutions and tests. Standardised tests are assumed to offer transparency and ‘equal’ treatment to all applicants. Procedural fairness requires that decisions are not based on the whim of individuals, but rather on a set of clear and unambiguous criteria that are consistently applied without bias. In this way, valid tests might be assumed to deliver just outcomes because they are ‘fair’. Messick (1989) highlights the social and politically situated nature of testing practices, emphasising the idea that ‘equality’ and ‘fairness’ are embedded in ideologies; they are not neutral concepts. He similarly asserts the need for ongoing evaluation of the outcomes of testing practices as a way of examining and evaluating the values and ideologies they promote. The validity of testing practices and procedures that appear to deliver ‘fairness’ may be undermined by actual consequences and social realisations.

However, Sen (2009) argues against the idea of the universal social values implicit in Messick’s theory, instead emphasising that multiple, conflicting and potentially irreconcilable values and fairness ideals are the norm in modern societies. Sen (2009) thus calls for the abandonment of an ideal, fixed formulation of ‘justice’ in favour of a view of justice in terms of actual consequences. This position is echoed by McNamara and Ryan’s (2011) questioning of the justice implications of creating and using tests as policy instruments. McNamara and Ryan suggest that the justice question involves asking if the test should exist in a given policy context - does (or would) the existence of a test promote better or worse social justice outcomes? This implies that the potential for tests to create injustices should represent a focal point of validity evaluations, rather than intended test purpose as imagined by test developers.

However, as McNamara and Ryan (2011) point out, this question is rarely asked by language testers. Testing companies, as suggested earlier, tend to distance themselves from the social and political issues raised by the uses of their instruments in policy contexts by adhering to the simplistic claim that the purpose of tests is to measure
language proficiency, as if this were somehow a value-neutral function. As noted in section 4.2, above, this lack of specificity also creates ‘validity chaos’ (Fulcher & Davidson, 2009), undermining any possibility to pursue test validation efforts as set out within conventional validation frameworks, which necessarily hinge on precise and constrained definitions of intended test purpose.

Consistent with arguments put forth by Sen (2009) and McNamara and Ryan (2011), it is argued here that to properly understand and evaluate the role of tests and the issues raised by their use in immigration and other policy contexts, test purpose and consequences, as perceived and experienced by those subjected to testing regimes rather than as intended by test developers, must constitute the focal point around which the appropriateness of testing practices is evaluated. The merit of such an approach is further supported by existing research into the use of tests in policy domains, which, as shown in chapter 3, demonstrates that individuals negotiate test and policy constraints in complex and unpredictable ways, according to their own perceptions of test purpose and their own intentions and goals. Such an approach offers a more nuanced and problematizing view of test constructs, uses and outcomes, by foregrounding the role of individual subjectivities and actions in shaping score meanings and test consequences, and thereby provides a means of shedding the condition of determinability that currently limits our capacity to identify and account for test impact in immigration policy contexts.

4.6 Conclusion

In this chapter, Messick’s (1989) consequential basis of validity and the different conceptions of fairness and justice in the language testing literature that have stemmed from his seminal work were examined. It was argued throughout this chapter that the unrealizability of the condition of determinability, which underlies theoretical frameworks in language testing and is also evident in Messick’s theory of validity, represents an obstacle that hinders efforts in the field to identify and account for the roles tests play in policy contexts. As set out in this chapter, a requirement to specify a defined intended test purpose as a precursor to identifying and evaluating
consequences has prevented the field from moving forward in describing how tests function in policy domains, and how these functions impact individuals and society.

In an effort to reframe the fairness debate and to shed this condition of determinability, Sen’s (2009) theory of justice and the implications this might have for testing was considered. The chapter concluded with the argument for a shift in the focus of test validation frameworks away from intended test purpose, which cannot in any case be adequately specified, towards foregrounding the role of test taker perceptions, feelings and actions, such that test purpose and the consequences of test use, as experienced by the individuals subjected to testing regimes rather than test developers or score users, represent the locus of evaluations of the appropriateness of testing practices. As set out in Chapter 1, the current study, as will now be presented across the next four chapters, supports this argument.
CHAPTER 5: RESEARCH METHODS

5.1 Introduction

As set out in the introductory chapter to this thesis, the current study is intended to shed light on questions of test use and test impact in the context of Australia’s skilled migration policy by examining in detail and over time the experiences of four individuals seeking to transition from temporary to permanent residency status as skilled migrants. Four case studies were undertaken; Each of the four individuals were former students of Australian tertiary institutions, and they had all been living in Australia for several years before commencing their efforts to acquire permanent residency. They were each required to demonstrate test scores of IELTS 7 or above in order to become eligible to make the transition to permanent status, and all experienced repeat test encounters in their attempt to achieve the required scores. This study aims to identify and explain how these four individuals perceived their experiences of the test and the score requirements across repeat encounters, how their perceptions of and interactions with the test influenced their decisions and actions, and how they perceived the impact of test use on their lives in this policy context.

In this chapter, the research methods adopted to address the aim of the study, as specified in Chapter 1 and summarised above, are described. The chapter is divided into two main sections, ‘Methodology’ and ‘Procedures’. In the first section, an account of the theoretical and epistemological underpinnings of the study design is provided. In the second section, participant details and the data collection, preparation and analysis procedures are described.

5.2 Methodology

Given the exploratory nature of the current study, the methodology adopted was a longitudinal, grounded theory-based approach that involved generating themes from an iterative analysis of a series of open-ended, in depth interviews conducted with each of the four migrant test takers. These four individual test taker cases constitute the primary
object of investigation in this study. Analysis was broadly targeted at identifying differences as well as similarities between their experiences of the test, to examine the extent to which individual subjectivities shape the role and effects of a language test in this policy space.

In this study, the four migrants’ cases were supplemented by a fifth ‘teacher perspective’, which consists of seven teachers’ accounts of their perceptions of the use of English language testing for skilled migration purposes in Australia, and its effects on the test takers they encountered in their IELTS preparation classes, as well as on their own practices as teachers. While the perceptions and experiences of migrant participants is the primary focus of the study, the perspective of test takers, particularly their criticism of testing regimes, is vulnerable to being dismissed as motivated by a lack of test success. Semi-structured, single interviews with seven teachers were designed to capture an alternative perspective on the use of language tests in the process of migrant selection, and to provide a means of interrogating the plausibility of migrant accounts as well as the credibility of the researcher’s interpretations of these accounts.

In this section, an overview of the methodological approach taken in the study is first provided, followed by an overview of the theoretical and epistemological assumptions associated with the in-depth interviewing and grounded theory-based approach adopted. While in-depth qualitative interviewing and grounded theory methods together constitute an integrated methodological approach, for the purpose of clarity, the two aspects of the approach are presented separately.

5.2.1 A case study approach

Case studies are a well-established approach in applied linguistics and elsewhere to explore the complexities and particularities that characterise individual experiences of a particular phenomenon (Duff, 2008, 2014; Stake, 2008; Yin, 2014). Case study approaches have been widely used in applied linguistics to investigate various aspects of language learning, including second language learning development (e.g. Duff et al., 2013; Lardiere, 2007), learner identities (e.g. Norton, 2013) and migrants’ experiences
of language learning (e.g. De Costa, 2010; Ricento, 2013). As Duff (2014) points out, case study research has contributed significantly to theory building in these areas of applied linguistics.

Stake (2008) makes a distinction between intrinsic case studies, whereby the case itself is of particular interest and does not necessarily reflect a broader issue or problem, and instrumental case studies, whereby the purpose is primarily to gain insights into a particular issue or problem. As he explains, an instrumental case “facilitates our understanding of something else. The case is still looked at in depth, its contexts scrutinized and its ordinary activities detailed, but all this helps us pursue the external interest” (2008, p. 123). The four individual migrant experiences that are the primary focus of this study are considered as ‘instrumental cases’ in Stake’s terms, and the ‘external interest’ that the cases are intended to shed light on in this study is test impact, manifest through the subjective experiences of individuals. A focus on four individual cases was adopted, rather than one-off interviews with a greater number of participants, in order to capture potential changes over time in participant perceptions as they interacted with the test across repeat encounters.

It is important to note that instrumental case studies are not necessarily viewed as typical or generalizable (Stake, 2008), and this study was not intended to provide any basis for claims of generalizability. Rather, by comparing four individual cases over the period of their transition from temporary to permanent residency status, the intention was to gain insights into how the language test requirement was perceived and accounted for at the individual level, and also to examine the extent to which important themes underlying individual experiences with the testing regime could be identified across cases. Identifying common themes across two or more cases provides a form of triangulation to support interpretations of data in qualitative research, and also provides an important first step in theory building (Duff, 2014; Stake, 2008; Yin, 2014).

In this study, five perspectives are examined, that of the four migrants and a fifth ‘teacher’ perspective. Teacher accounts of the test requirement and its associated consequences, according to their own experiences, were compared with the four
migrant cases, providing a further form of triangulation and an additional means of verifying the credibility of themes and interpretations. Full details of the procedures undertaken to ensure the credibility, dependability and transferability of the research findings (Lincoln & Guba, 1985), are detailed further below in the procedures section, under the heading ‘Ensuring trustworthiness’.

5.2.2 In-depth qualitative interviewing

In-depth qualitative interviewing is recognised across various disciplines as an effective and open-ended means of eliciting and exploring the thoughts, feelings and actions of individuals in relation to a particular phenomenon that they have experienced or are experiencing (Charmaz & Belgrave, 2012; Kvale & Brinkmann, 2009). This is particularly the case in longitudinal interview studies, where trust and rapport can develop between the researcher and interviewees, and opportunities exist to further clarify and explore issues raised in an interview in subsequent conversations (Grinyer & Thomas, 2012; Saldana, 2003). They are well suited to grounded theory research, which focuses on “what people do and the meanings they make of their actions and of the situations in which they are involved” (Thornberg & Charmaz, 2013).

In the current study, interviews are viewed as more than sites of data collection (Briggs, 2007; Mann, 2011; Talmy, 2010 & 2011). A critical approach to interviewing is adopted, which takes the view that interviews are a means of accessing participant accounts of their lived experiences, rather than factual reports of these experiences (Kvale & Brinkmann, 2009; Roulston, 2010). Interview data are thus interpreted as subjective accounts, rather than factual or truthful reports of the thoughts, feelings, and actions of participants. The focus is on the theoretical plausibility of participants’ accounts, compared with the rest of the data, rather than a focus on the accuracy or consistency of these accounts (Charmaz & Belgrave, 2012).
5.2.3 Grounded Theory methods

Grounded theory is a well-established approach to explain how individuals account for and make sense of their experiences, perceptions, and actions as they interact with their social worlds (Charmaz, 2006; Corbin & Strauss, 2008). Originally developed by the sociologists Glaser & Strauss (1967), ground theory-based approaches are aimed at developing interpretations and building theoretical frameworks that are grounded in the data, as the name suggests. Grounded theory methods are now widely used across a range of disciplines (Bryant & Charmaz, 2007), and have also been adopted, at least partially, by a number of researchers in language testing (Fox, 2004; Fulcher, 1996; Pill, 2013; Plakans & Burke, 2013).

According to Charmaz and Belgrave, in-depth interviewing and grounded theory are “emergent methods that combine flexibility and control” (2012, p. 348). Grounded theory methods rely on the researcher shaping the direction of interviews and analysis in order to tap into underlying processes without forcing particular outcomes. While the researcher signals their area of interest to interviewees and seeks to direct the flow of conversation around areas of interest, in-depth interviewing that is open-ended in nature also involves encouraging interviewees to speak freely and to introduce topics that they see as relevant. It is thereby thought to allow unexpected information and ideas to emerge (Kvale & Brinkman, 2009). In the current study, as described further below in the ‘Procedures’ section, the aim of the interview design was to enable unanticipated directions and new insights to emerge. The four migrant participants were told at the outset that the researcher was broadly interested in their test experiences and their thoughts about the test and score requirements, but that they should speak freely about their migration experiences.

Grounded theory methods also require that the researcher remain open to the data in order to allow for new and unexpected possibilities, and to avoid imposing pre-conceived ideas and pre-existing concepts onto the data (Bryant & Charmaz, 2007; Charmaz, 2006; Charmaz & Belgrave, 2012, Corbin & Strauss, 2008). As noted by
Charmaz and Belgrave (2012) and Thornberg and Charmaz (2013), the way in which this requirement is conceptualised varies according to different epistemologies underlying variants of grounded theory. Classic grounded theory, as originally conceived by Glaser and Strauss (1967), is underscored by a postpositivist, or objectivist epistemology, whereby the perspectives and beliefs of researchers are assumed to represent a source of bias which must be minimised as much as possible in order for ‘objective’ meaning to be located and extracted from the data. As such, classic grounded theory calls for engagement with relevant literature to be delayed until the analysis is near complete, allowing researchers to “approach the data uncontaminated by preconceived notions and theories” (Charmaz & Belgrave, 2012, p. 349).

More recently, a postmodern paradigm has been applied to grounded theory approaches (Clarke, 2005; Charmaz, 2006). This paradigm shift mirrors a postmodern turn in thinking more broadly across the social sciences (Corbin & Strauss, 2008), with which the current study is aligned. In a postmodern epistemology, multiple realities exist, and these realities are contingent, situated and fragmented, co-constructed in interactions between individuals and the social worlds they inhabit. In postmodern approaches to grounded theory, researchers actively engage with the literature and existing theoretical frameworks as a means of viewing the data through “multiple possible lenses”, not of “forcing the research into preconceived categories” (Thornberg & Charmaz, 2013, p. 163). Rather than seeking to exclude the researcher’s expertise and perspective from the process of analysis, postmodern grounded theory approaches thus emphasise the value of informed meaning making and knowledge construction in research.

In terms of an approach to coding and analysis, Charmaz and Belgrave (2012) characterise grounded theory as “an inductive, comparative, iterative, and interactive method” of building theory from data (p. 347), which relies on comparative analysis. Grounded theory methods involve three overlapping phases of analysis, initial coding, focused coding, and deriving theoretical concepts. Across these three phases, rigorous comparative analysis is undertaken between previously collected data and new data,
whereby data are compared with data and codes with codes in an iterative process that aims to ensure the best possible fit between themes, theoretical concepts, and data.

As described by Charmaz (2006), the process of grounded theory analysis begins with initial coding, which involves identifying potential themes and generating preliminary theme labels or codes. The second phase, focused coding, involves identifying patterns, similarities and differences across the themes initially coded, refining labels and theme definitions, and verifying or revising refined themes by working back and forth through the data. The focussed coding phase culminates with the development of final codes or themes. The third and final phase of analysis, deriving theoretical categories, involves identifying and explaining relationships between key themes to derive concepts that contribute to new theoretical advances. Details of the coding and analysis undertaken in the current study across these three iterative and overlapping phases is described further below, in the Procedures section.

To conclude this section of the chapter, a summary of the methodological underpinnings of the data collection and analysis procedures followed in the current study, as presented above, is provided below.

5.2.4 Summary

A longitudinal, in-depth interviewing and grounded theory-based approach that involved generating themes through iterative and overlapping phases of analysis was taken in the current study, to identify and explain how individuals account for and make sense of their experiences, perceptions and actions as they interact with the language test and score requirements for permanent residency. The epistemological position adopted in the current study is postmodern, in the sense that participant accounts and the experiences they depict are thought to be shaped by perceptions that are contingent, situated and potentially changing over time. Interview data are thus assumed to represent subjective accounts of individuals’ test experiences, rather than factual reports.
The methodological approach taken in the study allowed the researcher to develop trust with participants, to detect changes in their orientations towards their experiences over time, to iteratively interrogate interpretations and understandings of accounts, with participants and against the data, and to verify the robustness of emerging themes throughout the process of analysis. As noted above, the inclusion of an alternative stakeholder perspective, that of teachers, offered a further means of interrogating both the plausibility of migrant accounts and the appropriateness of researcher interpretations of these accounts.

5.3 Procedures

This section describes participant details, data collection, preparation, and analysis procedures followed in the study.

5.3.1 Participants

As noted in the introduction, five perspectives were examined in the study, that of four migrant test takers and a fifth ‘teacher’ perspective, derived from interviews conducted with seven English language teachers. The four test taker participant details are briefly described under the heading ‘Migrant participants’. Each individual is introduced separately in more detail in the next chapter, where findings are reported, to provide background and context to their individual accounts of their test experiences. Teacher participant details are described in the current section, further below. Pseudonyms are used throughout for all participants.

5.3.1.1 Migrant participants

Four adult migrants participated in the study, three female (Ana, Sharon, Yuki) and one male (Erfan). They all joined the study at various times in 2012, Sharon in March, Erfan and Yuki in August and Ana in September. All four migrants had studied at different universities in Australia. Ana, Erfan, and Sharon were on temporary post-study work visas after having completed postgraduate studies at Australian institutions. Yuki was in
her final year of an undergraduate nursing degree, and was on a temporary student visa. At the time the first interviews were conducted, Ana and Erfan were in the process of preparing for their first English test attempts for permanent residency, and Sharon and Yuki had already commenced their test taking efforts.

Sharon and Yuki were located via professional contacts at the researcher’s institution, and Ana and Erfan via professional contacts at a private English language college. They were each invited in person by the researcher to join the study and all remained involved until they had fulfilled the language test requirements needed to become eligible to apply for permanent residency.

Background information and an overview of each individual’s test attempts for permanent residency is provided in Table 5.1, below.
Table 5.1. Overview of migrant participant details

<table>
<thead>
<tr>
<th></th>
<th>Ana</th>
<th>Erfan</th>
<th>Sharon</th>
<th>Yuki</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country of origin</td>
<td>Colombia</td>
<td>Iran</td>
<td>Sri Lanka</td>
<td>Japan</td>
</tr>
<tr>
<td>First language</td>
<td>Spanish</td>
<td>Persian</td>
<td>Sinhala</td>
<td>Japanese</td>
</tr>
<tr>
<td>Year of arrival in Australia</td>
<td>2009</td>
<td>2009</td>
<td>2007</td>
<td>2008</td>
</tr>
<tr>
<td>Year of completion of Australian tertiary studies</td>
<td>2011</td>
<td>2011</td>
<td>2011</td>
<td>2012</td>
</tr>
<tr>
<td>Age at graduation</td>
<td>Mid 30s</td>
<td>Late 20s</td>
<td>Mid 20s</td>
<td>Early 30s</td>
</tr>
<tr>
<td>Field of study</td>
<td>Information Technology</td>
<td>Information Technology</td>
<td>Immunology</td>
<td>Nursing</td>
</tr>
<tr>
<td>Minimum IELTS score for PR*</td>
<td>IELTS 7</td>
<td>IELTS 7</td>
<td>IELTS 8**</td>
<td>IELTS 7/OET 8#</td>
</tr>
<tr>
<td>Date of joining study</td>
<td>September, 2012</td>
<td>August, 2012</td>
<td>March, 2012</td>
<td>August, 2012</td>
</tr>
<tr>
<td>Number of test attempts for PR</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Time between first and last test attempt</td>
<td>11 months</td>
<td>6 months</td>
<td>3 months</td>
<td>8 months</td>
</tr>
</tbody>
</table>

*PR = permanent residency

**Sharon required a higher score than the other three migrants because she went directly from school in Sri Lanka to university in Australia, whereas the others were able to gain credit in the Points Test for work experience in their home countries. As a result, they only needed 10 points for English to meet the Points Test pass mark, whereas Sharon needed 20 points.

# The Occupational English Test (OET), also accepted for skilled migration in Australia, is a specific purpose test for health professionals.

^ This figure refers to Yuki’s most recent efforts to gain permanent residency in Australia. She had also made several attempts at IELTS and the OET to meet the score requirements for skilled migration purposes prior to commencing tertiary studies in Australia, as detailed in the next chapter.
5.3.1.2 Teacher participants

Seven English language teachers participated in the study. All were very familiar with IELTS, having taught IELTS preparation courses for several years. These participants were accessed from two English language colleges, both affiliated with different universities, and were recruited by email via professional networks. Of the seven teachers, two were directors of studies, Claire and Jane, at each of the two colleges. As directors of studies, Claire and Jane were responsible for overseeing the English language and test preparation courses offered at their institutions rather than classroom teaching. Both were very familiar with the language requirements for permanent skilled migration in Australia, as demand was high at their institutions from individuals needing IELTS scores for the purpose of meeting the points test requirement, and both had worked as IELTS examiners for more than 10 years.

At Claire’s college, an IELTS preparation course aimed at scores of IELTS 8 had been introduced in response to a surge in demand following the introduction of these scores into migrant selection criteria. At Jane’s institution, the extra demand was absorbed into existing IELTS classes. Of the five remaining teachers, four worked at Claire’s college: Jenny, Michelle (also an IELTS examiner), Paul and Ruth, and they were all involved in teaching on IELTS preparation courses aimed at IELTS 8. The remaining teacher, Kate, who belonged to the same institution as Jane, was responsible for coordinating teaching for the IELTS preparation courses on offer, had experience teaching migrants seeking scores of IELTS 7 and 8 for permanent residency and was also training to become an IELTS examiner.

5.3.2 Data collection

As noted above, several interviews were conducted with each of the four migrant test takers in this study, compared to single interviews with each of the seven teacher participants. For this reason, data collection procedures are described separately for the migrant and teacher participants.
5.3.2.1 Migrant interviews

Data were collected over a period of around 18 months, beginning in late March 2012 and concluding in early October 2013. Within this period, a series of in-depth, open-ended interviews were conducted with each of the four participants as they went through the process of repeated language test attempts in their efforts to gain permanent residency. Interviews were conducted at various locations, including the researcher’s office, cafes, libraries, and parks, to suit the preference of participants. Initial interviews with each participant were guided in line with the research aims, outlined in chapter 1 and summarised at the beginning of this chapter. The researcher explained at the outset that the focus of the study was on participants’ test experiences in the context of their migration journey, their perceptions of the language test requirement, as well as how they felt about their language and their migration experiences more generally. After this brief orientation, participants were invited to talk freely about their migration experiences and life in Australia.

The schedule and duration of interviews is shown in Table 5.2, below. In total, 17 interviews were conducted with the migrant participants, which yielded 11 hours of recorded data and over 120,000 transcribed words. At the outset of the study, the intention was to conduct interviews with participants at regular intervals during their period of transition from temporary visas to permanent residency, either monthly or bi-monthly, as well as before and after each test attempt for permanent residency. In reality, scheduling interviews on a regular basis and around test taking episodes was more difficult than anticipated. The pressure of juggling already busy working lives combined with the pressures and demands of test taking and immigration meant that monthly or bi-monthly interviews were not always feasible for participants. Furthermore, at various points during the study, each of the participants dropped out of contact for a period of time, because, they told the researcher afterwards, they had been feeling too busy, too unhappy, or too preoccupied with the uncertainty of their situations to talk about their experiences. As a result, the interview schedule was determined by the availability and the willingness of participants as they manoeuvred through a difficult and taxing period in their lives, and varied from participant to
participant. Yuki was interviewed three times over a 6-month period, Sharon four times over an 11-month period, and Ana and Erfan each five times over a 13- and 9-month period, respectively.

Between interviews, regular contact was maintained with each of the participants over the period of the study, via email and phone. This regular contact was an important part of getting to know participants and of building mutual trust and rapport. It also allowed the researcher to draw on a broader context for understanding and interpreting interview accounts. A summary of the interview schedule and duration is provided in Table 5.2, below.

**Table 5.2. Interview schedule and duration**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Test dates</th>
<th>Interview date</th>
<th>Duration (m:s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ana</td>
<td>September 10, 2012</td>
<td>September 10, 2012</td>
<td>57:23</td>
</tr>
<tr>
<td></td>
<td>October 12, 2012</td>
<td>November 19, 2012</td>
<td>48:00</td>
</tr>
<tr>
<td></td>
<td>December, 2012</td>
<td>May 16, 2013</td>
<td>26:14</td>
</tr>
<tr>
<td></td>
<td>June, 2013</td>
<td>July 9, 2013</td>
<td>29:19</td>
</tr>
<tr>
<td></td>
<td>September, 2013</td>
<td>October 10, 2013</td>
<td>34:16</td>
</tr>
<tr>
<td>Erfan</td>
<td>August 3, 2012</td>
<td>August 3, 2012</td>
<td>47:55</td>
</tr>
<tr>
<td></td>
<td>October &amp; November, 2012</td>
<td>December 6, 2012</td>
<td>36:07</td>
</tr>
<tr>
<td></td>
<td>February, 2013</td>
<td>February 6, 2013</td>
<td>60:09</td>
</tr>
<tr>
<td></td>
<td>April, 2013</td>
<td>April 18, 2013</td>
<td>36:20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 7, 2013</td>
<td>34:51</td>
</tr>
<tr>
<td>Sharon</td>
<td>February, 2012</td>
<td>March 20, 2012</td>
<td>19:32</td>
</tr>
<tr>
<td></td>
<td>April, 2012</td>
<td>April 27, 2012</td>
<td>18:32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 4, 2012</td>
<td>23:55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>February 25, 2013</td>
<td>24:53</td>
</tr>
<tr>
<td>Yuki</td>
<td>May, July (2), 2012</td>
<td>August 20, 2012</td>
<td>43:48</td>
</tr>
<tr>
<td></td>
<td>Nov, Jan (2), 2012</td>
<td>March 6, 2013</td>
<td>57:29</td>
</tr>
</tbody>
</table>
5.3.2.2 Teacher interviews

Seven one-on-one interviews were conducted between August and October, 2012, with each of the seven teachers. The interviews were conducted at participants’ places of work. Interviews ranged in duration from 17 to 38 minutes, with an average length of approximately 25 minutes, yielding 3 hours of data in total (just over 26,000 transcribed words). Interviews were semi-structured and focussed broadly on two main areas:

- The ways in which the new test score requirements for permanent residency impacted their teaching practices and experiences
- Their opinion about the test requirements for skilled migration, including their perceptions of test purpose

They were also asked:

- if there been a change in the kinds of students enrolling in their test preparation courses as a result of the new requirements;
- if the new requirements were having an impact on their teaching strategies and priorities and if so, in what ways;
- what advice they were giving students about how to prepare for and succeed on the test;
- if they believed the courses they were teaching were likely to equip students to achieve scores of 7 or 8 on IELTS (needed for permanent residency).

5.3.3 Data preparation

Interview data were audio recorded and transcribed by the researcher. Features such as false starts, hesitations, fillers (e.g. er, uh, um), laughter, repetitions and contracted word forms (e.g. you’re, cos) are included in the transcripts. Non-standard English features used by participants were also maintained in the transcripts. As already explained, the focus of the study was primarily on content of the interviews, and thus micro-details such as pause lengths and overlapping speech were not transcribed. Transcripts were labelled with participant pseudonyms, date, location, and duration of
interview, and lines were numbered. Samples transcript segments from the first interview with each migrant participant are included as Appendices 1 to 4. The conventions for the transcription of interview data are provided in Table 5.3, below.

Table 5.3. Transcription conventions

<table>
<thead>
<tr>
<th>Format</th>
<th>Example</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>underlining</td>
<td><em>I got 6.5 in every single band</em></td>
<td>Syllable stressed by speaker</td>
</tr>
<tr>
<td>single quotation marks</td>
<td><em>I told him ‘oh I just got here 15 days ago and we need money’</em></td>
<td>Speech reported by speaker</td>
</tr>
<tr>
<td>hyphen</td>
<td><em>he did the six mo- six month English course</em></td>
<td>Work incomplete, false start</td>
</tr>
<tr>
<td>dash</td>
<td><em>that’s what I- he feels so grateful</em></td>
<td>Speakers utterance incomplete</td>
</tr>
<tr>
<td>three dots</td>
<td><em>before one month...I used to live</em></td>
<td>Hesitation</td>
</tr>
<tr>
<td>comma</td>
<td><em>If you’re going to apply for your residency and you are married, and if your partner decides not to take the test</em></td>
<td>Marking off sense groups within speech flow (cf. written punctuation)</td>
</tr>
<tr>
<td>Full-stop</td>
<td><em>we just, we just had two bags full of dreams.</em></td>
<td>Indicating completion of an utterance</td>
</tr>
<tr>
<td>Question mark</td>
<td><em>how can I explain to you?</em></td>
<td>Upward, questioning intonation</td>
</tr>
<tr>
<td>@</td>
<td><em>He could barely say hello @ @</em></td>
<td>Laughter (@ = 1 beat; @@ = 2 beats, etc)</td>
</tr>
<tr>
<td>Round brackets</td>
<td><em>we have to show to the Australian government this amount</em> (she writes down $40,000)</td>
<td>Researcher commentary</td>
</tr>
<tr>
<td>Double round brackets</td>
<td><em>((sigh))</em></td>
<td>Peripheral sounds and non-verbal actions</td>
</tr>
<tr>
<td>Square brackets</td>
<td><em>we were like move on, move on, study pals [XXX]</em></td>
<td>Unclear syllable/word</td>
</tr>
</tbody>
</table>

5.3.4 Data analysis

All transcripts were imported into NVivo for Mac (version 10.2.2), with separate NVivo projects set up for analysis of the migrant and teacher datasets. Analyses of migrant and teacher interview data involved generating themes from iterative phases of coding, consistent with a grounded theory approach, as described in the methodology section, above. This approach involved two phases, initial and focused coding, through which final themes and were generated, and a third phase, deriving theoretical concepts,
based on the combined analyses of thematic relationships identified in the migrant and teacher datasets.

The process of generating final themes and deriving thematic relationships was first conducted on the migrant dataset across two stages, which are set out below under the heading ‘Analysis of migrant data’. As noted previously, the teacher perspective was intended to serve as a form of triangulation, to verify the plausibility of migrant accounts and researcher interpretations of these accounts. In light of this, the results of the analysis of migrant data provided the starting point for analysis of the teacher dataset, as explained further below under the heading ‘Analysis of teacher data’.

5.3.4.1 Analysis of migrant data

As noted above, analysis of migrant data was conducted in two stages. The first stage involved generating key themes from the entire migrant dataset, rather than on a case-by-case basis. These key themes provided the framework through which to examine each individuals’ perceptions, decisions, and actions in relation to their test experiences, which constituted stage two of the analysis. This second stage involved examining each migrant case individually, and culminated with the third and final phase of the analysis of migrant data, deriving thematic relationships to explain how test taker perceptions shape decisions and actions, and vice-versa, in each individual case. An overview of each stage is provided, below, and the findings from each stage are presented in chapter 6.

Stage one – generating a common set of key themes

As mentioned above, the first stage of analysis involved iterative rounds of initial coding and focused coding across the entire migrant dataset, rather than on a case-by-case basis. This resulted in the generation of key themes that captured a common set of perceptual and action processes that individuals engaged in to make sense of and respond to their test experiences. Although the phases of initial and focussed coding involved in generating these themes were iterative and overlapping, they are described separately below, for clarity.
Initial coding

In the initial coding phase, the researcher examined interview transcripts line-by-line, identifying segments of the data in which migrants described or alluded to perceptions, feelings, decisions and actions associated with the test and their experiences of it. Segments consisted of single words, partial lines, whole lines or any number of consecutive lines of the transcripts. Throughout the phases of initial and focused coding, the data were scrutinized with questions such as ‘what is happening in the data?’, ‘what is the participant’s main concern?’, ‘what type of action or process is involved here?’, and ‘what is this a consequence of?’ (Thornberg & Charmaz, 2013).

Coding were aimed at capturing processes that participants engaged in to situate and give meaning to their test-related experiences. The researcher decided where relevant segments of data began and ended for coding purposes, and irrelevant data were discounted. For example, during the course of interviews, participants at times spoke about topics that were unrelated to the research focus of the current study, or were beyond the scope of the current study, and these data were not coded.

Table 5.4, below, provides an example of initial coding, including the researchers’ notes in relation to specific segments of data. As can be seen, the codes were detailed and specific in the initial coding phase. The questions listed above helped to generate ideas about how to characterise these processes more broadly, and helped highlight potential relationships between pieces of data, which informed the subsequent generation of final themes.
Table 5.4. Example of initial coding process

<table>
<thead>
<tr>
<th>Initial codes</th>
<th>Data (S=Sharon; R=Researcher)</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeking expert advice in test preparation</td>
<td>61. S: Yeah, so um I’ve done this um preparation course</td>
<td>‘their’ writing style – specific to the test, contrasts with writing style relevant to study demands. Lines 86-90 indicate she took prep course to identify attributes of IELTS 8 writing performance/identify why she didn’t achieve these scores</td>
</tr>
<tr>
<td></td>
<td>62. R: Uh ha where did you-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>63. S: for the IELTS here and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64. R: Where did you do the preparation course?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>65. S: That was at (redacted)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>66. R: OK</td>
<td></td>
</tr>
<tr>
<td>Identifying a specific IELTS writing style</td>
<td>67. S: Um and I noticed that the reason that their essay writing style is different</td>
<td></td>
</tr>
<tr>
<td></td>
<td>68. R: Mm hm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>69. S: like it’s very specific but then I think-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70. R: Different to what you’ve learnt in school?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>71. S: Yes, yes.</td>
<td></td>
</tr>
<tr>
<td>Seeking expert advice in test preparation</td>
<td>82. R: Uh ha, and so when, you did you do the preparation course right before you sat the test</td>
<td>Response to previous score outcome, links to ‘Identifying unfamiliarity with essay style’ (lines 67 – 71)</td>
</tr>
<tr>
<td></td>
<td>83. again?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>84. S: I- no, after</td>
<td></td>
</tr>
<tr>
<td></td>
<td>85. R: After?</td>
<td></td>
</tr>
<tr>
<td>Planning repeat test attempt</td>
<td>86. S: because I want to redo it again</td>
<td>Expert advice sought because of belief that test success rests on a particular formula, links to ‘Identifying unfamiliarity with essay style’ (lines 67 – 71)</td>
</tr>
<tr>
<td>Linking expert advice to achieving score of 8</td>
<td>87. R: Yep</td>
<td></td>
</tr>
<tr>
<td></td>
<td>88. S: so that I can get 8 in every band.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>89. R: Mm hm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90. S: Yeah</td>
<td></td>
</tr>
</tbody>
</table>

Initial codes, such as those shown in Table 5.4, were iteratively revised as further interview transcripts from the same and other migrants were examined. The process of refinement continued throughout the analysis process until a set of focused codes were derived, which could be applied across the dataset, as described further under the heading ‘Focused coding’, below.
Focused coding

Focused coding was an iterative process that involved going back and forth between initial codes within and across participant interview transcripts to identify similarities, patterns and common themes, and to draw out differences between groups of codes in order to arrive at a set of common themes. Emerging themes were re-examined and revised throughout the focused coding stage, to maintain a close and robust link between data and codes.

As mentioned above, the aim was to establish a final set of key themes that described the perceptual and action processes that occurred across the four cases as migrants interacted with the test and score requirements over the course of their test taking journeys, rather than to capture individual’s particular thoughts, feelings, or specific actions, per se. The ultimate aim of stage one was to generate themes that would serve as a framework through which to explain the role of individual subjectivities in shaping how migrants interacted with and responded to the test and score requirements, on a case-by-case basis, in stage two.

To illustrate the process of moving from initial codes to final themes, which occurred during the focused coding phase, an extract from the data shown above in Table 5.4, lines 67-71, will be taken as example. The data in Table 5.4 is from the first interview with Sharon. As per the researcher’s annotation next to lines 67-71, Sharon’s later explanation of why she took the IELTS preparation course, in lines 86-90, indicates that she was using the IELTS preparation course to identify the reason why she failed to achieve IELTS 8 on a previous test attempt, and she decides this is because “their essay writing style is different” (line 76), which is initially coded as ‘Identifying differences between own language use and IELTS writing test construct”. An examination of further transcripts from subsequent interviews indicated that other participants tended to explain their lack of score success in terms of a deviation from the required IELTS writing conventions, as they perceived them, as shown in the example below from an interview with Ana:
“The problem is I was writing too much, too many paragraphs” (Ana, interview 4, line 6)

In the example above, Ana is not identifying a specific IELTS writing style, but she seems to be doing the same thing as Sharon in the example provided in Table 5.4 - attributing meaning to her previous score outcomes. Further examination of the data showed that different participants attributed unsuccessful (as well as ultimately successful) score outcomes to different factors in a number of ways, including by specifying attributes of ideal test performances, as Sharon did, and by evaluating their own performances, as Ana is doing in the examples provide above. These observations led to the generation of a key theme ‘interpreting score outcomes’, intended to encompass the underlying processes participants engaged in when attributing different meanings to scores.

In summary, stage one of the analysis led to the generation of five key themes that were intended to capture common perceptual and action processes central to migrants’ accounts of their test experiences. These key themes - ‘Self-appraising’, ‘Interpreting score outcomes’, ‘Understanding target scores’, ‘Responding affectively’ and ‘Practising agency’ - are presented, with a gloss and example data, at the beginning of the migrant perspectives section of the results in chapter 6.

**Stage two – explaining individual perspectives**

As noted above, the key themes generated through stage one of the analysis process constituted the framework through which individual perceptions, feelings, decisions and actions were examined on a case-by-case basis, in stage two of the analysis. Stage two also involved a third phase of analysis, which involved identifying and explaining interrelationships between themes in order to trace the interactions between individuals’ perceptions and actions, as manifest in each migrant case. In chapter 6, the way in which key themes and their interrelationships characterise individual perspectives is explained and demonstrated with thick descriptions of interview extracts in relation to each migrant’s experiences, so as to enable the researcher’s interpretations to be evaluated by the reader, thereby enhancing the transferability and trustworthiness of findings.
5.3.4.2 Analysis of teacher data

As mentioned above, the results of the analysis of migrant data provided the starting point for analysis of the teacher dataset. This was intended to enable the teacher perspective to serve as a form of triangulation, as previously noted. As with migrant data, analysis of teacher data involved processes of initial and focused coding. However, in this case, key themes derived from the migrant dataset, modified to reflect the alternative perspective of the teacher group of participants (for example, the theme ‘self-appraising’ was modified to ‘appraising test takers’), were taken as initial codes.

Similar to the coding procedures outlined above in relation to the migrant dataset, teacher interview transcripts were examined line by line, relevant segments of data were identified, and themes were applied. As with the migrant dataset, relevant segments consisted of single words, partial lines, whole lines or any number of consecutive lines of the transcripts, and the beginning and end point of segments were determined by the researcher. The analysis process was conducted iteratively, with the researcher moving back and forth within and between transcripts to identify similarities and patterns that were consistent with the amended themes, and to identify and explain differences. Further revisions of theme labels and definitions were made to ensure a close fit between themes and data, and new themes were derived where the coverage of themes generated from the migrant data was inadequate to capture the nuances of the teacher perspective.

This analysis process generated five key themes: Appraising test takers’ English use; understanding score meanings; characterising the test taker perspective; negotiating score meanings; and losing agency in the classroom. The first two of these themes closely mirrored the themes derived from the migrant dataset, and the latter three were specific to the teacher perspective. In chapter 7, these themes, derived from the teacher analysis, are presented, defined, and exemplified, and relationships between themes, which constitutes findings from the third phase of analysis of the teacher dataset, are identified and explained.
Whereas in chapter 6, key themes which apply to the entire migrant dataset are first presented and defined, then elaborated in relation to each individual migrant case in four separate sections, chapter 7 is organised around themes, rather than individual participants. This reflects the aim associated with the teacher analysis, to provide a means of verifying the credibility of migrant accounts by presenting an alternative stakeholder group perspective.

To complete the third phase of analysis of a grounded theory approach, Deriving theoretical concepts, findings from the migrant and teacher datasets are discussed in combination in chapter 8, where the key themes and thematic relationships identified in chapters 7 and 8 are modelled to characterise test impact as a dynamic process, with test taker subjectivities constituting the nexus.

To conclude the Procedures section, and this chapter, details of the measures taken to ensure trustworthiness are presented below.

5.3.5 Ensuring trustworthiness

Trustworthiness in qualitative research involves three dimensions, credibility, dependability, and transferability (Lincoln and Gruba, 1995). In terms of credibility, as already described above, an alternative stakeholder perspective was included in the study, to provide a means of verifying the plausibility of migrant accounts and the appropriateness of researcher interpretations of these accounts. Further to this, due to the longitudinal nature of the study, the researcher was able to check understandings and interpretations of previous interview content with migrant participants in subsequent interviews, and against subsequently collected data. In terms of dependability, the appropriateness of key themes and their applicability to the dataset was verified through second coding of approximately 10% of the data, conducted independently by two colleagues. Differences in applying key themes was discussed until agreement was reached, and theme definitions were amended to resolve ambiguities, to the extent possible. To enhance the transferability of findings, in the
results chapters, extracts are accompanied by thick descriptions of coding decisions, as noted, to enable the reader to evaluate the researcher’s judgments.

5.4 Conclusion

In this chapter, the research methods adopted to address the aim of the study were described. The chapter began by explaining the methodological approach adopted in the study, including its theoretical and epistemological underpinnings. The study procedures were then described, including participant details and the data collection, preparation and analysis procedures followed. Findings from the study are presented over the next two chapters. In Chapter 6, the themes derived from the migrant dataset are first described, and then each migrant case is described separately. Chapter 7 presents the perspectives of the seven teacher participants concerning the use of English test scores for migrant selection purposes are described, with a focus on how they characterise migrant test takers seeking permanent residency and the experiences of these test takers.
CHAPTER 6: INDIVIDUAL EXPERIENCES - ANA, ERFAN, SHARON AND YUKI

6.1 Introduction

This chapter presents findings from the grounded theory analysis of interview data collected from the four migrants who participated in the current study: Ana, Erfan, Sharon, and Yuki. As set out in the previous chapter, analysis involved identifying themes that characterised participants’ accounts of their perceptions, feelings, decisions, and actions in relation to their test experiences as they negotiated the score requirements needed to secure permanent residency in Australia. This chapter and the next (educator perspectives) provide the basis for a discussion in chapter 8 of the theoretical framework derived from these findings, the dynamics of test impact.

Throughout the periods of their involvement in this study, each participant experienced repeat test encounters. As demonstrated in this chapter, across these encounters, key themes of self-appraising, interpreting score outcomes, understanding target scores, and practising agency were continually revisited as participants engaged in iterative and interrelated processes of comparing and evaluating new and previous experiences over time to try to make sense of the test and to better direct their actions towards achieving their score goals. Participants were heavily invested, both emotionally and materially, in the life they had built in Australia and in their efforts to acquire permanent residency, which for each of them hinged ultimately on test success. Test stakes were thus very high, and responding affectively to test experiences over time represented a further key theme, which interacted with self- and test-related perceptions to shape decisions and actions throughout test-taking journeys.

In this chapter, key themes, derived from the entire migrant dataset, are first briefly defined. Following this, the experiences of each individual, Ana, Erfan, Sharon, and Yuki, are presented in four separate sections. In each case, individuals' experiences of learning and using English are first set out as a means of shedding light on the self-perceptions and beliefs about language learning that underlie each person's perceptions of their test
experiences. As demonstrated in this chapter, each individual oriented to the test by generating understandings of target scores, which are influenced by how they perceive test purpose throughout their test taking journey and how they perceive the relevance of the test to the demands they face as migrants living and working in Australia. They then conduct self-appraisals of their English abilities in relation to these demands and in relation the English usage they associate with scores of IELTS 7 and 8. These interacting processes of self-appraising and understanding target scores prompt affective reactions and shape each person's interpretations of their own score outcomes, which, in turn, serve as a means to iteratively evaluate prior self-appraisals and understandings, and to inform decisions and actions as individuals proceed through repeat test encounters in an effort to achieve their score and migration goals.

The chapter concludes with a summary section, where it is argued that subjective experiences and time as variables interact to influence score meanings and to shape test consequences as individuals negotiate the constraints and uncertainties associated with being subjected to the testing regime in this policy context. In chapter 8, the way in which these iterative cycles of perceiving, feeling and acting play a role in negotiating different test impacts across the four individual cases is established, and the validity and fairness issues raised by these test taker experiences of the test in the context of Australia's skilled migration policy, supported by accounts provided by teachers (reported in chapter 7), are discussed.
6.2 Key themes

As mentioned above, five key interrelated themes were identified as characterising participants’ accounts of perceptions, feelings, decisions and actions in relation to the test: Themes ‘self-appraising’, ‘interpreting score outcomes’ and ‘understanding target scores’ relate to participants’ perceptions, ‘responding affectively’ to feelings and ‘practising agency’ to decisions and actions. These themes are presented and defined below.

**Self-appraising English abilities**

The theme ‘self-appraising’ refers to participants’ self-perceptions and self-evaluations of their English skills, test performances, and capacities to use English effectively in various contexts, including on the test, to realise their goals.

**Interpreting score outcomes**

The theme ‘interpreting score outcomes’ refers to participants’ accounts of their efforts to attribute meaning to the test scores they receive, in terms of their English skill level or their test performance quality, for example.

**Understanding target scores**

The theme ‘understanding target scores’ involves two separate but interrelated aspects. The first aspect relates to participant perceptions of test purpose within skilled migration policy, and the second to their perceptions of the test construct – their interpretations of the meaning of IELTS 7 or 8 in terms of the skills and performance attributes they associate with the test – and these scores and their perceptions of the relevance of these to the English demands they face living and working in Australia.
Responding affectively

The theme ‘responding affectively’ refers to expressions of feelings and emotions, moods and attitudes that are associated with self-perceptions, and perceptions of the test, score outcomes and target scores.

Practising agency

The theme ‘practising agency’ refers to participants’ accounts of purposeful decision-making and action responses in relation to their English language learning and test taking efforts. Two main types of agency were identified in the analysis: Practising strategic agency, whereby individuals felt able to engage in directly goal-oriented decisions and actions (they felt able to set and adhere to an overall plan of action representing a direct pathway to goal achievement); and Practising tactical agency, whereby individuals make decisions and take actions aimed at creating pathways around obstacles that cannot be immediately overcome, or in response to obstacles or setbacks that arise unexpectedly.

6.3 Overview of individual migrant cases

The experiences of the four migrants involved in this study, Ana, Erfan, Sharon and Yuki, are described over the remainder of the chapter. Each case begins with an introductory subsection, in which the individual’s background and migration trajectory is outlined. Following this, a summary of their accounts of ‘Experiencing learning and using English’ is provided. In the case of Ana and Erfan, who were both interviewed prior to commencing their test attempts for permanent residency, a short subsection describing their initial strategic approach to the test is then presented under the heading ‘Before test attempt 1’. Then, participants’ accounts of their repeat test encounters are described under the main heading ‘Experiencing repeat test encounters’. This subsection is organised according to how interviews coincided with test attempts. For example, interviews were conducted with Ana and Sharon after each test encounter, so their experiences are presented under subheadings ‘After test attempt 1’, ‘After test attempt 2’ etc. Where more than one test attempt takes place in between interviews, findings are grouped accordingly (for example, ‘After the first two test attempts’, etc.,
in Erfan’s case and ‘After the first three test attempts’, etc., in Yuki’s case). To conclude each case, a summary of individual experiences is provided.

Conventions

Data from interview transcripts are presented in this chapter in italics. Most researcher turns and back channelling have been omitted, although the researcher’s interview questions are included where necessary, in italics and within square brackets after the initial ‘R:’. References to interview transcripts are provided in square brackets at the end of each quote. These references include interview number and line numbers. Omission of a section of text within a participant’s turn is indicated with an ellipsis (**`). When this occurs on a separate line, a whole turn or turns have been omitted. The symbol ‘…’ represents a short pause. Researcher commentary on the data is presented within the text in normal font within square brackets. These conventions are illustrated in the example from the first interview with Erfan, below:

“I just see it [the use of IELTS 8 in selection criteria] as a very fine filter. I don’t want to think about it in this way but because of, it gets in my mental way, yeah… but yeah, to me it works, it’s just like a filter [R: what’s being filtered?] The number of people *** Apparently, they think they can’t say that we don’t like people I don’t know, older than 40 or from these countries or whatever”

[Interview 1, 503-509]

Samples transcript segments from the first interview with each migrant participant are included as Appendices 1 to 4. Segments include all of the data shown in extracts related to the first interview in each case. Each individual case will now be presented, in turn, below.
6.4 Ana

6.4.1 Introduction

Ana arrived in Australia with her husband in 2009, with the plan to become an English speaker, gain employment in her field and apply for permanent residency as a skilled migrant. She was in her mid-30s at the time, and had a bachelor-level qualification in Information Technologies with several years of experience working in her profession in Colombia, her home country. Ana identifies Spanish as her mother tongue. Neither she nor her husband had any experience learning English prior to their arrival. For their first month in Australia, they stayed with a Colombian family who arranged cleaning jobs for them with a Spanish speaking employer.

After six months, Ana enrolled in a 20-week English language course, which she completed while working as a cleaner. At the end of the 20 weeks, Ana felt her English had improved enough to enable her to seek out a better job. She left the cleaning position and started work in a customer service role in a large supermarket, which required her to use English daily with colleagues and customers. Although she was gaining confidence in her English skills, Ana still felt they were inadequate to pursue skilled work in her profession and so decided to enrol in a master’s degree at university. Ana had satisfied the English language requirements for university entry with her completion of the 20-week English language course, and so did not need to take IELTS.

Ana continued working in the same supermarket position on a part-time basis while completing her master’s degree. At first, she experienced English-related difficulties in her studies but she felt that she managed to overcome these over her time at university. By the time she graduated, she felt she had acquired legitimacy as an English speaker and she began attempting to find work in her profession. Ana submitted several job applications but they were all rejected prior to the interview stage. Frustrated and under financial pressure, she began looking for alternative work. She found a full-time position as an administrator in a freight company, where she was working over the period of her participation in this study.
Ana began taking IELTS for the purpose of acquiring permanent residency in October 2012. The first interview with Ana was conducted a month earlier, in September 2012. By that time, she had fulfilled all of the points test criteria besides English language. She needed scores of IELTS 7 to submit an expression of interest to apply for a permanent skilled visa. Ana achieved these scores on her fourth test attempt and immediately submitted an expression of interest. She was granted permanent residency in January 2014; just over a year after her test efforts began.

Before moving on to Ana’s repeat test encounters, key themes related to Ana’s account of experiences of learning and using English prior to her first test attempt for permanent residency are presented below. Ana’s account of these experiences illustrates her sense of becoming a legitimate English speaker before her first test attempt, and provide background to the next sections, ‘Before test attempt 1’, in which Ana’s initial orientation to the test requirements is outlined, and ‘Experiencing repeat test encounters’, in which Ana’s interpretations of and responses to test experiences in the context of her repeated efforts to achieve her score goal are detailed. To conclude the section on Ana, a summary of her experiences is provided.

6.4.2 Experiencing learning and using English

As mentioned, Ana arrived in Australia with no English language knowledge, with the intention of learning the language, gaining skilled employment and remaining in the country permanently. As demonstrated in this subsection, she attributes her lack of English to factors outside her control - a lack of language learning opportunities during childhood - and attributes her successes in overcoming these obstacles and becoming a functional English user to her own hard work and determination since arriving in Australia. In locating the key to success internally, Ana assumes a sense of control over her language learning outcomes, which, together with the mindset that effort and perseverance yields success, also underlies Ana’s orientation to the test score requirements, described in the following subsection, 'Before test attempt 1'.
Ana’s perceptions of her experiences learning and using English during her first three years in Australia are illustrated in Extracts 6.4.1 to 6.4.5, below. These extracts are grouped according to the three main time frames that Ana refers to in her accounts of these experiences: before university (Extract 6.4.1), at university (Extracts 6.4.2 and 6.4.3), and after university (Extracts 6.4.4 and 6.4.5). The five key themes outlined in the introduction to this chapter are exemplified in this subsection, and key relationships between themes, relevant to her later accounts of repeat test encounters, are first demonstrated in relation to individual extracts within each time period, and then summarised at the end of this sub-section.

6.4.2.1 Before university

Extract 6.4.1

“As many Colombians, because, no one...how can I explain to you? When we were kids our parents had no vision about the world so they didn’t worry about our English, so we never studied English when we were kids. But they did their best, they didn’t know how useful is English, how useful English become in our life in the future. So, because we didn’t speak English I started working as a cleaner***After one year working as a cleaner I said enough, and I think my English is functional enough, and it’s time for me to move on”

[Interview 1, 225-243]

Extract 6.4.1, above, illustrates a relationship between themes ‘self-appraising’ and ‘practising agency’. In Extract 6.4.1, Ana expresses a negative self-appraisal of her English ability at the time of her arrival in Australia (we didn’t speak English). She attributes responsibility for not speaking English to her parents and Colombians of her parents’ generation, who failed to recognise the value of English during her childhood (our parents had no vision about the world ... so we never studied English). This indicates a lack of agency as Ana positions herself as disadvantaged due to past circumstances beyond her control. She then explicitly links negative self-appraising (because we didn’t speak English) to practising tactical agency (I started working as a cleaner); highly constrained by her lack of English, Ana is unable to plan a direct approach to achieving
her migration goal. Instead, she takes on unskilled work as a means of financially supporting herself so that, in the future, she can access opportunities to develop proficiency in English, which will enable her to pursue skilled employment opportunities and permanent residency.

After a year of working as a cleaner, during which time she also completed an English language course, Ana’s self-appraising of her English skills is more positive (my English is functional enough). This self-perceived English language development increases Ana’s sense of agency and prompts her to take action (it’s time for me to move on). Ana characterises ‘moving on’ as goal-oriented action, a step towards achieving her aims of establishing a career and remaining permanently in Australia, which suggests she is practising strategic agency.

6.4.2.2 At university

Despite feeling positive about her English language development and accessing what she perceived to be a better job opportunity (a customer service role in a supermarket), Ana is still not confident enough to start applying for positions in her field and so decides to undertake a Master’s degree. The interrelated themes of ‘self-appraising’ and ‘practising agency’, highlighted above, as well as the theme ‘responding affectively’ are evident in Ana’s account of her experiences at university, shown in Extract 6.4.2, below.

Extract 6.4.2

“\textit{I was afraid of my English level so I decide to study, I wasn’t ready.... Even when I started my Uni, I wasn’t ready. I don’t know how I did...I never failed any, any subject, because I studied very hard... But I used to cry, after classes I used to cry because I couldn’t understand a word what they say to me. But your brain is wonderful, our brains are wonderful. At the beginning, I couldn’t understand anything, but after 3 months or 4 months, everything becomes easier, you know...I passed, and I learned a lot of things...}” [Interview 1, 293-318]
In Extract 6.4.2, Ana expresses negative self-appraisals related to her capacity to move directly into skilled employment (I wasn’t ready) and also in relation to her initial ability to cope with the English demands of university study (I wasn’t ready; I couldn’t understand a word). These negative self-appraisals are associated with affective responses (I was afraid; I used to cry). As in her account of experiences before university, negative self-appraising (I was afraid of my English level) is again linked to practising tactical agency (so I decide to study); Ana hoped studying at university would provide the time and resources needed to overcome the English language obstacles she perceived to be limiting her access to skilled employment. Ana attributes her successful outcomes at university explicitly to the tactical agency she practised to overcome the English language obstacles that she faced (because I studied very hard). At the end of Extract 6.4.2, she describes a transformation in her experience of using English at university (everything becomes easier), which implicitly indicates a shift towards positive self-appraising, and is directly linked to the achievement of obtaining a degree.

Taken together, Extracts 6.4.1 and 6.4.2, above, demonstrate that Ana’s self-appraising varies in relation to different goals. Ana’s self-appraisal of her English skills is positive in relation to achieving the intermediary goal of accessing a better job (Extract 6.4.1) but negative in relation to achieving her primary goal of transitioning into skilled employment (Extract 6.4.2). Her self-appraising shifts from negative to positive as her English skills develop in the context of her university studies and the further intermediary goal of gaining a master’s degree. The interaction between themes ‘self-appraising’ and ‘practising agency’, illustrated by Extract 6.4.1 and 6.4.2, can be summarised as follows: Ana positions herself as agent in relation to her language learning and employment goals; she acts purposefully, identifying intermediary goals against which to conduct self-appraisals as she negotiates the obstacles she encounters, and attributing successes internally, to her own tenacity and hard work. Throughout her accounts, positive self-appraisals are associated with practising strategic agency (directly goal oriented actions/decisions); Ana associates learning English with an increasing capacity to take deliberate and strategic action towards achieving her intended employment and migration outcomes. By contrast, negative self-appraisals are
associated with tactical agency, as Ana strives to create pathways around obstacles she is unable to immediately overcome.

Extract 6.4.3, below, relates to Ana’s first ever IELTS experience, which was prior to commencing her test taking efforts for permanent residency. Ana took the test in the middle of the last semester of her master’s degree, in September 2011, to satisfy the English requirements for a temporary post-study visa. This temporary visa would allow her work rights for a further 18 months after graduation. Her brief account of her test result on this occasion suggests a relationship between themes ‘self-appraising’, ‘understanding target scores’ and ‘interpreting score outcomes’ that, as will be shown further below under the heading ‘After university’, underlies her perceptions of the score requirements for permanent residency.

**Extract 6.4.3**

“So you have to show to them that you are functional, so you are asked to get six. In every single band... I got six point five in every single band”

[Interview 1, 588-594]

Extract 6.4.3, above, indicates a relationship between Ana’s understanding of target scores, in this case IELTS 6, and her process of self-appraising. She perceives test purpose to be to verify that migrants possess a ‘functional’ level of English for the purpose of employment (you have to show to them that you are functional, so you are asked to get six). In this extract, she places emphasis on the “point five” and the fact that she scored above the minimum requirement. This indicates that she is interpreting her own score outcome to mean that she has proven she is at least ‘functional’, and further suggests that she perceives scores as a meaningful evaluation of her English level. This becomes explicit in Extract 6.4.5, further below, when she uses the term ‘functional’ again, this time explicitly to describe her capacity to communicate in English in the context of workplace interactions.
6.4.2.3 After university

Extracts 6.4.4 and 6.4.5, below, provide additional evidence of the relationship between themes ‘self-appraising’ and ‘understanding target scores’ mentioned above. Ana is referring to her social and work contexts in these extracts, as they were at the time she provided the account. Her perceptions of successful integration into the local community, socially and professionally, are drawn upon to support an implicit claim in Extract 6.4.5, that her English skills are already adequate and the score requirements for skilled migration are too restrictive.

Extract 6.4.4

“The funny thing is my English, my, my friends are Aussie. I consider them my friends. My neighbour, he’s very Aussie. And Sara* is super Aussie. And Jill*, from work, she’s super super Aussie, and Dave* as well... This is a different thing than the listening IELTS test. Even if they use slang with me, I can understand everything.” [Interview 1, 720-735]

*pseudonyms used

In Extract 6.4.4, Ana contrasts her experiences communicating with her “Aussie” friends and her associated positive self-appraisal (Even if they use slang with me, I can understand everything) with her experience of the IELTS listening test. In so doing, she provides an implicit negative evaluation of the test construct, questioning its relevance to ‘real world’ communication.
Excerpt from the document:

"Because this is a, this is an English country, everyone speaks English and the government has to measure that in one or another way... the thing is how they change the score so drastically... from 6 to 7 or 8... I am functional. I work in an Australian company. And if you ask them, I think, they have no problems with me and my communication. I haven’t heard, I haven’t heard any complaint about me. I haven’t heard any complaint about me from my customers, because I’m in contact with them every day" [Interview 1, 888-910]

In Extract 6.4.5, Ana is criticizing the introduction of scores of IELTS 7 and 8 into skilled migrant selection processes. Her comment “they change the score so drastically... from 6 to 7 or 8” illustrates an understanding of scores of IELTS 7 and 8 as arbitrary and unjustified, representing a level of English beyond that required to communicate and engage in work interactions. Her self-appraisal “I am functional” in Extract 6.4.5 echoes her earlier interpretation of score outcomes, in Extract 6.4.3, and serves as evidence in support of her criticism of the test requirements. She further asserts her ‘functional’ English level and her successful integration by citing her employment in an Australian company and her proven capacity to communicate effectively in English with her colleagues and customers on a daily basis.

6.4.3 Summary of themes: Experiencing learning and using English

As demonstrated by Extracts 6.4.1 and 6.4.2, Ana’s self-appraising of her English abilities was contingent on different intermediary goals (finding work in Australia, changing jobs and succeeding at university) and her primary work-related goal of gaining a job in Australia commensurate with her skills and qualifications. The extent to which she feels able to directly pursue her goals influences the type of agency she practises; positive self-appraisals of her English skills are associated with practising strategic agency, whereas negative self-appraisals are associated with practising tactical agency as Ana
perceives the need to find new paths to overcome difficulties or to avoid insurmountable obstacles (see Table 6.1).

Table 6.1. Self-appraising and practising agency: Linking Ana’s self-perceptions and actions

<table>
<thead>
<tr>
<th>Goals</th>
<th>Self-appraising</th>
<th>Practising agency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Finding work</td>
<td>we didn’t speak</td>
<td>I started working as a cleaner</td>
</tr>
<tr>
<td>Changing jobs</td>
<td>my English is functional enough</td>
<td>it’s time for me to move on</td>
</tr>
<tr>
<td>Gaining skilled employment</td>
<td>I was afraid of my English level; I wasn’t ready</td>
<td>so I decide to study</td>
</tr>
<tr>
<td>Succeeding at university</td>
<td>I wasn’t ready; I couldn’t understand a word</td>
<td>I studied very hard</td>
</tr>
</tbody>
</table>

Extracts 6.4.3 to 6.4.5 illustrated a relationship between Ana’s process of self-appraising and her understanding of target scores (see Table 6.2). As summarised in Table 6.2, Ana questions the appropriateness of the score requirements for skilled migrant selection based on the sense of legitimacy she derives from her previous IELTS scores and her experiences using English in social and work interactions, as manifest in her positive self-appraising.

Table 6.2. Self-appraising and understanding target scores: Linking Ana’s self-perceptions and test perceptions

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Self-appraising</th>
<th>Appraisal</th>
<th>Understanding target scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test scores:</td>
<td></td>
<td>Implicit</td>
<td>[I scored &gt; 6; I am functional]</td>
</tr>
<tr>
<td>you have to show to them that you are functional, so you are asked to get six... I got six point five in every single band</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social interactions:</td>
<td></td>
<td>I can understand everything</td>
<td></td>
</tr>
<tr>
<td>Even if they use slang with me</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work interactions:</td>
<td></td>
<td>I am functional</td>
<td>they change the score so drastically...from 6 to 7 or 8</td>
</tr>
<tr>
<td>they have no problems with me and my communication... I haven’t heard any complaint about me</td>
<td></td>
<td></td>
<td>[Target scores &gt; functional; Implicit: scores are too high]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.4.4 Before test attempt 1

Ana’s approach to her first test attempt was shaped by her self-appraisals and understanding of target scores, outlined in the previous section. She situates target scores of IELTS 7 and 8 above the level of English she needs to communicate effectively in everyday social and work interactions, and so perceives achieving such scores as representing a significant challenge. Her effort to adopt a positive attitude in her approach to the test at the outset, shown below, mirrors the mindset that effort and perseverance yield success evident in her account of experiences learning and using English, where she attributes her capacity to overcome obstacles to hard work and tenacity in the face of adversity. This approach to the test also reflects a perception of test purpose as a legitimate evaluation of her English skills, which she maintains throughout the first three of her four test encounters, despite her rejection of the appropriateness of the score requirements for permanent residency. Ana’s account of her initial approach to the test, exemplified below in Extracts 6.4.6 and 6.4.7, is characterised by relationships between themes ‘self-appraising’, ‘practising agency’, and ‘responding affectively’.

Extract 6.4.6

“Everything that happens in your life, you create it. If I think I am not able or I am not capable I won’t get it. I have to brush, wash my brain every morning, you can do it, you will improve, you can do better” [Interview 1, 1023-1031]

In Extract 6.4.6, above, Ana situates self-doubt, rather than the test score requirements, as the main obstacle to realising her migration goal (If I think I am not able or I am not capable I won’t get it). She thus identifies avoiding negative thinking as a key strategy in her approach to the first test encounter. Themes ‘self-appraising’ and ‘practising agency’ are linked in
Extract 6.4.6 as Ana’s positive self-affirmations (you can do it, you will improve, you can do better) represent both positive self-appraisals and the outcomes associated with practising strategic agency to avoid negative thinking (I have to brush, wash my brain every morning).

Extract 6.4.7

“…he [her husband] depends on my results, and we have built a life here, and now my residency depends on that exam…um, but if you want to get good score you have to apply all your skills, it’s not about English skills. It’s about attitude, you have to be very positive, you have to practice, but the, but you have to think, you must start thinking that you are able to pass the exam. You have to encourage yourself, you have to be your own cheer leader. But I’m afraid of. I’m shaking. I’m very worried. But life goes on, and you have to, you have to do your best.” [Interview 1, 1295-1323]

In Extract 6.4.7, above, Ana highlights the high stakes she associates with test outcomes (we have built a life here, and now my residency depends on that exam) and the challenge she faces in exercising control over negative affective responses prompted by the uncertainty of her situation (I’m afraid. I’m shaking. I’m very worried). Consistent with her account of experiences learning and using English, she links overcoming obstacles and achieving positive outcomes (but if you want to get a good score...) with practising strategic agency (you have to be positive...you have to encourage yourself) and maintaining positive self-appraisals (you must start thinking that you are able to pass the exam).

Taken together, Extracts 6.4.6 and 6.4.7 demonstrate Ana’s sense of agency and responsibility in determining her own fate as she plans for her first test attempt for permanent residency. Ana positions herself as agent, assuming responsibility for overcoming obstacles, and associating positive outcomes with her capacity to exercise control over her thoughts and feelings. This was also the case in her account of her prior experiences learning and using English. She locates the main obstacle to success internally; specifically, her own fears and self-doubt, and associates the capacity to
exercise control over these negative feelings with test success. As is shown in the subsection below, her first score outcomes undermine these perceptions, prompting a loss of agency as Ana shifts from strategic to tactical action in response to unexpected score outcomes.

6.4.5 Experiencing repeat test encounters

Prior to her first test attempt for permanent residency, Ana’s experiences learning and using English had reinforced her perceptions of a relationship between effort and success. Thus, Ana believed that her test scores were likely to improve through test preparation and using English daily. However, on her first attempt, Ana’s scores on all skills except listening were no better than they had been when she took the test more than a year earlier to gain a temporary post-study work visa. On her second test attempt, a month and a half after the first, her score for speaking increased compared to the first attempt, but her scores dropped on the remaining three skills, including listening. Five months passed before Ana made her third attempt, on which she achieved her target scores of IELTS 7 or above on all skills except writing. On the fourth and final test encounter, another four months later, Ana achieved these scores on all four skills and so was able to apply for a permanent visa. A comparison of her score outcomes across the four encounters is provided in Table 6.3, below.

Table 6.3. Ana’s IELTS scores

<table>
<thead>
<tr>
<th></th>
<th>Oct 2012</th>
<th>Dec 2012</th>
<th>May 2013</th>
<th>Sep 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>6.5</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Writing</td>
<td>6.5</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Listening</td>
<td>8</td>
<td>6</td>
<td>7.5</td>
<td>8</td>
</tr>
<tr>
<td>Reading</td>
<td>6.5</td>
<td>6</td>
<td>7.5</td>
<td>8</td>
</tr>
</tbody>
</table>

Ana remained confident from the outset that she was able to achieve target scores on the objectively scored reading and listening parts of the test, despite fluctuating test outcomes, and rarely mentioned these skills in her accounts of her repeat test experiences. By contrast, her score outcomes for speaking and especially writing
confounded her and she sought expert help to identify how to improve these skills. To this end, she hired three consecutive private tutors, a different one after each unsuccessful test attempt as ‘failures’ prompted her to question prior feedback and seek out new help.

In this subsection, Ana’s responses to score outcomes from each test encounter will be detailed, in turn. As will be shown, Ana’s experiences across repeat test encounters are characterised by her use of feedback, which was different and sometimes conflicting as she moved from tutor to tutor. Ana evaluated feedback and integrated her evaluations into processes of self-appraising, interpreting score outcomes, and understanding target scores. These processes were associated with shifts in the types of agency she practised and in her affective responses, which shaped subsequent test outcomes and influenced subsequent perceptions, feelings, and actions, as described below.

6.4.5.1 After test attempt 1

As mentioned, Ana’s score outcomes from her first attempt were inconsistent with her expectations. She interpreted score outcomes to mean there was a difference between how she was performing on the test and the quality expected by examiners, but was unable to identify specific aspects of her performances that needed modification. In response to this uncertainty, she sought expert feedback from a private tutor who was also an IELTS examiner.

Extracts 6.4.8 to 6.4.11, below, illustrate Ana’s response to the feedback she receives from the private tutor. Extract 6.4.8, which shows Ana’s evaluation of the feedback she receives, highlights interrelationships between themes ‘understanding target scores’, ‘interpreting score outcomes’, ‘self-appraising’ and ‘practising agency’ (see also Table 6.4, further below).

Extract 6.4.8

“*She told me that my writing wasn’t good enough to get a 7, wasn’t, was not, not good enough to get a 7. Um, also my speaking because my*
pronunciation was awful *** I think they follow a pattern, you know what I mean, a pattern? So if she thinks that, I think the rest will think the same, so I, what else can I do, you know? I have to continue [working with her]”

[Interview 2, 58-72]

As Extract 6.4.8 shows, the feedback Ana received from the tutor indicated the existence of a substantive difference between the level of her speaking and writing skills and the level associated with target scores. The way in which Ana integrates this feedback into a process of interpreting score outcomes, self-appraising and practising agency, is set out in Table 6.4, below.

**Table 6.4. Ana’s use of feedback from tutor 1**

<table>
<thead>
<tr>
<th>Receiving feedback</th>
<th>Evaluating feedback</th>
<th>Understanding target scores</th>
<th>Interpreting scores</th>
<th>Self-appraising</th>
<th>Practising agency</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>my writing wasn’t good enough ... my pronunciation was awful</em></td>
<td><em>they [IELTS examiners] follow a pattern...the rest will think the same</em></td>
<td>Implicit: Consistent measure, meaningful indicators of English skills</td>
<td>Implicit: scores reflect English skill deficiency</td>
<td>Implicit: speaking and writing skills below target score level</td>
<td><em>what else can I do...I have to continue</em></td>
</tr>
</tbody>
</table>

As shown, Ana resigns herself to accepting the tutor’s negative judgment of her writing and speaking ability, despite its inconsistency with her prior experiences using English, because she believes that other IELTS examiners “will think the same”, suggesting she perceives the test to be a consistent instrument. Her action response represents a shift from practising strategic agency to tactical agency; given that her own strategic approach to the test failed to yield success, Ana defers decision-making to the tutor. She assumes the tutor, as an IELTS examiner, has access to insider knowledge of scoring criteria and so she feels she has no choice but to continue working with her (what else can I do...I have to continue).

The three short Extracts, 6.4.9 to 6.4.11, below, indicate Ana’s feelings in relation to the feedback she receives, illustrating the theme ‘responding affectively’. As shown, the negative shift in Ana’s self-appraisal prompts negative feelings about English and a loss
of motivation (Extract 6.4.9), a loss of confidence in her English abilities (Extract 6.4.10) and a loss of self-confidence (Extract 6.4.11).

Extract 6.4.9

“I’ve started hating **English**...and those stupid topics about **vitamins and schools**...” [Interview 2, 148-152]

Extract 6.4.10

“I have started feeling **silly** most of the time. Especially with my speaking. Every time I have to think twice” [Interview 2, 183-187]

Extract 6.4.11

“I have this feeling that everything is wrong with me” [Interview 2, 855]

6.4.5.2 After test attempt 2

As noted above, Ana’s preparation test attempt 2 was guided by her first private tutor. On this second attempt, Ana’s speaking score improved slightly but her writing score went down compared to her first test encounter. This prompted her to re-evaluate the feedback upon which she had based her preparation efforts, as shown in Extract 6.4.12, below.

Extract 6.4.12

“There is no doubt that the other woman was stealing my money. She didn’t teach me nothing new.” [Interview 3, 367-369]

Ana reverses her previous evaluation of the tutor’s feedback, dismissing it as meaningless in light of her second test score outcomes. Again feeling at a loss concerning how to interpret her failure to achieve the score she needed in writing, Ana seeks out a second tutor (Extract 6.4.13), and this time finds an English teacher with no experience as an IELTS examiner. Extract 6.4.13 illustrates a relationship between themes ‘interpreting score outcomes’ and ‘practising agency’; the action of seeking out
new expertise represents practising tactical agency, as unexpected scores outcomes prompt Ana to try a different approach to locating the key to test success.

Extract 6.4.13

“...I was desperate because I didn’t know what else I could do in order to improve my writing...so I prayed to find someone like him...and I googled”
[Interview 3, 462-469]

Ana’s shift in tactics, represented by abandoning the first tutor and seeking out new help, led to a positive shift in self-appraisals as the positive feedback the second tutor provides helps Ana regain confidence in her English abilities. An interaction between themes ‘self-appraising’ and ‘practising agency’ is illustrated by interview excerpts presented in Table 6.5, below. The second tutor’s positive feedback aligned well with Ana’s prior experiences learning and using English. As shown in Table 6.5, this helped Ana to situate test preparation as part of further expanding her English language skills and enabled her to re-establish a sense of control over her own fate. A return to positive self-appraising and a renewed sense of agency allowed Ana to re-establish her initial strategic approach to the test, in which she identified self-doubt as the main obstacle she faced, and she re-adopts maintaining positive thinking as a key strategy.

Table 6.5. Ana’s use of feedback from tutor 2

<table>
<thead>
<tr>
<th>Receiving feedback</th>
<th>Evaluating feedback</th>
<th>Self-appraising</th>
<th>Practising Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>Strategic</td>
<td></td>
</tr>
<tr>
<td>he truly believes</td>
<td>he’s a wonderful</td>
<td>My battle is not</td>
<td>I’ve been studying</td>
</tr>
<tr>
<td>I am ready</td>
<td>teacher/</td>
<td>with English</td>
<td>writing... I try to think</td>
</tr>
<tr>
<td></td>
<td>I have learned things</td>
<td>anymore I think</td>
<td>this is not an imposition,</td>
</tr>
<tr>
<td></td>
<td>that I didn’t know</td>
<td>my battle is my,</td>
<td>this is something you can use the rest of your</td>
</tr>
<tr>
<td></td>
<td></td>
<td>my confidence</td>
<td>life</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and my nerves</td>
<td></td>
</tr>
</tbody>
</table>

Interview 3, 294  Interview 3, 132/253  Interview 3, 145-8  Interview 3, 357-360

Practising strategic agency after the second test attempt thus involved conceptualising test preparation as a meaningful activity (this is not an imposition, this is something you can use the rest of your life). Ana’s effort to embrace a positive attitude about the test
is further reflected in Extract 6.4.14, below, in which she explicitly reasserts the link that she established in her account of her experiences learning and using English, between developing her English skills and accessing better job opportunities.

Extract 6.4.14

“I want to make it right. It’s not just for the test. I want to get a better job. If I want to get a better job, I need to improve my er language skills”

[Interview 3, 217-223]

Thus, in allowing Ana to regain a sense of control and purpose, the feedback she receives from the second tutor is associated with a shift away from the negative affective responses and loss of motivation prompted by the first tutor’s feedback (I’ve started hating English, Extract 6.4.9), towards a renewed sense of intrinsic motivation concerning her English learning.

6.4.5.3 After test attempt 3

On her third test attempt, Ana achieves her target of IELTS 7 or above in all skills except writing, despite focusing significant attention on this skill with her second tutor. In response, Ana moves on to a third tutor, this time another IELTS examiner.

Extracts 6.4.15 to 6.4.17, below, illustrate a relationship between themes ‘interpreting score outcomes’ ‘self-appraising’ and ‘understanding target scores’, as feedback provided by the third tutor triggers a significant shift in Ana’s perceptions of score meanings. Before meeting this third tutor, Ana had been interpreting her failures to reach IELTS 7 in writing to mean that her writing skills were deficient and in need of further development. However, as shown in Extract 6.4.15, below, the third tutor signals that her writing skills are adequate and alerts Ana to the existence of an arbitrary formula against which performances are judged. This prompts a shift in Ana’s interpretation of score outcomes; rather than
indicating a problem with her English level, she begins to associate poor scores outcomes with a failure to deliver the expected formulaic performance.

**Extract 6.4.15**

“The problem is I was writing too much, too many paragraphs. OK? You should write four paragraphs instead of five and you should have just one point or two per paragraph. No more than two. And I was writing like three points per paragraph, it’s too much...

***

The intro and conclusion shouldn’t have more than two sentences...you should write about 90 and 100 words per paragraph. So it should be 25 or 30 words, 60 words conclusion and introduction, and hund- and 200 the two paragraphs...The sentences I was writing were too short. They should have at least 17 words. From 17 to 25 words...

***

... he told me something very interesting: ‘Ana, you’re a good writer but just tell them what they want to see just for the test and you’ll be alright’

[Interview 4, 6-89]

Besides this advice from the third tutor, Ana also receives information from two other test takers that prompts her to suspect examiner subjectivity is a determining factor in score outcomes. As shown in Extract 6.4.16, below, the first test taker tells her that a certain test centre is associated with higher scores. The second tells her that his scores went up from IELTS 6 to IELTS 7 in one week.

**Extract 6.4.16**

“Every- Apparently, everyone gets good scores when they take the test in [name of test centre]...I was studying with Colombian boy and he needed just 6 and he took the test 7 times...The last time he took it he took it at [name of test centre] so he got the score he needed. And, the other guy, he took the test, for example, this Saturday ... and he got 6 and he took the test next er week, weekend and he got 7 in writing...
...how can, how could you improve your scores, score in one week? It’s not possible. Or may-Er I don’t kn- That’s why I truly believe that the writing is subjective. It depend of the examiners mood, it depends on his- Or- What if, if they have 300 papers and maybe they are being told not to give too many 7s” [Interview 4, 279-304]

As illustrated by Extract 6.4.17 below, Ana’s suspicions are confirmed by the third private tutor:

**Extract 6.4.17**

“...he told me ‘sometimes it’s luck Ana...because they work under a lot of pressure’...”

***

...That’s why he told me about, ‘it’s just ticking, ticking the boxes Ana’”

[Interview 4, 316-334]

Extract 6.4.18, below, illustrates the theme ‘practising agency’. Ana’s new perceptions of the test lead her to shift the focus of her test preparation away from developing her English skills towards practising adherence to the formula the tutor prescribes. Since Ana perceives this approach to be linked to achieving target scores, her action represents strategic agency.

**Extract 6.4.18**

"I am writing what they want to read and I am trying to include the things that he told me to do, to include" [Interview 4, 336-337]

Thus, as indicated by extracts 6.4.15-6.4.18, Ana integrates tutor 3’s feedback, supplemented by information from other test takers, into revised interpretations of score outcomes, self-appraisals, and understandings of target scores. These shifting perceptions prompt Ana to adopt a new strategic approach to the test. The relationships between themes, highlighted above, are presented in Table 6.6, below.
Table 6.6. Ana’s use of feedback from tutor 3

<table>
<thead>
<tr>
<th>Interpreting score outcomes</th>
<th>Self-appraising</th>
<th>Understanding target scores</th>
<th>Practising agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was writing ... too</td>
<td>you’re a good</td>
<td>the writing is</td>
<td>I am writing what</td>
</tr>
<tr>
<td>many paragraphs</td>
<td>writer</td>
<td>subjective/</td>
<td>they want to read</td>
</tr>
<tr>
<td>[too many] points per</td>
<td></td>
<td>it’s just ticking,</td>
<td></td>
</tr>
<tr>
<td>paragraph... sentences</td>
<td></td>
<td>ticking the boxes</td>
<td></td>
</tr>
<tr>
<td>were too short</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extract 6.4.15</td>
<td>Extract 6.4.15</td>
<td>Extract 6.4.16/17</td>
<td>Extract 6.4.18</td>
</tr>
</tbody>
</table>

Extracts 6.4.19 and 6.4.20, below, illustrate the theme ‘responding affectively’. Ana’s shift in strategic approach, as shown above, is associated with a cynical perception of the test (it’s just ticking the boxes), which undermines her sense of test preparation as a meaningful activity. She no longer associates improving test scores with further refining her English skills, as she had after the second test attempt, and this shift in perception prompts her to feel demotivated (see Extract 6.4.19).

Extract 6.4.19

“I don’t want to write...No, ah, ah, I... I have to start...I know that I have to start working on it but every time that I take a pen and a piece of paper I feel ‘here we go again’ and...” [Interview 4, 210-213]

This cynicism is also associated with feelings of frustration, as the ‘benefits’ of test preparation become limited solely to score outcomes (Extract 6.4.20):

Extract 6.4.20

“This one is the last time I take the test. After that I don’t know what’s going to happen because I’m wasting, I’m wasting my money and my time and my patience”

[Interview 4, 340-344]
6.4.5.4 After test attempt 4

Ana’s overall score success on her fourth attempt, after modifying her approach to test preparation according to advice received from her third tutor, reinforces the understanding of test scores she derived from the third tutor’s feedback. As illustrated by Extract 6.4.21, below, Ana feels certain that target scores represent the particular performance formula this tutor provided.

Extract 6.4.21

“One of the most important things about getting 7 in writing is free-error sentences.

***

In IELTS the sentences must be in-between 17 and 24 because it’s since there’s perfect length for a complex sentence” [Interview 5, 105-122]

Interestingly, however, in Extract 6.4.22, Ana attributes her test success to the effectiveness of her original strategic approach, rather than the revised approach she took in the lead up to her final test encounter, as recommended by the third tutor.

Extract 6.4.22

“At the end of the process I understood, finally, that I had enough English to pass the test, or to get the marks I needed in order to apply for residency...You have to learn how to control your emotions, you have to be calm during the test, it’s very important because if you are nervous or if you allow yourself to put a lot of pressure that day, it’s not going to work”

[Interview 5, 82-87]

As shown in Extract 6.4.22, Ana takes ownership of the positive test outcome, expressing a positive self-appraisal (I had enough English to pass the test) and associating the realisation of her score goals with her capacity to control negative affective responses (You have to learn how to control your emotions, you have to be calm during the test...).
Excerpts 6.4.23 and 6.4.24, below, illustrate Ana’s affective responses to her successful score outcomes. As shown, she expresses positive feelings about English (Extract 6.4.23) and a positive shift in motivation (Extract 6.4.24) as she re-asserts the intrinsic value of putting effort into developing her English skills.

**Extract 6.4.23**

“I love Australia, I love English. Even it’s, even if it’s still broken, I love English, I love English from the bottom of my heart” [Interview 5, 225-227]

**Extract 6.4.24**

“I’m going to start er writing...this time I’m going to enjoy it because I’m not going to be tested it’s more like OK, English is my second language and I’m expecting to spend all my life in this country and I want to speak proper English. You have to honour your language, no?” [Interview 5, 281-293]

In the end, as shown above, Ana’s final test success has two main effects on her perceptions: Firstly, she believes that she located the specific formula for success, thanks to advice from the third tutor, and so is finally able to overcome uncertainty and fix meaning to target scores. Secondly, she recovers a sense of certainty in relation to her English skills and her legitimacy as an English user. After a year of upheaval and fluctuations, Ana sums up the impact of her experiences as follows:

**Extract 6.4.25**

“I was in prison for one year. I used to wake up with that pressure, like, when I opened my eyes, like ‘oh my God, the test’, every day. Yes and because my husband as well, he has got a good job and he’s really happy in that company and everyday he used to come home at night and he was like, ‘oh I had a great day at job’ and I was thinking if I don’t pass this fucking test we’re going to be sacked of this country. I felt like a loser, you know” [Interview 5, 569-581]
6.4.6 **Summary of Ana’s experiences**

Shifting self- and test perceptions influence feelings, decisions, and actions throughout Ana’s test taking journey, as she attempts and reattempts to establish certainty concerning score meanings from one test encounter to another. In this way, the impact of the test on Ana’s life emerges not only as a result of the existence of the score requirements within skilled migration policy, but as a result of interactions between Ana’s subjectivity and the test over time. Interrelationships between Ana’s thoughts, feelings, decisions and actions, set out in relation to Ana’s experiences in terms of the five key themes: self-appraising, understanding target scores, interpreting score outcomes, practising agency and responding affectively, are illustrated in Figure 1.1, below.

**Figure 6.1: Summary of thematic relationships (Ana)**
As demonstrated in relation to Ana’s experiences, and highlighted in Figure 6.1, above, she draws on test and real world experiences using English in order to conduct self-appraisals, interpret scores and understand target scores (indicated by blended shading in Figure 6.1). Her repeat test encounters are characterised by a perception that although the meaning of target scores eludes her, it remains fixed and stable over time and reflects actual English ability, albeit at a level above and beyond what she sees as necessary to communicate effectively in her social and work contexts in Australia. As Ana experiences them, target score meanings are positioned as constant (indicated by the solid border in Figure 6.1) and she is in a state of flux; her language behaviour is in need of modification and correction. She relentlessly pursues different experts’ feedback from test encounter to test encounter, as a means of discovering the qualities of a successful performance, of attributing meaning to her own score outcomes, and of devising appropriate action responses to bring about success.

After her first test attempt, the feedback Ana receives from the first expert undermines her sense of legitimacy as an English speaker; her negative self-appraising extends beyond the test as she adopts the deficit perspective assumed by this expert. From then on, different feedback is associated with an iterative cycle of shifting interpretations of score outcomes and shifting self-appraisals, which prompt Ana to practise different types of agency (indicated by the broken borders in Figure 6.1). Nonetheless, even after the third tutor’s advice, when she begins to suspect examiner subjectivity as a source of interference and variability in scoring, she remains committed to the belief that achieving target scores depends on her own capacity to avoid mistakes and to produce a performance consistent with implicit and relatively stable criteria prioritised by examiners.

To conclude this section on Ana, a summary of the main shifts in her thoughts, feelings, decisions and actions over time, categorised according to the five themes mentioned above, is presented in Table 6.7, below.
<table>
<thead>
<tr>
<th>Themes</th>
<th>Before test 1</th>
<th>After test 1</th>
<th>After test 2</th>
<th>After test 3</th>
<th>After test 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-appraising</strong></td>
<td>you can do it, you will improve... [E6.4.1]</td>
<td>my writing wasn’t good enough... my pronunciation was awful [E6.4.8]</td>
<td>I am ready / My battle is not with English anymore [T6.5]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of appraisal</td>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Understanding target scores</strong></td>
<td>Implicit: scores provide consistent measure and meaningful information about English skills</td>
<td></td>
<td></td>
<td></td>
<td>the writing is subjective [E6.4.16]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>it’s just ticking, ticking the boxes [E6.4.17]</td>
</tr>
<tr>
<td><strong>Interpreting score outcomes</strong></td>
<td>Implicit: scores reflect English skill deficiency</td>
<td></td>
<td></td>
<td>Implicit: scores represent performance issue</td>
<td>I was writing too many paragraphs ... sentences were too short [E6.4.15]</td>
</tr>
<tr>
<td><strong>Practising agency</strong></td>
<td>you have to be positive [E6.4.7]</td>
<td>what else can I do... I have to continue [E6.4.8]</td>
<td>I’ve been studying writing... I try to think this is not an imposition, this is something you can use the rest of your life [T6.5]</td>
<td>I am writing what they want to read...[E6.4.18]</td>
<td></td>
</tr>
<tr>
<td>Type of agency</td>
<td>Strategic</td>
<td>Tactical</td>
<td>Strategic</td>
<td>Strategic</td>
<td>Strategic (revised)</td>
</tr>
<tr>
<td><strong>Responding affectively</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling unmotivated</td>
<td></td>
<td>Feeling motivated</td>
<td>Feeling unmotivated</td>
<td></td>
<td>Feeling motivated</td>
</tr>
<tr>
<td>Feeling motivated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E=Extract; T=Table
6.5 Erfan

6.5.1 Introduction

Erfan arrived in Australia in 2009 to undertake a master’s degree in computer science. He was aged in his mid-20s at the time. He had completed undergraduate studies in Iran, his home country, and had gained three years of experience working in his profession prior to coming to Australia. He came to Australia with the intention of pursuing a PhD after completing his master’s degree. He described the option of gaining permanent residency as a skilled migrant, which at the time was linked to studying in Australia, as the “icing on the cake” that had led him to choose Australia as a destination.

Erfan identifies Persian as his first language. He also speaks Russian, because his family lived in Russia for four years when he was a child, and has some basic knowledge of Arabic. He learned English during his school years in Iran but had very little exposure to the language outside the classroom. Before leaving Iran, he took IELTS to try to satisfy the English language requirements for direct entry into university in Australia. He scored IELTS 6 overall, below the minimum required score of 6.5, so he completed a ten-week English language course in Australia prior to commencing his master’s studies, which served as an alternative entry pathway into university. He successfully completed his master’s degree, achieving a mix of distinction and high distinction grades, but was unable to secure an international student scholarship for doctoral studies.

Without a scholarship, Erfan would have had to pay international student tuition fees in addition to covering his living costs while completing a PhD. During his master’s studies, his parents in Iran had provided funds for his university fees and had been providing money to cover his living expenses in Australia, which Erfan estimated to amount to around 100,000 Australian dollars. However, Erfan was becoming concerned about what he perceived to be worsening economic and political conditions in Iran. He felt that his parents were experiencing increasing hardship, and he was unable to ask them for further financial support. At the same time, he was unwilling to return to live in Iran and
felt enormous pressure to realise his study and migration goal, given his parent’s already significant contribution and the time and effort he had thus far committed.

Due to the financial difficulties associated with pursuing further studies as an international student, Erfan decided he needed to first apply for permanent residency in Australia. Permanent resident status would enable him to undertake a doctorate as a local student, exempt from tuition fees. He again took IELTS, this time for the purpose of securing a temporary post-study visa, which would allow him 18 months to work and to accumulate the points needed to qualify for a permanent visa. He achieved an average score of IELTS 6.5 on this occasion, above the IELTS 6 required.

Erfan’s priority then became to secure employment as soon as possible. After spending three months looking for full-time work as an IT professional without success, Erfan was offered a position as a research assistant at the university where he had completed his master’s degree. The job was part-time and so not ideal, but it was related to his area of study and with his financial situation becoming desperate, he decided to accept. He had been working there for just over three months when he took part in his first interview for this study, in September 2012, a month before his first IELTS attempt for the purpose of accessing permanent residency.

Erfan needed IELTS 7 to accumulate the minimum points total for eligibility to submit an expression of interest in applying for a permanent skilled visa, but was aiming for IELTS 8 to push his points total above the minimum. He took the test four times in a six-month period but was unable to achieve IELTS 8 in all four skills. On the fourth and final attempt, he achieved a mix of 7s and 8s and decided to try submitting an expression of interest on the basis of these scores. After waiting almost two years for his application to be processed, he was finally granted permanent residency in 2015. Erfan’s experiences are presented below under three main headings: ‘Experiencing learning and using English’, ‘Before test attempt 1’ and ‘Experiencing repeat test encounters’. To conclude the section on Erfan, a summary of his experiences is provided.
6.5.2 Experiencing learning and using English

The theme ‘self-appraising’ is central to Erfan’s accounts of his experiences learning and using English since arriving in Australia, exemplified in Extracts 6.5.1 to 6.5.4, below. As demonstrated in this subsection, Erfan makes a key distinction in relation to language identity, which underlies his process of self-appraising. He distinguishes between English ‘native’ speakers and other ‘foreigners’, and on this basis, makes a further distinction between the English skills needed to engage in social interactions with English ‘native’ speakers, which he feels he lacks, and the English skills needed to function in university and work contexts, which he feels he possesses. These self-appraisals derive from how Erfan feels in interactions (Extracts 6.5.1, 6.5.2, and 6.5.3), thus evidencing a relationship between his process of self-appraising and the theme ‘responding affectively’. His self-appraisals, in turn, influence his perceptions of the test and score requirements, highlighting a further relationship with the theme ‘understanding target scores’ (Extract 6.5.4).

Extract 6.5.1

“I believe, actually I feel, I felt two hurdle for meeting friends, meeting other people and making friends. The very first one for me was financial situation *** the other part was obviously the language barrier *** I still don’t consider it to be sufficient to meet whenever I want and get whatever job I want. I still feel pressure and not comfortable in conversation *** And actually because since I didn’t have any friends I didn’t have any opportunity to practice speaking”

[Interview 1, 91-112].

Extract 6.5.1, above, demonstrates that Erfan perceives his English speaking level to be inadequate for initiating social interactions and engaging in everyday conversations. He is referring exclusively to interacting with ‘native’ speakers in making this negative self-appraisal, as becomes explicit across Extracts 6.5.2 and 6.5.3, further below. He believes that his inability to use English for social purposes creates what he calls a “language
barrier”, which limit his access to social networks and also to more suitable job opportunities. As shown in Extract 6.5.1, Erfan’s negative self-appraisals of his English skill level (I still don’t consider it to be sufficient...) are derived from negative affective responses associated with interacting in English (I still feel pressure and not comfortable in conversation), thus highlighting a relationship between themes ‘self-appraising’ and ‘responding affectively’. His account also suggests a perpetual cycle of exclusion; being without friends means he lacks access to the opportunities for learning that he feels are needed to develop his English speaking skills. Conversely, possessing inadequate English speaking skills means he is unable to take action to make friends and create learning opportunities.

Erfan’s self-appraisal of his English skills in the study domain, by contrast, is positive, as shown below in Extract 6.5.2.

**Extract 6.5.2**

“I felt very comfortable within the classes and within the university as a whole and I didn’t fi-, I didn’t have any single problem related to language so I thought that it might be a sign of improvement in my language skills... yeah... but definitely the set of language skills that you use within the university is totally different about what you would use in social interactions and making social connections and when I graduated from university when I found it tough to make friends I realised that oops maybe I shouldn’t be that much confident” [Interview 5, 193-203]

Extract 6.5.2, above, again highlights a relationship between themes ‘self-appraising’ and ‘responding affectively’. This time, in the context of his university studies, Erfan associates positive feelings about communicating in English (I felt very comfortable) with positive self-appraisals of his English level (I didn’t have any single problem related to language). However, he explicitly contrasts the language skills needed to function in the university context with those required to establish social ties (“but definitely the set of language skills that you use within the university is totally different about what you would use in social interactions”). This perception of two distinct sets of skills emerges
in light of his continued failure to make Australian friends despite his capacity to meet the demands of his master’s studies (“when I graduated from university when I found it tough to make friends”), triggering a loss of confidence in his English skills.

Extract 6.5.3, below, highlights the distinction that Erfan perceives between interacting in English with other ‘foreigners’, which he feels able to do within the university setting and in the context of his employment, and interacting with ‘native’ speakers of English, which he desires but feels unable to accomplish. This extract further highlights the context-dependent nature of Erfan’s self-appraising, which is contingent on the language identity he assigns himself and his interlocutors.

**Extract 6.5.3**

“Every foreigner person has kind of same level of lexical resources or common language that they use, it’s very different from when you speak to a native person *** Communicating with other students was, even though they all were foreigners and we didn’t have common language apart from English, communicating with them was very easy yeah but when we step out of that small group everybody struggled *** just didn’t feel confident” [Interview 5, 290 – 319]

As Extract 6.5.3, above, shows, Erfan’s concern about his inability to establish social relationships relates specifically to engaging with ‘native’ speakers of English (“it’s very different from when you speak to a native person”). In Extract 6.5.1, shown earlier, when Erfan says “I still feel pressure and not comfortable in conversation”, he is thus referring to communicating with people he identifies as English ‘native’ speakers; those who constitute the group he wishes to access as opposed to other outsiders, like himself, with whom he experiences no difficulty engaging with in English.

Despite negative self-appraisals of his ability to use English to build relationships with people who are English ‘native’ speakers, according to his perceptions, Erfan situates himself as a functional user of English for the purpose of skilled migration. Extract 6.5.4, below, illustrates a link between the themes ‘self-appraising’ and ‘understanding target
scores’, which is further evidenced by Erfan’s account of his perceptions of IELTS and the score requirements, detailed in the next subsections.

**Extract 6.5.4**

“So after working, after living for 4 years and university ... someone definitely has the ability to function in this country, but it might not be as high as 8 but it should be sufficient, at least, I think” [Interview 1, 772-775]

In Extract 6.5.4, above, Erfan identifies his employment and university experiences as the relevant indicators of his capacity to use English to live and work in Australia. He questions the need to use English test scores as additional criteria, and perceives scores of IELTS 8 to represent a level of English above that required to adequately function.

**6.5.3 Summary of themes: Experiencing learning and using English**

As set out above, the theme ‘self-appraising’ is central across Erfan’s accounts of his experiences learning and using English. A summary of the interaction between this theme and themes ‘responding affectively’ and ‘understanding target scores’ is provided in Table 6.8, below.

**Table 6.8. Summary of themes: Erfan's experiences of learning and using English**

<table>
<thead>
<tr>
<th></th>
<th>With 'Native' speakers</th>
<th>With 'Foreigners'</th>
<th>For skilled employment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responding affectively</td>
<td>Self-appraisal</td>
<td>Responding affectively</td>
<td>Self-appraisal</td>
</tr>
<tr>
<td>(Negative)</td>
<td>(Negative)</td>
<td>(Positive)</td>
<td>(Positive)</td>
</tr>
<tr>
<td><em>I still feel pressure, not comfortable in conversation</em></td>
<td><em>I still don’t consider it to be sufficient</em></td>
<td><em>I felt very comfortable</em></td>
<td><em>I didn’t have any single problem related to language</em></td>
</tr>
<tr>
<td>[Extract 6.5.1]</td>
<td>[Extract 6.5.1]</td>
<td>[Extract 6.5.2]</td>
<td>[Extract 6.5.2]</td>
</tr>
<tr>
<td></td>
<td><em>I found it tough to make friends</em></td>
<td><em>...communicating with them was very easy</em></td>
<td><em>someone definitely has the ability to function in this country</em></td>
</tr>
<tr>
<td></td>
<td>[Extract 6.5.2]</td>
<td>[Extract 6.5.3]</td>
<td>[Extract 6.5.4]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>it might not be as high as 8 but it should be sufficient</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Implicit:</strong> IELTS 8 is too high</td>
</tr>
</tbody>
</table>

148
As shown in Table 6.8, above, Erfan conducts self-appraisals of his English skills by evaluating how he feels about using English, which depends on his perceptions of his own language identity compared to the identities of other speakers (‘native’ English speakers or ‘foreigners’). His self-appraisals change according to the different feelings he associates with each type of interlocutor, with negative affective responses leading to negative self-appraisals, and vice-versa. These self-appraisals shape Erfan’s perceptions of target scores, which he views as too high in light of his own experiences using English in the contexts he associates with skilled migration, university, and work.

In the next subsection, ‘Before test attempt 1’, Erfan’s initial strategic approach to the test is outlined. As demonstrated below, his perceptions of his previous IELTS experiences, prior to commencing university and for a temporary post-study work visa, inform his perceptions of test quality and of the role of the test within skilled migration policy. These perceptions, in turn, shape the type of agency Erfan practises in an effort to realise his score goals.

6.5.4 Before test attempt 1

Erfan’s perceptions of IELTS and his associated approach to his first test encounter for permanent residency are exemplified in Extracts 6.5.5 to 6.5.8, below. As highlighted in the previous subsection, Erfan makes a distinction between his perceived inability to use English in social settings with native speakers and his capacity to use English effectively in the context of his university studies and employment. In this subsection, Erfan’s self-appraising is specifically in relation to the test. Erfan’s shifting self-appraisals of test performances and affective responses across previous test encounters lead to a sense that the test lacks consistency (Extract 6.5.5). This perception combines with Erfan’s understanding of target scores as exclusionary to influence the type of agency he practises in his approach to the score requirements (Extracts 6.5.6 and 6.5.8). Erfan’s decisions and actions are also shaped by his feelings in response to his understanding of target scores and the associated uncertainty he experiences (extract 6.5.7).
Extract 6.5.5

“Um...actually it’s not a very bad test but I think that the thing about IELTS is it varies from one exam to another exam, it varies dramatically. I mean, in one exam you feel very, very confident and very on top of the material and you feel very relaxed and the next exam you don’t feel nothing like that ... ...
To me it’s not a very standard test”

[Interview 1, 289-300]

Extract 6.5.5, above, highlights interactions between Erfan’s perceptions of IELTS and themes ‘self-appraising’ and ‘responding affectively’. Referring to two earlier IELTS encounters, Erfan identifies inconsistencies in the test (it’s not a very standard test), to which he attributes a shift from positive self-appraisals (you feel...on top of the material) and positive feelings (you feel very confident) to negative self-appraisals and feelings (the next exam you don’t feel nothing like that).

Extract 6.5.6 below, illustrates a relationship between themes ‘understanding target scores’ and ‘practising agency’. Erfan identifies exclusion as the purpose behind the arbitrary use of scores of IELTS 8 within selection criteria for permanent skilled migration. As discussed in relation to Extract 6.5.4 in the previous subsection, he associates such scores with a level of proficiency above and beyond that required to use English effectively. Here, in Extract 6.5.6, he suggests that scores function as a “filter”, masking a hidden, discriminatory agenda. He responds to this perception by attempting to exercise control over his own thinking as a means of eliminating what he identifies as a potential obstacle to goal achievement, a negative attitude towards the score requirements. Resisting negative thinking thus represents a form of practising strategic agency; in order to remain focused on pursuing his goal, he makes a conscious decision to avoid becoming demotivated by the perceived unfairness of the score criteria.
Extract 6.5.6

I just see it as a very fine filter. I don’t want to think about it in this way but because of, it gets in my mental way, yeah, but yeah, to me it works, it’s just like a filter [R: what’s being filtered?] The number of people, the number of applicants. Apparently, they think they can’t say that we don’t like people I don’t know, older than 40 or from these countries or whatever” [Interview 1, 503-509]

Extract 6.5.7, below, illustrates a relationship between themes ‘practising agency’ and ‘responding affectively’.

Extract 6.5.7

“I don’t want to think about it, it is a thing that I have to do and I have no way to escape so I don’t want to mess my mind around with that, I don’t want to blame anyone. It’s not very good, it’s very difficult and you know the whole, my life depends on it” [Interview 1, 681-690]

Although Erfan contests the appropriateness of the score criteria (Extract 6.5.6, further above), as shown in Extract 6.5.7, above, he feels he must submit to the test requirements (it is a thing that I have to do and I have no way to escape). This perception that his capacity to practise agency is constrained by policy prompts negative feelings (It’s not very good, it’s very difficult), particularly given the stakes he attaches to test success (my life depends on it). He views resistance as not only pointless but also self-destructive, thus he adopts the strategic approach already described above, resisting negative thinking (I don’t want to think about it... ...I don’t want to mess my mind around with that, I don’t want to blame anyone).

As shown in Extract 6.5.8, below, Erfan’s self-appraisal of his ability to reach his score goal on the first attempt is unequivocally negative. Because of the high stakes involved, he sees repeat test attempts as the only viable course of action.
Extract 6.5.8

“[R: Do you think that you’ll get the score you need?] No, no, 100% not. [R: So, what will you do after that?] Register again. Every month. It’s the only way to, it’s the only thing I can do” [Interview 1, 533-541]

Extract 6.5.8, above, illustrates a relationship between themes ‘self-appraising’ and ‘practising agency’ that derives from Erfan’s perceptions of and feelings about the test and score requirements, set out in this subsection. Given Erfan’s perceptions of the test as inconsistent (Extract 6.5.5), of the score requirements as exclusionary (Extract 6.5.6), and of the profoundly high stakes attached to score success (Extract 6.5.7), he decides on a strategy of repeat test attempts to exploit or at least mitigate the impact of test inconsistency (more attempts mean more chances of encountering favourable test materials). Moreover, he strives to resist negative thinking as a means of remaining focused on achieving his goals (Extracts 6.5.6 & 6.5.7).

As shown in the next subsection, ‘Experiencing repeat test encounters’, he maintains these two key strategies throughout his test-taking journey. To conclude this subsection, Table 6.9, below, summarises the relationships between themes that characterise Erfan’s strategic approach to the test.

Table 6.9. Erfan’s strategic approach to the test: Linking perceptions, feelings and actions

<table>
<thead>
<tr>
<th>Understanding target scores</th>
<th>Responding affectively (negative)</th>
<th>Self-appraising (negative)</th>
<th>Practising agency (Strategic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>it’s not a very standard test [Extract 6.5.5]</td>
<td>It’s not very good, it’s very difficult ... my life depends on it [Extract 6.5.7]</td>
<td>I don’t want to think about it [Extracts 6.5.6 &amp; 7]</td>
<td></td>
</tr>
<tr>
<td>I just see it as a very fine filter [Extract 6.5.6]</td>
<td>No, no, 100% not [Extract 6.5.8]</td>
<td>I don’t want to blame anyone [Extract 6.5.7]</td>
<td>Register again. Every month [Extract 6.5.8]</td>
</tr>
</tbody>
</table>
6.5.5 Experiencing repeat test encounters

As mentioned at the outset, Erfan made four test attempts within a six-month period as part of his effort to reach his target scores. He scored below IELTS 7 for writing on his first three attempts, and this was the skill area that he focused mainly on in his test preparation. On all four attempts, Erfan achieved 8 or above for listening and at least 7 for speaking and reading. On the third attempt, Erfan scored 8 on all skills except writing. On his fourth and final attempt, as mentioned at the outset, he scored a mix of 7s and 8s. At this point he decided to submit an expression of interest with his existing points test total, rather than pursuing scores of IELTS 8 in all skills. Erfan’s scores across repeat attempts are shown in Table 6.10, below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>Writing</td>
<td>6</td>
<td>6</td>
<td>6.5</td>
<td>7</td>
</tr>
<tr>
<td>Listening</td>
<td>8</td>
<td>8.5</td>
<td>8</td>
<td>8.5</td>
</tr>
<tr>
<td>Reading</td>
<td>7.5</td>
<td>7</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

A central theme in Erfan’s accounts of his experiences across repeat test encounters, as in his accounts of experiencing learning and using English and his initial approach to the test, is ‘self-appraising’, which interacts with other themes in ways consistent with those already demonstrated. The theme ‘interpreting score outcomes’, which becomes relevant as he begins taking the test for permanent residency, also derives from central processes of self-appraising. Interpretations of score outcomes stem, in part, from Erfan’s perceptions of inconsistencies between self-appraisals of his capacity to reach target scores, which improve with test preparation and practice, and self-appraisals of his ability to use English to establish social relationships with native speakers of English, which remain negative regardless of test outcomes. These interpretations interact with his understanding of target scores to prompt different feelings and actions throughout Erfan’s test taking journey, as demonstrated below.
6.5.5.1 *After the first two test attempts*

By the second interview, Erfan had taken the test twice in two months, as per his strategic plan. His interpretation of his score outcomes across his first two attempts is presented in Extract 6.5.9, below. As shown in Extract 6.5.9, he attributes his score outcomes for speaking, reading and listening to improvements in his test performance, and his poor writing scores to a lack of improvement on that skill, thus evidencing a relationship between themes ‘self-appraising’ and ‘interpreting score outcomes’. However, he is careful to note that he does not view these scores as being related to his ability to use English outside of the test.

**Extract 6.5.9**

“*Apart from writing I think that I see that it’s improving. It’s not the language itself it’s just a skill that you have to have for the exam, so it’s improving. But for the writing I can’t see much, much of improvement*” [Interview 2, 123-128]

Erfan perceives the test to represent a domain of English use in and of itself, and his self-appraisals of test performances are separate and mutually exclusive from his self-appraisals of his ability to use English in real world contexts. In Extract 6.5.10, below, Erfan reiterates this view by alluding to an inconsistency between his positive speaking score outcomes and negative self-appraisals of his ability to use English in social interactions.

**Extract 6.5.10**

“I don’t see it’s very good way of measuring the amount of, I don’t know, the ability of English a person has, it’s some sort of framework for testing, so someone might get a very good result on the test but is could cannot speak, cannot interact very well, and the opposite might also be true” [Interview 2, 517-521]
These processes of ‘self-appraising’ and ‘Interpreting score outcomes’ interact to prompt Erfan to take particular action. He decides to prepare for the writing test by rehearsing performances in response to IELTS practice writing tasks, under test conditions, and comparing performances with target score samples. These rehearsals represent practising strategic agency. Attempts to model exemplar performances reflect a perception that score outcomes correspond to performance attributes that are specific and exclusive to the test, and these efforts are directly oriented towards achieving target scores (see Extract 6.5.11).

**Extract 6.5.11**

“First I write an essay with the timer on and then I compare the sample answer with my answer. It’s usually terrible, it’s usually a disaster ... ... obviously it’s not as good as it should be” [Interview 2, 370-378]

As shown in Extract 6.5.11, above, the strategic action Erfan takes (i.e. rehearsing and comparing his writing performances to sample performances) reinforces his understanding of target scores and feeds back into his process of self-appraising, as Erfan situates the quality of his own writing in relation to the performance quality he perceives to be associated with target scores.

The relationships between themes outlined above in Extracts 6.5.9 to 6.5.11 is summarised, below, in Table 6.11.
Table 6.11. Summary of thematic relationships: linking Erfan’s perceptions and actions

<table>
<thead>
<tr>
<th>Self-appraising</th>
<th>Interpreting scores</th>
<th>Practising agency (Strategic)</th>
<th>Understanding target scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking test/social domain:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflicting</td>
<td>≠ ability to use English</td>
<td>Rehearse &amp; Compare</td>
<td>IELTS 7/8 = level above that which he possesses [implicit]</td>
</tr>
<tr>
<td>Writing test:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>It’s not the language itself</td>
<td>I write an essay with the timer on and then I compare the sample answer with my answer</td>
<td></td>
</tr>
<tr>
<td>I can’t see much, much of improvement</td>
<td>= quality of test performance it’s just a skill that you have to have for the exam</td>
<td></td>
<td></td>
</tr>
<tr>
<td>It’s usually terrible, it’s usually a disaster ... ... obviously it’s not as good as it should be</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Extract 6.5.12, below, highlights a further interaction between themes ‘self-appraising’ and ‘understanding target scores’, as the difference Erfan perceives between the quality of his writing and the quality of writing associated with target scores prompts him to again question the appropriateness of the English requirements for permanent skilled migration.

Extract 6.5.12

“It shouldn’t purely filter people according to their language ability. So, it’s a professional migration ***I shouldn’t be expected to be an English writer. I can, I can write technical report very well but I can’t write an essay about a social thing. So, the irrelevant thing is this bit *** it’s not fair to expect me to have the same sort of ah skills as a native speaker. For me that all the thing that I do, I just write the technical report, I enter a document in some piece of software *** and these things are true even in my first language, I don’t write anything in my first language as in just expressing myself or discussing an issue. I don’t speak publicly” [Interview 2, 531-575]
As shown in Extract 6.5.12, above, Erfan understands target scores to align with an unrealistically high level of English ‘native speaker’ writer proficiency, which he feels that second language users are not able to achieve (*I shouldn’t be expected to be an English writer; it’s not fair to expect me to have the same sort of ah skills as a native speaker*). He contrasts a positive self-appraisal of his ability to fulfil the writing demands in his job (*I can write technical report very well*) with a negative self-appraisal of his ability to produce a target score writing test performance (*I can’t write an essay about a social thing*) to highlight that the test lacks relevance to the English demands he faces in his work and life in Australia. He further asserts the irrelevance of the test tasks by claiming they bear no resemblance to tasks he has experienced outside of the test, even in his first language (*these things are true even in my first language, I don’t write anything in my first language as in just expressing myself or discussing an issue*). These perceptions are consistent with his earlier understanding of target scores as discriminatory, a ‘filter’ set to exclude on the basis of arbitrary and unfair language requirements (see Extract 6.5.6, earlier).

Extracts 6.5.13 to 6.5.15, below, illustrate how themes ‘self-appraising’ and ‘understanding target scores’, highlighted above in relation to Extract 6.5.12, interact with themes ‘practising agency’ and ‘responding affectively’ to characterise how test consequences emerge in Erfan’s case. As shown below, Erfan’s perceptions of the test and his understandings of scores lead to decisions and actions that are counterproductive, both personally and professionally, but that appear nonetheless necessary in order to overcome the test score obstacle that he faces. Erfan’s perception that the test is irrelevant to the English demands he faces in his life in Australia, shown in Extract 6.5.12, above, is associated with a perception that using English in social and work contexts is not a useful form of test preparation.

Accordingly, Erfan decides he must forgo opportunities to engage in social interactions in order to maximise the time he has available to spend on practising writing test performances, as shown in Extract 6.5.13.
Extract 6.5.13

“Ah, I don’t know what can I do, actually I wanted to engage in a social club. Initially I wanted to be engaged in my free times but right now I think it’s better if I spend it on practicing more rather than being socialising. So very first priority is that IELTS thing, after that everything is hopefully going to be much easier” [Interview 2, 337-345]

In deciding to avoid social activities, Erfan co-constructs or at least exacerbates negative consequences associated with the test requirement. Firstly, the decision prolongs his social isolation, a major and ongoing concern (see Extract 6.5.14, below). In addition, he misses out on opportunities to develop his English conversation skills, which he feels he needs for building relationships and for accessing a wider range of job opportunities (see Extracts 6.5.1-6.5.3, earlier).

Extract 6.5.14

“About a week ago I was feeling very bad *** my future, my, my state is not... I don’t know what will happen to me *** I can’t imagine to go back after this four years spending here. This is the, this is the, my first concern. The other that is, that I wanted badly to do PhD here but because of that whole PR thing I can’t do that... so that I feel that I’m losing time for that. The other one is probably the financial pressure... and my fate in social life sometimes it makes me very anxious and I wonder how things led to here *** I really hope that it’s just temporal because I hate to see that in the time that everything is sorted out, I got my PhD and PR and I got a good income I still living alone by myself, it’s not great” [Interview 2, 448-497]

As demonstrated in Extract 6.5.14, above, Erfan’s prolonged social isolation and the general uncertainty created by the score requirements are associated with negative affective responses (I was feeling very bad; my future, my, my state is not... I don’t know what will happen to me; my fate in social life sometimes it makes me very anxious), and represent negative impacts on Erfan’s sense of well-being.
Furthermore, in the lead up to the third test attempt, Erfan decided to take additional action to increase the time resources available for test preparation, and asked his employer to reduce his work hours (see Extract 6.5.15, below). This decision again illustrates Erfan’s commitment to prioritising his test efforts, regardless of the personal (see Extract 6.5.14, above) and, in this case professional impact on his life. As shown in Extract 6.5.15, he began to view time spent at work as hindering his efforts to achieve score success. In response to this perception, he practises tactical agency to eliminate this newly identified obstacle.

**Extract 6.5.15**

“I send an email to my supervisor and asking him to change my working hours *** So because my plan was if, if he wouldn’t accept this uh my condition I just leave here and work four hours in McDonalds or something like that to be, to free up my time during the day. Because there is no point for me to working here without making any progress” [Interview 3, 128-176]

Again, this action response conflicts with Erfan’s priorities beyond the test and arguably, with the aims of skilled migration policy selection criteria, one of which is supposedly to prioritise work experience in Australia in addition to language and other professional skills (see Chapter 2.4). Besides adding to the financial pressure Erfan suffers, the perceived need to prioritise test preparation over accumulating professional experience thus creates a further absurdity; Erfan is foregoing experience that would be likely to improve access to employment opportunities in order to rehearse writing skills that he feels are completely irrelevant to his professional domain.

**6.5.5.2 After test attempt 3**

Erfan’s perceptions of his test experiences after the third IELTS attempt were, for the most part, consistent with his account of the two previous test attempts, apart from a positive shift in his self-appraisals of his writing test performance (Extract 6.5.16).
Extract 6.5.16 illustrates a link between themes ‘self-appraising’ and ‘interpreting score outcomes’, as shown below.

**Extract 6.5.16**

“I got 6.5 but it was a very good essay so I should have got at least 7 for that essay but because I didn’t address the question directly I lost the mark for task achievement” [Interview 4, 153-156]

Extract 6.5.17, below, illustrates a relationship between the themes above and the theme ‘responding affectively’. In Extract 6.5.17, Erfan is comparing this most recent writing performance to previous experiences, and attributes the positive shift in his self-appraisal to changes in his mood. From the outset, as shown earlier under the subheading ‘Before test attempt 1’, Erfan identifies resisting negative thinking as key form of strategic agency linked to achieving his test score goals. However, as he progresses though repeat test encounters, the weight of the uncertainty of his situation and the pressure and loneliness he feels threaten his ability to remain positive about the test. As shown, Erfan explicitly locates fluctuations in his mood as a source of variations in his writing test performances, associating negative feelings with poor performances, and being “in a good mood” with good performances.

**Extract 6.5.17**

“I think it depends on various of things, first of all it depends, hugely depends on my mood, so if I, if I’m not feeling right I write pretty crappy things. Yeah but, yeah if I’m in a good mood and if the topic is something that I really have something to say it’s easy...” [Interview 4, 200-204]

Extract 6.5.18 further highlights the impact of ongoing uncertainty about his future on Erfan’s psychological and emotional state. Managing his affective responses to the uncertainty created by the test score requirements becomes a major issue for Erfan over time, as he feels increasingly incapacitated by feelings of anxiety, as shown below.
Extract 6.5.18

“I felt really terrible and it lasts for I don’t know a month or two...Yeah, sometimes it just happens. I really looking for an escape from this situation, it’s...I don’t know” [Interview 4, 542-558]

6.5.5.3 After test attempt 4

As noted earlier, Erfan achieved at least IELTS 7 on all skills on his fourth test attempt, and decided to submit his expression of interest for permanent residency with these scores. After his four test experiences, despite improvement in his test scores, he maintains the perception that score outcomes are not related to his actual ability to use English in real world contexts and continues to situate his (and other ‘foreigners’) English skills in a deficit position relative to the skills of ‘native’ speakers. As illustrated in Extract 6.5.19, he feels vulnerable to the judgment of native speakers, an issue which, over his time in Australia, has hindered his efforts to realise social integration with those he perceives to be members of the local community.

Extract 6.5.19

“I just feel pressure when I’m talking to native people...native speaking people. I just feel, I, I, I ah I don’t want to make silly mistakes...because it might give them a wrong impression of myself” [Interview 5, 340-344]

Not surprisingly, as shown in Extract 6.5.20, Erfan associated overcoming the test score obstacle with a transformation in his emotional state, which, as highlighted above, had become increasingly disturbed over the course of his test-taking journey.

Extract 6.5.20

“I was feeling good about the exam but there is always something deep down in your heart, which is, which always worries you, what happened what happens if you fail this time again. Yeah, but it didn’t happen...Ah, I just yeah, felt a very deep happiness” [Interview 5, 16-25]
6.5.6 Summary of Erfan’s experiences

As demonstrated in this section, Erfan perceives the English skills involved in engaging in social interactions with ‘native’ speakers to be distinct from those involved in engaging with other ‘foreigners’. He also distinguishes between communicating in English in the contexts of study and employment, which, according to his self-perceptions, he manages effectively, and using English to make friends and build social networks with members of the local population, which he feels unable to do. These self-appraisals derive from his feelings about interacting in different contexts; he feels uncomfortable in conversations with ‘native’ speakers so he negatively self-appraisal his English skills in the social domain. By contrast, he feels perfectly comfortable communicating in the contexts of university and his workplace, which are linked to positive self-appraisals in these domains.

Erfan’s real world experiences conflict with his test experiences, which informs his understanding of target scores and his interpretations of score outcomes. Throughout the process of engaging in repeat test attempts, Erfan viewed his English skills as stable and unchanging, regardless of his test experiences and score outcomes. Already seeing himself as capable of using English for the purpose of skilled migration, he perceived test tasks to be irrelevant to the communicative demands he faces in his life in Australia. He viewed target scores and his own score outcomes as arbitrary and disconnected from the level of English he needed and possessed, respectively. These perceptions were associated with suspicions that the use of test scores in skilled migrant selection processes masked a discriminatory agenda, serving to filter out and exclude those from ‘undesirable’ cultural and linguistic backgrounds. Erfan’s perceptions of score meanings, in turn, shaped his processes of self-appraising and responding affectively specifically in relation to the test and score requirements, and influenced his decisions and actions as he progressed through his test taking agency. His attempts to exercise agency and play a role in determining his own fate, in spite of the constraints imposed by the policy and the test, led to decisions and actions that were, at times, inconsistent with his broader
goals, such as establishing a social network and accumulating work experience in Australia.

To conclude the section on Erfan, Figure 6.2, below, illustrates the interaction between his experiences outside the test, in social, study and work domains and his experiences with the test, as described above. As indicated by the solid border around themes in real world domains, Erfan’s self-appraisals and feelings about his ability to use English remained relatively stable. Test scores represent the point of insertion of the test into Erfan’s life, and are thus situated in the overlap between the two circles, representing the real world and test domains which Erfan views as distinctly separate. The tension that emerges as understandings of target scores and interpretations of score outcomes conflict with Erfan’s real world experiences leads to a view that score meanings are arbitrary and irrelevant, as noted above. This prompts an iterative cycle of self-appraising, responding affectively, and practising agency, in which each process is contingent on test outcomes, as indicated by the broken line borders in Figure 6.2.

Figure 6.2: Summary of thematic relationships (Erfan)
6.6 Sharon

6.6.1 Introduction

Sharon is from Sri Lanka and identifies Sinhala as her mother tongue, but has spoken and been educated in English since early childhood. Upon finishing high school in Sri Lanka, she came to Australia in 2007 to undertake a Bachelor of Science. After successfully completing her Bachelor degree in 2010, she decided to undertake an honours year specialising in Immunology. She graduated from this in 2011, obtained a temporary post-study work visa and gained full-time employment in her field as a researcher in an immunology laboratory at a university. Sharon had been in her job for 6 months and living in Australia for over four years at the time she joined this study. Her intentions were to acquire permanent residency and to pursue PhD studies under the supervision of the professor heading the laboratory where she worked.

Sharon’s country of origin is not one of the countries identified as English-speaking within Australian immigration policy, which meant that, despite speaking English for most of her life, Sharon was obliged to provide evidence in the form of test scores of her English proficiency in order to obtain international student visas to complete her tertiary qualifications in Australia. She sat IELTS twice for this purpose, once before arriving in Australia in 2007, and again in Sri Lanka after completing her three-year Bachelor degree so that she could renew her visa and return to complete her honours year. Her original IELTS scores were, by this time, no longer valid. Both times she scored IELTS 7.5 or above on all skills, higher than the IELTS 6.5 required. She was able to reuse the second set of scores in her application for a temporary post-study work visa, but because of the policy changes introduced in 2011, described in chapter 2, these scores were not high enough to deliver the points she needed to meet the minimum total required for eligibility to apply for permanent residency – she now needed IELTS 8 in all skills. As noted in Chapter 5 (see footnote to Table 5.1), Sharon required a higher score than the other three migrants because she went directly from school in Sri Lanka to university in Australia,
whereas the others were able to gain credit in the Points Test for work experience in their home countries. As a result, they only needed 10 points for English to meet the Points Test pass mark, whereas Sharon needed 20 points.

Sharon made two test attempts in an effort to achieve these scores, in February 2012 and again in April 2012. Although she achieved overall scores of at least IELTS 8 on both occasions\(^2\), she was unable to achieve higher than IELTS 7.5 on the writing subtest. As will be described further below, when she received the scores from her second encounter, Sharon decided to abandon the idea of further test attempts. Instead she sought legal advice and was alerted to a new state-sponsored visa option, available in the state where she was living and working. She was advised that she could meet the requirements for this particular visa with scores of IELTS 7, although IELTS 8 would improve her chances of selection. She decided to lodge her permanent residency application in July 2012 with her existing IELTS scores and was granted one of these visas in mid-2013, approximately 1 year later. Prior to becoming a permanent resident, Sharon had applied for and been granted an international student scholarship to pursue PhD studies.

Four interviews were conducted with Sharon, three in 2012 in March, April and May, and a fourth interview in February 2013. When the first interview was conducted, Sharon had already made her first test attempt for permanent residency purposes, and had completed an IELTS preparation course after receiving her test scores. The second interview was conducted just prior to her second test attempt, and the third interview was conducted after the second test attempt but before Sharon had received her scores. After this attempt, which turned out to be unsuccessful, Sharon avoided further interviews as she was starting her PhD research program and had heavy work commitments. At the time when the final interview was conducted, she was waiting to hear the outcome of her permanent visa application, which she had submitted over six months earlier.

---

\(^2\) Sharon reported that she received IELTS 8s and 8.5s on the other skills, but could not remember the specific scores by skill.
Sharon’s experiences are presented below under two headings: ‘Experiences learning and using English’ and ‘Experiencing repeat test encounters’.

6.6.2 Experiences learning and using English

As already mentioned, Sharon started speaking English in early childhood and completed a significant part of her education in Sri Lanka in English (Extract 6.6.1).

Extract 6.6.1

“We spoke English at home since I was very little and then I went into an English speaking school, though I studied in my mother tongue but I was, I did English literature and all that sort of thing so I was taught English from a very young age” [Interview 1, 119-137]

In addition, she successfully completed tertiary studies in Australia, which included writing a 10,000-word thesis in her fourth and final year, for which she received a high distinction grade. As also mentioned, Sharon was able to gain full-time employment in her field after graduating. Extract 6.6.2, below, highlights Sharon’s sense of legitimacy as an English speaker and illustrates a relationship between themes ‘self‐appraising’ and ‘responding affectively’, as Sharon expresses frustration over the requirement to provide evidence of her legitimacy by way of test scores.

Extract 6.6.2

“It’s pretty frustrating when I mean, yeah I can communicate the whole point is to be able to communicate in English properly and I don’t have a communication problem as such... ...like I use English more often than I would use my mother tongue, so...” [Interview 1, 370-379]

Not surprisingly, Sharon felt that her history in the language as well as her university and employment success represented sufficient evidence of her ability to communicate effectively in English. As shown in Extract 6.6.2, Sharon believes that the purpose of the
test in this policy context should be to ensure potential migrants are able to communicate in English *(the whole point is to be able to communicate in English properly)*. Although she supports test use for this purpose, she feels that the test fulfils no meaningful or justifiable role in her case; she sees herself as a competent English user, as illustrated by her positive self-appraisal *(I can communicate)*, and self-identifies as an established speaker of the language *(I use English more often than I would use my mother tongue)*. These perceptions prompt a negative affective response *(It’s pretty frustrating)*, as mentioned above.

As described in the next subsection, Sharon’s sense of legitimacy as an English speaker remains constant throughout her test experiences. For this reason, she rarely engages in explicitly self-appraising her English skills during her accounts of her repeat test experiences, except to assert the inappropriateness of the test score requirements in her case, as she is doing in Extract 6.6.2, above. Nonetheless, this sense of legitimacy underlies Sharon’s interpretations of score outcomes and understanding of target scores across her two test encounters, and shapes the types of agency she practices to realise her intended migration outcomes, as set out below.

### 6.6.3 Experiencing repeat test encounters

As noted at the outset, Sharon’s score goal was IELTS 8 and she made two attempts at achieving these scores. Both times, her scores on the three skills apart from writing were IELTS 8 or above, but her writing score was IELTS 7.5. A table of score is not presented in Sharon’s case, as she did not provide scores by skill, simply reporting that she achieved a mix of IELTS 8s and 8.5s. On both occasions, Sharon had taken IELTS Academic, even though the General test is also accepted for skilled migration purposes. Her reasons for this were that she felt more familiar with academic writing and also, she felt certain that she had attained a level of mastery in academic writing during her graduate studies, especially through her thesis writing. She thus expected to achieve her score goal on the first test attempt. As described in this subsection, her lack of success prompted her to distinguish between her capacity to use English and her capacity to achieve her target scores, which in turn influenced her perceptions of the test and score requirements. As
is set out below in relation to each of Sharon’s two test encounters, the theme ‘self-appraising’ interacts with themes ‘interpreting score outcomes’ and ‘understanding target scores’ in characterising Sharon’s perceptions of the test. Sharon’s responses to her perceptions of the test, as will be shown, evidence a relationship between these three themes and the themes ‘responding affectively’ and ‘practising agency’.

6.6.3.1 After test attempt 1

As mentioned above, Sharon’s first writing score outcome triggered a new process of self-appraising, specific to the test, as she realised her writing skills did not readily translate to score success. In response to her test result, Sharon sought expert advice to help her conceptualise the attributes of an IELTS 8 writing performance. To this end, she enrolled in an IELTS preparation course. This action represents practising tactical agency; Sharon began to think that skills specific to the test were needed to supplement her writing skills, and she sought out expert information in order to locate an appropriate strategic approach to achieving her score goal.

As shown below, the information Sharon obtained through attending this course helped her identify why her writing performance had not been awarded a score of IELTS 8. Furthermore, advice she received, from both teachers and other test takers, shaped her perceptions of score meanings, thus signalling a relationship between the decision to take an IELTS preparation course (and thus the theme ‘practising agency’) and themes ‘self-appraising’, ‘interpreting score outcomes’ and ‘understanding target scores’, as shown across Extracts 6.6.3 to 6.6.7, below. Furthermore, Extract 6.6.8 illustrates a relationship between Sharon’s self-appraisal of her English skills (as opposed to her test taking skills) and themes ‘understanding target scores’ and ‘practising agency’, as she puts in place a strategic approach to preparation for her next test attempt, based on the understanding of target scores she derives from information and advice obtained via the IELTS preparation course.

Extract 6.6.3, below, highlights a relationship with the action Sharon takes and the themes ‘self-appraising’ and ‘interpreting score outcomes’. As shown, through her
involvement in the IELTS preparation course, Sharon attributes her below target writing score to a failure to adopt the specific style of writing expected on IELTS. This indicates, albeit implicitly, that Sharon’s self-appraising of her test performance becomes negative (she realises her performance was not suited to examiner expectations). Moreover, in signalling this difference between her own style of writing and that required on IELTS, her account suggests that the test construct lacks relevance to the writing she engages in outside of the text.

**Extract 6.6.3**

“I noticed that the reason that their essay writing style is different, like it’s very specific” [Interview 1, 67-69]

Based on her perception that a different style of writing is needed on IELTS, Sharon identified developing effective test taking strategies, rather than learning or improving English skills, as crucial to achieving target scores in writing. Extracts 6.6.4 and 6.6.5, below, illustrate the theme ‘understanding target scores’.

In Extract 6.6.4, below, Sharon identifies an ability to think and write quickly as the main skill needed to achieve IELTS 8, rather than language knowledge. This understanding of target scores is reinforced by information she receives from teachers of the IELTS preparation course, as shown in Extract 6.6.5, further below.

**Extract 6.6.4**

“It’s more of a requirement of speed rather than being um familiar with the language I think. I mean the, the test is easy, it’s just that there’s a lot of pressure simply because a lot of things are squeezed into an hour or so… so that’s why it’s difficult but I think otherwise if you actually sit down and rationally have time to think, it’s not a difficult test” [Interview 1, 411 – 423]
Extract 6.6.5

“Plus when I went for the preparation class, the lecturers, the teachers rather, um who was, who were taking the classes were telling us they themselves have an issue with the test because um of the time pressure. Like, you have to do everything very, very quickly so and even like they were saying even if they were going to sit for it I don’t think they’ll get a very high score either simply because of that fact” [Interview 1, 476-483]

In citing the IELTS preparation course teachers, who apparently suggested that even they would have difficulty producing an IELTS 8 writing performance under test conditions, in Extract 6.6.5, above, Sharon is critiquing the artificiality of the test construct.

Similarly, Extracts 6.6.6 and 6.6.7, below, provide further evidence of a perception of target scores as representing an unreasonably high level. In extract 6.6.6, the theme ‘understanding target scores’ interacts with implicit ‘self-appraising’ to form the basis of an unfairness claim in relation to the score requirements.

Extract 6.6.6

“So, it’s a good way of gauging somebody’s ability to communicate, but what I think is you shouldn’t have a very high score, because after all, um, English is not the mother tongue of most of us so it’s not fair to think us of being of very, very good high standard as maybe a native speaker would” [Interview 1, 469-474]

As shown in Extract 6.6.6, while Sharon views English tests, such as IELTS, to be appropriate tools for verifying that would-be migrants from non-English speaking backgrounds are able to communicate in English, she perceives scores of IELTS 8 to represent an unrealistic and unattainable level of English for people from non-English language backgrounds. In so doing, she disassociates herself from ‘native’ speaker norms and although she has spoken English since early childhood, she aligns herself with the group for whom “English is not the mother tongue”.

170
Extract 6.6.7 demonstrates that Sharon’s perception of target scores (as mirroring idealised native speaker norms) is reinforced by advice she received from teachers in the IELTS preparation course she undertook. These teachers, she explains, were suggesting as a strategy that test takers return to their home countries to take IELTS, to avoid being judged against these unachievable standards.

Extract 6.6.7

“So what I heard was that um, through like IELTS training teachers, um they were saying that it, it’s, a lot of people have come to them saying that they get better scores when they actually sit the, sit for their test through their own home country rather than say for ex- Australia for example, because you, you tend to get compared to like a native speaker with a native accent and that sort of thing so your scores tend to get lower” [Interview 2, 2-9]

Extract 6.6.8, below, highlights a further interaction between themes ‘**self-appraising’**, ‘**understanding target scores’** and ‘**practising agency’**.

Extract 6.6.8

“For me, it’s more like I’m preparing for the exam; I’m not learning anything um but I mean, I don’t know, maybe because I usually sort of converse in English commonly, maybe that’s the reason, but that’s h- just how I feel maybe somebody else would say ‘oh no, I’ve been learning English and I’ve been expanding my knowledge’ but for me it’s more like, sort of, getting ready for the exam and trying to meet what the exam expects of you” [Interview 2, 129-137]

As shown in Extract 6.6.8, above, Sharon sees no value in test preparation beyond achieving the scores she needs to become eligible for permanent residency, given that she self-identifies as an English speaker (*because I usually sort of converse in English commonly*), not as a language learner (*I’m not learning anything*), indicating an implicit positive self-appraisal of her English skills. The perception that her test efforts are
meaningless apart from fulfilling the policy demands derives from an understanding of target scores as representing a style of writing and an approach to writing, in terms of speed, that are irrelevant to the actual writing demands she faced or faces in the contexts of graduate study and employment, respectively. The understanding of target scores she arrives at on the basis of the IELTS preparation course, combined with her sense of legitimacy as an English speaker, prompt her to practise strategic agency as she implements a new approach to reaching her score goals (for me it’s more like, sort of, getting ready for the exam and trying to meet what the exam expects of you).

The interactions between themes that characterises Sharon’s response to her first unsuccessful test encounter, as described above, are summarised in Table 6.12, below.

**Table 6.12. Summary of relationships between themes after Sharon’s first test attempt**

<table>
<thead>
<tr>
<th>Practising agency (Tactical)</th>
<th>Interpreting score outcomes</th>
<th>Self-appraising</th>
<th>Understanding target scores</th>
<th>Practising agency (Strategic)</th>
</tr>
</thead>
</table>
| IELTS preparation course     | their essay writing style is different | Performance not suited to examiner expectations | Construct:  
- Writing speed  
  It’s more of a requirement of speed  
- Native English norms you tend to get compared to like a native speaker with a native accent ... so your scores tend to get lower | I’m preparing for the exam  
- it’s more like, sort of, getting ready for the exam and trying to meet what the exam expects of you |
|                             |                             |                             |                             |                             |
|                             |                             |                             |                             |                             |
|                             |                             |                             |                             |                             |

**Construct:**
- Writing speed
  It’s more of a requirement of speed
- Native English norms you tend to get compared to like a native speaker with a native accent ... so your scores tend to get lower

**Fairness issues:**
- Time constraints they [IELTS prep teachers] themselves have an issue with the test because um of the time pressure
- ‘Native’ standard is exclusionary it’s not fair to think of being of very, very good high standard as maybe a native speaker would
As shown in Table 6.12, above, Sharon interprets her score outcome on the basis of information and advice from the IELTS preparation course, which leads to a distinction between self-appraising of her English skills and self-appraising specifically in relation to the test. This, in turn, informs her perceptions of target scores meanings. Her understanding of target scores relates primarily to aspects of the test construct that she views as irrelevant and inappropriate, and the fairness issues that emerge as a result. Regardless of her critical view of the test, Sharon’s revised perceptions of score meanings (as disassociated from her English skills and irrelevant to the writing demands she faces) form the basis of a new strategic approach to the test.

6.6.3.2 After test attempt 2

Interviews were conducted with Sharon both before and after she received her test scores from the second test attempt. Extracts 6.6.9 to 6.6.12, below, refer to Sharon’s perceptions of her test experiences prior to receiving score outcomes. These extracts illustrate relationships between themes ‘self-appraising’ and ‘responding affectively’ (Extracts 6.6.9 and 6.6.10), and highlight further developments related to the theme ‘understanding target scores’ (Extracts 6.6.11 and 6.6.12), as Sharon’s perceptions are influenced by links she identifies between the information she received in the IELTS preparation course and anecdotal evidence from other test takers.

Extract 6.6.9

“"I mean I’m pretty confident I can do it because I’ve got the scores that I needed before [prior to commencing graduate study] so unless, somebody’s English improves as they’re studying English, it certainly can’t go down, right?” [Interview 3, line 319 - 321]

In Extract 6.6.9, Sharon’s expresses feeling confident about achieving her target score in writing, which implies a positive self-appraisal of her test performance. Her optimism is based on a perception that test score outcomes will mirror improvements in her academic writing skills, which she developed throughout her graduate studies. Thus, having taken a revised approach to the IELTS writing task, she expects a score increase
compared to when she took the test for her student visa, which was prior to writing her thesis.

In contrast to these positive feelings she experienced in relation to her test performance (Extract 6.6.9), as shown in Extract 6.6.10, below, the uncertainty of her migration situation and the apparent pointlessness of the need to meet arbitrary score requirements in her case (see Extracts 6.6.2 and 6.6.8, further above) are associated with negative affective responses; feelings of worry and frustration.

**Extract 6.6.10**

“I’m really sick and tired of you know worrying about this, and having to do this and it’s just…so annoying” [Interview 3, 193-196]

In relation to the theme ‘understanding target scores’, Sharon’s frustration is exacerbated by further uncertainty about the integrity of the test, which emerges on the basis of suggestions made by teachers during the IELTS preparation course. She remains unsure, for example, of the extent to which her chances of score success may be hindered by taking the test in Australia. In Extract 6.6.11, below, Sharon refers to anecdotal evidence she has heard from other test takers and contacts in Sri Lanka, which supports the credibility of advice provided by teachers in the IELTS preparation course, who were recommending that test takers return to their home countries as a way of improving score outcomes (by avoiding being subjected to comparisons with English ‘native’ speaker norms – see Extract 6.6.7, earlier).

**Extract 6.6.11**

“Some were... actually I know a few people who did go back um to do their test and come back so...Yeah and I think one or two I heard did manage to get the score they needed when they sat it through home

***

Oh I’ve heard from people in Sri Lanka too, like it’s easier to score high”

[Interview 3, 523-530; 543]
In addition, Sharon hears rumours about examiners in Australia deliberately scoring performances below IELTS 7 and 8 to increase the profitability of the test. She is unsure of whether or not to believe the rumours, but sees the association of particular scores with high stakes migration outcomes as creating a motive for corruption on the part of testing companies, given the likelihood that test takers will make repeated test attempts rather than abandoning their migration goals.

**Extract 6.6.12**

“Maybe, I don’t know, maybe it’s just this is what I hear people telling but I don’t know how truthful it is, but that they’re sort of not giving you a good mark purposefully so that you’ll be sitting the test over and over again because usually it will be for migration purposes, so you’ll be doing it over and over again and then that way they’ll earn more money” [Interview 3, 547-553]

As shown across Extracts 6.6.9 to 6.6.12, above, prior to receiving her score outcomes from the second test attempt, Sharon’s perceptions of the test and the use of IELTS scores for skilled migrant selection had become increasingly cynical. These perceptions were associated with negative affective responses, regardless of her feeling confident about her chances of achieving target scores. Extracts 6.6.13 to 6.6.15, below, which relate to Sharon’s perceptions after receiving scores from her second test encounter, illustrate interactions between themes ‘interpreting score outcomes’, ‘self-appraising’ and ‘practising agency’, which emerge in light of Sharon’s increasingly negative perceptions of the test and score requirements (‘understanding target scores’). As shown across these extracts, her interpretation of score outcomes triggers a negative shift in her self-appraisal of her capacity to achieve her score goal in writing (Extract 6.6.13). This prompts her to practise tactical agency; she decides to seek advice from a migration lawyer to find ways around the IELTS 8 score requirement she faced, which leads to a decision to abandon further test attempts (Extract 6.6.14). Extract 6.6.15, in which Sharon explains the impact of the uncertainty of her migration situation, sheds further light on her decision to apply for a permanent visa on the basis of her existing scores.
Extract 6.6.13, below, illustrates an interaction between themes ‘self-appraising’ and ‘interpreting score outcomes’ and ‘understanding target scores’. Reflecting on her lack of success on her second test attempt, Sharon attributes her writing score outcome to her inability to adequately deal with the time constraints she faced in the test, as well as to difficulties associated with having to write by hand.

Extract 6.6.13

“I haven’t done exams for a while that’s why maybe, because I used to do really well at them so yeah, I would be able to it’s just that I’ve been, I haven’t like written, done any written work as such for a long time, it’s always with the computer… I’ve realised after sort of doing honours and all it’s always like, I tend to think more, and then write something, as opposed to what I would have done before, where I would have just you know, kind of just gone, gone ahead and written stuff out. So it slows me down…” [Interview 4, 460-485]

As shown in Extract 6.6.13, Sharon situates the skills needed to achieve scores of IELTS 8 as purely test-taking skills that can only be maintained through test practice (I haven’t done exams for a while that’s why maybe, because I used to do really well at them). She associates the behaviours relevant to producing quality writing in the context of her graduate studies and employment (it’s always with the computer…I tend to think more, and then write something) with negative self-appraising of her ability to succeed on the test (so it slows me down). This further evidences a perception that the style of writing required in the exam is inconsistent with the writing skills she developed through her studies, and highlights the absurdity of the test requirements in her case.

Extract 6.6.14, below, illustrates a relationship between Sharon’s self- and test-perceptions, described in relation to Extract 6.6.13, above, and themes ‘practising agency’ and ‘responding affectively’. Sharon no longer believed she possesses the skills needed to achieve her score goals, although, as noted above, her negative self-appraisal is restricted to her capacity to engage the required test-taking behaviours. She therefore responded by practising tactical agency; she decided to meet with a migration lawyer,
as a means of exploring if other avenues were available for gaining permanent residency. The migration lawyer alerted Sharon to the possibility of applying for a state-sponsored permanent visa, which had recently become available in her skill area. The lawyer advised her that although she met the minimum requirements for this visa with her existing scores, the same competitive selection process applied whereby, if the number of applicants exceeds the number of visas available, applicants are ranked according to points total. The lawyer suggested that she would have a better chance of selection if she repeated the test and achieved scores of IELTS 8. She decided, however, not to take this advice, as shown in Extract 6.6.14, below.

**Extract 6.6.14**

“I was in fact told [by the migration lawyer] to sort of repeat the test and I was like, nah I can’t do it again… … I’m not going to do it again… …it’s really time consuming and it’s quite expensive it’s like 400 dollars or something a test and yeah, it’s, it’s pretty intensive and I just don’t like exams anymore”
[Interview 4, 440-458]

Sharon’s decision to reject the lawyer’s advice and abandon further test-taking efforts, shown in Extract 6.6.14, represents another instance of practising tactical agency that derives from Sharon’s process of self-appraising, described further above in relation to Extract 6.6.13. Score outcomes from the second test encounter convinced Sharon that achieving IELTS 8 would, at best, require repeated attempts in order to re-establish prior test-taking habits, without any guarantee of success. She thus decides to risk applying for a permanent visa on the basis of her existing scores. In Extract 6.6.14, she attributes this decision to the negative time (it’s really time consuming) and financial impacts of test taking (it’s quite expensive), as well as to her affective response to her lack of success in realising her score goals thus far (I just don’t like exams anymore).

Extract 6.6.15, below, may provide further explanation for Sharon’s decision to risk applying for permanent residency with her existing scores. As shown above, her test experiences created unforeseen uncertainties as she became aware that score success depended on factors outside her control, including examiner subjectivity, as well as on
test-taking skills that she could no longer readily access. Extract 6.6.15 demonstrates that the uncertainty of her visa status created a sense of transience that affected Sharon’s decision-making outside the test context and prevented her from establishing stability in her everyday life. She finds this “quite unsettling”, and, on the basis of her comments in Extract 6.6.14, above, she likely perceives that continuing with repeat test taking will only exacerbate the negative impacts of the lack of certainty already affecting her.

**Extract 6.6.15**

“I live a bit of a nomadic lifestyle, whereas if I knew that I was able to stay here I would have, you know, moved to a bit of a more homely environment *** It’s a bit unsettling, quite unsettling, so pretty much I’ve been living in student accommodation even after I finished my degree. I would have loved to sort of um you know moved to a bit of a more homely place, I didn’t quite get to doing that because I wasn’t quite sure whether I can stay here for good” [Interview 4, 535 – 561]

Sharon’s responses to her unexpected score outcomes on the second attempt are summarised in Table 6.13, below.
Table 6.13. Summary of relationships between themes after Sharon’s second test attempt 2

<table>
<thead>
<tr>
<th>Responding affectively</th>
<th>Self-appraising</th>
<th>Understanding target scores</th>
<th>Practising agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I’m pretty confident I can do it</td>
<td>Implicit: positive self-appraisal of writing performance</td>
<td>one or two did manage to get the score they needed when they sat it through home</td>
<td></td>
</tr>
<tr>
<td>I’m really sick and tired of you know worrying about this, and having to do this and it’s just...so annoying</td>
<td></td>
<td>they’re not giving you a good mark purposefully/ you’ll be doing it over and over again/that way they’ll earn more money</td>
<td></td>
</tr>
<tr>
<td>After scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(I tend to think more, and then write something) So it slows me down</td>
<td>Implicit: Irrelevant construct</td>
<td>Tactical</td>
<td>Abandoning further test attempts</td>
</tr>
<tr>
<td>I just don’t like exams anymore</td>
<td></td>
<td>Seeing a lawyer</td>
<td>I’m not going to do it again</td>
</tr>
</tbody>
</table>

6.6.4 Summary of Sharon’s experiences

Information and advice gained during the IELTS preparation course informed Sharon’s conceptualisations of the performance attributes associated with writing scores of IELTS 8, leading to the view that test skills were central to success but also alerting her to the possibility that scores were derived from implicit and indeterminate ‘native’ speaker norms. Although Sharon perceived both her English and her academic writing skills to be highly proficient and well established, she began to doubt her capacity to achieve IELTS 8 due to a lack of test practice, which exacerbated the difficulty she associated with task time constraints. Furthermore, she increasingly saw such scores as inaccessible to non-‘native’ speakers, which contributed to a growing sense of unfairness, and prompted feelings of impatience and frustration in relation to the test and the policy, particularly as Sharon began to question the integrity of examiners and to suspect ruthless, exclusionary, and profit-seeking motives. These perceptions and feelings shaped Sharon’s decisions and actions, as demonstrated above.
As shown across Sharon’s two test attempts, her self-appraisals of her test performances shifted as she integrated information and advice from teachers and other test takers about target score meanings with her interpretations of score outcomes in an iterative cycle. Figure 6.3, below, illustrates these relationships between themes that characterised Sharon’s perceptions in relation to the test, and the link between these perceptions and her feelings and actions. As indicated by the orange shading, Sharon views the test as a distinct context with distinct style and norms of language use, and her self- and test-perceptions in this context do not impact her sense of legitimacy as an English user or her self-appraisals of her ability to use English in real world contexts.

**Figure 6.3. Summary of thematic relationships (Sharon)**
6.7 Yuki

6.7.1 Introduction

Yuki is from Japan and identifies Japanese as her first language. She completed nursing qualifications in Japan and worked there in her profession for three years before coming to Australia for the first time in 2008. Yuki initially decided to move to Australia to study English, with the intention of gaining the language skills needed to access employment opportunities in her field and of remaining permanently. She arrived on a one-year student visa, having enrolled in an English language course of the same duration. While she was completing the course, she was also attempting to satisfy the English test score requirements for professional registration as a nurse. Her aim was to meet the requirements and locate skilled work before her temporary student visa expired, so she could transition directly to a permanent skilled visa.

However, by the end of the one-year period, Yuki had not been able to achieve the scores needed to register to work as a nurse and so was unable to apply for permanent resident status. This marked the beginning of a complex migration trajectory, as she was then required to leave Australia and seek alternative visa options. She returned to Japan for 8 months before deciding to apply to undertake undergraduate nursing studies in Australia. Even though she already possessed nursing qualifications, tertiary study in Australia offered an alternative pathway to permanent residency, and Yuki thought replicating her qualifications at an Australian institution would also provide an opportunity to further develop English skills relevant to work in her profession.

Yuki received a conditional offer to study nursing at a university in Australia, but she was unable to achieve the scores of IELTS 6.5 needed to satisfy the English language requirements for entry. She decided to defer enrolment for a year, and instead returned to Australia on a working holiday visa for the year of her deferment. She hoped to use the time to improve her English, with the aim of achieving the score requirements for university entry before her visa expired. Over the year, she spent around six months
working full time in a restaurant and undertook IELTS preparation classes for three months in total, the maximum study time allowed under her visa conditions. Over this period, she also attempted IELTS ten times, but without success.

With her working holiday visa and university offer soon to expire, Yuki again returned to Japan in late 2010. Back in Japan, she made further attempts at IELTS and was able to achieve the scores she needed for university entry. She returned to Australia to begin an undergraduate nursing degree in early 2011. Having received credit for her Japan-based qualifications, her full-time course length was reduced from three years to two. She completed her degree in late 2012 and graduated in March 2013.

Already a seasoned test taker, Yuki began attempting to meet the English test score requirements for permanent residency in 2012, during the final year of her nursing degree. Her first test attempt for this purpose was in May that year, on the Occupational English Test (OET). While the other three migrants in this study were only taking IELTS for the purpose of obtaining permanent residency, Yuki was attempting both IELTS and the OET. The OET is a specific purpose English test designed for use in the health professions, including nursing. As noted in chapter 2, at the time this study was conducted, the OET and IELTS were the only two English tests recognised for the purpose of skilled migration in Australia. Yuki’s score goal was IELTS 7 or OET B. IELTS 7 and OET B are considered equivalent in the points test; both representing ‘proficient’ English, associated with 10 points.

The first interview with Yuki took place in August 2012. At this time, she was studying full-time, juggling two part-time hospitality jobs and engaged in ongoing English test preparation and test-taking efforts. She had made three unsuccessful test attempts for permanent residency purposes (2 OET and 1 IELTS attempt) before the first interview, and between the first and second interviews, she made three further test attempts, twice again at the OET and once again at IELTS. After the second interview, she made another three test attempts, with an IELTS attempt in between two OET attempts. On the second of these OET attempts, in January 2013, she finally achieved her target scores.
This was her ninth English test encounter within a nine-month period over 2012/13 and marked more than 20 attempts since deciding to undertake an undergraduate degree in Australia to pursue the study-migration pathway to permanent residency. These test encounters were in addition to the attempts she made during her first year in Australia in 2008 and in Japan afterwards, the number of which she was unable to recall. The final interview was conducted in March 2013, just after she had applied for nursing registration and just before she submitted her application for a permanent visa. Yuki was granted Australian permanent residency in March 2014.

Yuki’s experiences learning and using English since arriving in Australia, outlined below, are inseparable from her repeat test-taking experiences, given that she encountered test score obstacles at various stages in her convoluted migration trajectory and dedicated herself to making as many test attempts as possible. For the purpose of clarity and to maintain consistency with the presentation of the three previous individual cases, the subsection, ‘Experiencing learning and using English’, refers to Yuki’s accounts of experiences prior to her most recent direct attempts to gain the scores needed for permanent residency, which commenced in 2012. The following subsection ‘Experiencing repeat test encounters’, as in the previous three individual cases, deals solely with these latter repeat-test taking experiences, nine in total, since May 2012.

### 6.7.2 Experiences learning and using English

Compulsory English classes had been part of Yuki’s high school education in Japan, but she perceived her language learning during her time at school to be minimal. She also spent four months studying English in the United Kingdom after finishing her nursing studies in Japan. As noted above, upon arriving in Australia, Yuki’s experiences of learning and using English became intertwined with test taking experiences, and her migration trajectory became complicated by her inability to achieve the English test scores needed to directly pursue a skilled migration pathway.
Themes ‘self-appraising’, ‘interpreting score outcomes’, and ‘practising agency’, as illustrated in this subsection, characterise Yuki’s accounts of her pre-2012 experiences. As shown across Extracts 6.7.1 to 6.7.6, below, these themes are interrelated; Yuki’s self-appraisals and interpretations of scores over time influence her decisions and actions as she seeks to overcome English language obstacles and realise her migration intentions. Furthermore, self-appraising and interpreting score outcomes are shaped by self-perceptions and by perceptions about language learning and achieving test success. As demonstrated in this subsection, Yuki consistently situates herself as responsible for test outcomes. She does not criticise the test and does not question the meaningfulness of scores. Rather, she assumes that scores reflect her English skills and she interprets poor score outcomes to indicate a deficit on her part; inadequate language development, which she must take action to remedy in order to reach her professional and migration goals.

By the end of her first year in Australia, as previously noted, Yuki was unable to achieve the scores needed to register as a nurse (IELTS 7 or OET B in all four skills), a necessary precursor to applying for a permanent visa. Extract 6.7.1, below, illustrates a relationship between themes ‘self-appraising’ and ‘interpreting score outcomes’.

**Extract 6.7.1**

“At the time my English wasn’t really good, so I couldn’t speak properly, so yeah, that’s why I couldn’t pass” [Interview 1, 9-10]

As shown in Extract 6.7.1, Yuki’s uses negative self-appraisals of her English skills during this period (my English wasn’t really good; I couldn’t speak properly) to explain her test scores, which reflects an interpretation of score outcomes as meaningful indicators of her English abilities at the time (that’s why I couldn’t pass).

Extracts 6.7.2 and 6.7.3, below, illustrate a relationship between these themes and the theme ‘practising agency’. In light of her inability to achieve the scores required for professional registration, while back in Japan, Yuki sets the new goal of undertaking a second nursing degree in Australia, which is associated with a lower score requirement
(IELTS 6.5). As demonstrated in Extract 6.7.2, this shift in focus represents practising tactical agency; Yuki’s overarching goal remained the same, to become a nurse and a permanent resident in Australia, but she establishes this new, intermediary goal as a means of overcoming the English language obstacle that test scores helped her identify.

**Extract 6.7.2**

*I decided, because I couldn’t pass English test for registration as a nurse, so um I decided to go to uni... So the first goal become going to uni* [Interview 1, 16-19]

After engaging in self-directed study and test attempts over an eight-month period in Japan, Yuki was unable to achieve the English scores of IELTS 6.5 needed for university entry. As a result, she decided to defer her studies for a year and to re-enter Australia on a twelve-month working holiday visa (Extract 6.7.3).

**Extract 6.7.3**

“*Then, um, so I came back to Melbourne as a working holiday, and I was thinking to go to uni while my working holiday is in um, how do you say? Ah, useful? How do you say? It’s like a…* [R: OK, so while it was still valid?] *Valid, yeah, that word, I was thinking that word*” [Interview 1, 112-115]

Taking this further detour in her migration pathway again represents practising tactical agency. As shown in Extract 6.7.3, above, Yuki’s aim had been to meet the score requirements for university entry before her working holiday visa expired; a working holiday visa provided a means for Yuki to re-immerse herself in the English-speaking environment of Australia, which she believed would help her improve her English skills to the required level.

Over this year long visa period, Yuki tried IELTS ten times, without success. As shown in Extract 6.7.4, below, although her speaking and listening scores improved, she was unable to achieve the score she needed for writing.
In Extract 6.7.4, above, a relationship between themes ‘self-appraising’, interpreting score outcomes’ and ‘practising agency’ is further evidenced. As shown, Yuki associates positive test scores outcomes over this period with a positive shift in self-appraisals of her speaking and listening skill development, which again suggests that she is interpreting score outcomes as meaningful indicators of her English level. Yuki links her speaking and listening skill development to her immersion in an English-speaking environment, the outcome of practising tactical agency, as noted earlier, and implicitly to the regular skill practice this action enabled. She compares these positive developments to her lack of improvement in writing, a skill she rarely practised in the context of her restaurant work and everyday life in Australia over this period, and attributes her failure to achieve score success to this perceived skill deficit.

Table 6.14, below, provides a summary of the relationships between themes demonstrated so far across Extracts 6.7.1 to 6.7.4.

Table 6.14. Self-appraising, interpreting scores, and practising agency: Linking Yuki’s self-perceptions, test perceptions and actions

<table>
<thead>
<tr>
<th>Self-appraising</th>
<th>Interpreting scores</th>
<th>Practising agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative:</td>
<td></td>
<td>Tactical:</td>
</tr>
<tr>
<td>my English wasn’t really good, → that’s why I couldn’t pass → so um I decided to go to uni [Extract 6.7.1]</td>
<td>[Extract 6.7.1]</td>
<td>[Extract 6.7.2]</td>
</tr>
<tr>
<td>so I couldn’t speak properly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(unable to achieve scores for university entry)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative:</td>
<td></td>
<td>Tactical:</td>
</tr>
<tr>
<td>My writing is not really good +</td>
<td>yeah but writing just reduced my overall score</td>
<td>I came back to Melbourne as a working holiday [Extract 6.7.3]</td>
</tr>
<tr>
<td>Positive:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The implicit link between immersion and achievement, which Yuki perceives in relation to test score success as well as language skill development, is explicitly illustrated in Extracts 6.7.5 and 6.7.6, below. These extracts further demonstrate the theme ‘practising agency’, and highlight two key forms of strategic agency that Yuki practises over her initial periods in Australia: (i) repeating the test (immersion in the text context); and (ii) surrounding herself with Australian English speakers (immersion in English speaking contexts).

As already mentioned, Yuki made ten test attempts during her second year in Australia. Her explanation for taking the test so many times is provided in Extract 6.7.5, below.

**Extract 6.7.5**

“I took so many times just, just, because I’m…I was studying um IELTS practice, test practice, but I thought if I just have a little bit of break, I just don’t remember or I’m not really get, I’m not really familiar with the format of the exam, so I just thought it’s a good idea to take continuously, if, until I can reach the certain level I need, so yeah” [Interview 1, 126-129]

As shown in Extract 6.7.5, Yuki regarded the action of repeat test taking as a valuable form of test preparation, through which she increased her familiarity with the test tasks. Repeat test taking thus represents practising strategic agency in Yuki’s case; she gaining practice and test familiarity through immersing herself in the test context and views this as a means of eventually realising her score goal.

Despite repeating the test ten times while in Australia on a working holiday visa, Yuki was still unable to meet the score requirements for university entry and, as mentioned earlier, was forced to again return to Japan at the end of this visa period. As noted, she managed to achieve the scores she needed in Japan, and returned to Australia in early 2011 to begin her nursing studies.

When she returned to Australia, Yuki made every effort to surround herself with Australians, to gain as much practice using English in everyday situations as possible. She
decided, for example, to live in a shared house with Australians rather than in international student accommodation. As shown in Extract 6.7.6, below, she relates the opportunities for practice this form of immersion provides directly to her goal of becoming a skilled English user in her professional domain, which she sees as necessary if she is to succeed in the practical component of her university studies:

Extract 6.7.6

“I live in a share house... I live with Aussie people, so I just speak English there, yeah. I try to speak English cos when I'm doing a clinical placement if you don’t speak English properly you can’t communicate with um like a buddy nurse, or you know educator, or even patients. That's really hard and then some students just um fail because of that so I just decided to speak English as much as possible” [Interview 1, 240-245]

The two key forms of strategic agency Yuki practises in an attempt to realise her language learning and migration goals (Extracts 6.7.5 and 6.7.6), which are in addition to the tactical agency that characterises her complex migration pathway (Extracts 6.7.2 and 6.7.3), remain central to her approach to meeting the score requirements for permanent residency, described in the next subsection ‘Experiencing repeat test encounters’.

6.7.3 Experiencing repeat test encounters

Yuki, as noted in the introduction, had made three English test attempts for the purpose of gaining permanent residency in 2012 before the first interview took place. The first three attempts were followed by three more between the first and second interviews, and three more again between the second and third interviews. Yuki was taking both IELTS and the OET, and her score outcomes across the 9 attempts she made as part of her most recent attempt to gain permanent residency are shown in Table 6.15 and Table 6.16, below.
Table 6.15. Yuki’s IELTS scores

<table>
<thead>
<tr>
<th></th>
<th>Jul 2012</th>
<th>Sep 2012</th>
<th>Jan 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>6.5</td>
<td>6.5</td>
<td>6</td>
</tr>
<tr>
<td>Writing</td>
<td>5.5</td>
<td>6</td>
<td>5.5</td>
</tr>
<tr>
<td>Listening</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Reading</td>
<td>6.5</td>
<td>8</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Table 6.16. Yuki’s OET scores

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Writing</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
</tr>
<tr>
<td>Listening</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Reading</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
</tbody>
</table>

Yuki’s perceptions, feelings, decisions and actions in relation to her test experiences over her nine encounters are presented in three groups of test attempts, below: ‘After 3 test attempts’; ‘After 6 test attempts’; and ‘After 9 test attempts’.

6.7.3.1 After 3 test attempts

After the first three test attempts, Yuki’s account of her experiences focussed predominately on the speaking test. She had previously achieved her target scores of IELTS 7 on speaking, but this time received a lower score.

Extract 6.7.7, below, illustrates a shift in the relationship between themes ‘self-appraising’ and ‘interpreting score outcomes’, compared to that which characterised Yuki’s accounts of previous test experiences, described in the previous subsection, whereby she linked score outcomes directly to her English skill level. By the time she received scores from her third test attempt, Yuki had spent almost four years in Australia. She had, by this time, many Australian friends and had accumulated experience and developed confidence communicating in English with people she identifies as Australian. As shown in Extract 6.7.7, Yuki thus begins to distinguish
between her English skills and her test performances, associating the latter but not necessarily the former with test score outcomes.

**Extract 6.7.7**

“I’m using English in daily life here, so speaking sometime you know…it depends on the topic of the exam though but um yeah speaking is not really huge problem for me, yeah [R: Why do you think you got a lower score on speaking this time?] I think because the topic was in um IT, and I’m not really IT person, I have no idea about what I can talk to the um examiner, so I was quite um not really active speaking so, yeah, so that’s why I think um, because our conversation doesn’t expand” [Interview 1, 194-201]

As illustrated in Extract 6.7.7, above, Yuki’s interpretations of score outcomes have become independent of self-appraisals of her English skills, which are positive (*speaking is not really huge problem for me*), and dependent instead upon self-appraisals of her test performance (*I was quite um not really active speaking*), the quality of which she sees as contingent on her topic knowledge rather than her English level (*I’m not really IT person... that’s why I think um, because our conversation doesn’t expand*).

This perception that score outcomes depend on test topic also shapes Yuki’s perceptions of the score requirements. Extracts 6.7.8 to 6.7.10, below, demonstrate that Yuki holds implicitly conflicting views on the appropriateness of the link between topic knowledge and score outcomes.

**Extract 6.7.8**

Extract 6.7.8, below, illustrated the theme ‘understanding target scores’. As shown in Extract 6.7.8, while Yuki supports the notion that nurses need to be able to communicate in English, she contests the appropriateness of requiring test scores of IELTS 7. She further challenges the need to achieve such scores in all four skills in one sitting, on the basis that scores depend, at least in part, on factors beyond English proficiency, such as topic knowledge and also the physical condition of test takers on the day of the test.
Extract 6.7.8

“I think, yeah, a certain level should be um like a borderline to be a nurse, but it’s a bit harder for, like 7 is quite hard, I mean IELTS 7 is quite hard and then, you know, one time, you have to, like four sections you have to pass in one time, so that is, that is the a bit hard part... Maybe just like three times, four times you can take, and then if you can pass all sections, eventually, that’s good enough I think... Yeah, cos it depends on the topic, and yeah, like the same as the OET, it really depends on the topic and it depends on your condition, like your physical condition as well....” [Interview 1, 379-387]

In Extract 6.7.8 above, as noted, Yuki questions the English score requirements for nursing registration, situating test topics as potential obstacles to test success. By contrast, in Extract 6.7.9, below, Yuki describes the different topics she encounters on IELTS and the OET in positive terms.

Extract 6.7.9

“I need a bit of ah practice a bit more, yeah, especially speaking, because I’m somet-, I still stuck. I’ve been speaking English about nearly 4 years now, over, I still stuck what I want to say to them or some unfamiliar situation, I can’t explain it, so yeah, I need it, I need practice.... ...IELTS is quite helping I think because it’s more general topics so you can talk-, you can- how can I say, get used to talking Engl-speaking English in general situation, not like a medical situation... ...OET speaking is maybe helpful, because um it’s more like a real situation in a hospital so it’s just explain to something to the patient or um persuading him or her to be doing something in a situation so that’s helpful” [Interview 1, 391-403]

Extract 6.7.9, above, illustrates a relationship between themes ‘self-appraising’ and ‘understanding target scores’, which highlights the ambivalent stance Yuki takes towards the score requirements. Although she positively self-appraises her ability to speak English (Extract 6.7.7, further above), in Extract 6.7.9, she provides a negative self-
appraisal of her capacity to handle unfamiliar topics in everyday and professional contexts \textit{(I still stuck what I want to say to them or some unfamiliar situation, I can’t explain it)} and identifies a need for further practice on this basis \textit{(I need practice)}. Yuki describes the different test topics she encounters on the OET and IELTS as providing useful practice for communicating in social and work settings, respectively, which she associates with improving her capacity to effectively engage in interactions in diverse situations. This suggests that her understanding of target scores, as partly dependent on topic \textit{(Extract 6.7.8, further above)}, is consistent with her perception of the communicative demands she faces outside the test.

Conversely, Yuki identifies using English in her studies and daily life as valuable forms of test practice and preparation, as shown in Extract 6.7.10, below, which further evidences a perception that the test construct was well aligned to the actual communication demands she faces.

\textbf{Extract 6.7.10}

\begin{quote}
\textit{“I still use English for you know assignment, or school, or something... ...even just the normal daily usage of English just, is still part of practice of my English test”} [Interview 1, 454-457]
\end{quote}

Taken together, Extracts 6.7.7 to 6.7.10, above, illustrate mixed self-appraisals, which are mirrored by an ambivalent attitude towards the score requirements (see Table 6.17, below).
Table 6.17. Summary of relationships between themes after Yuki’s first three test attempts

<table>
<thead>
<tr>
<th>English skills</th>
<th>Test performance</th>
<th>Interpreting score outcomes</th>
<th>Understanding target scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive:</td>
<td>Self-appraising</td>
<td>Negative:</td>
<td>Test construct is</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>inappropriate/unfair:</td>
</tr>
<tr>
<td></td>
<td>speaking is not</td>
<td></td>
<td>IELTS 7 is quite hard,</td>
</tr>
<tr>
<td></td>
<td>really huge problem for me</td>
<td></td>
<td>four sections you have</td>
</tr>
<tr>
<td></td>
<td>I was quite um not really active speaking</td>
<td></td>
<td>to pass in one time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I still stuck what I want to say to them or some unfamiliar situation, I can’t explain it</td>
<td>Test construct is</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>appropriate/fair:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IELTS is quite helping,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>you can get used to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>speaking English in</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>general situation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As summarised in Table 6.17, on the one hand, Yuki expresses confidence in her speaking ability and attributes poor score outcomes to a lack of topic knowledge, rather than English language skills, which she links to an implicit claim of unfairness concerning the need to achieve target scores in all four skills in one sitting. At the same time, she provides a negative self-appraisal of her ability to handle unfamiliar topics in interactions, and identifies the diverse topics she encounters on IELTS and the OET as helpful and relevant to achieving her English learning goals outside of the test, in her personal, academic, and professional domains. She further identifies using English outside of the test as a valuable form of test preparation. This reciprocal relationship between developing English skills and improving test performances thus signals that she still perceives the test to be providing, at least to some degree, a meaningful measure of her English abilities.
6.7.3.2 After 6 test attempts

By the second interview, Yuki had made two attempts at IELTS and four OET attempts. On her second IELTS attempt, Yuki achieved above IELTS 7 on the receptive skills but failed to reach her target score in writing and again in speaking, despite feeling confident with the speaking topic on this occasion. This unexpected speaking outcome confounded Yuki and she requested that the test be remarked, but her score remained the same. On the writing subtest, by contrast, she was not perplexed by her lack of success since over her multiple IELTS encounters, she had never scored above IELTS 6.

In Extract 6.7.11, below, Yuki is explaining her writing score outcome. As described below, the extract illustrates an interrelationship between themes ‘self-appraising’, ‘interpreting score outcomes’ and ‘understanding target scores’ that is consistent with her perception of the test and scores as meaningful measures of her English skills. Yuki attributes her score outcomes to a perceived deficit in her academic writing skills. She maintains this negative self-appraisal despite experiencing no difficulties writing in the context of her university studies. While noting the apparent contradiction, Yuki does not question the appropriateness of her test score outcomes. Instead, she adjusts her understanding of target scores; she assumes that the standard of writing associated with IELTS 7 is higher than that which she is expected to produce in her university assignments.

Extract 6.7.11

“I’m not really, I’m not really good at academic writing, but I normally write academic writing for my assignments for uni so I’m not really sure about the... Because I don’t have any problem in the uni assignment, then IELTS, I think it’s more strict, I guess, yeah that’s why I can’t get much score, yeah”

[Interview 2, 65-68]

Extracts 6.7.12 to 6.7.14, below, relate to Yuki’s OET experiences. As shown in Extract 6.7.12, on three of her four OET encounters thus far, Yuki has achieved her target scores
(OET B) on three of the four skills, failing to succeed on a different skill each time (first writing, then reading, then listening).

**Extract 6.7.12**

"I have to have B for all subject. But every time I couldn’t, I couldn’t pass like a different subject... ...it’s just one subject I can’t pass... ...and every time different subject” [Interview 2, 88-95]  

Extract 6.7.13 below, illustrates themes ‘**self-appraising**’, ‘**responding affectively**’, and ‘**practising agency**’ and indicates a decoupling of the relationship between ‘self-appraising’ and ‘interpreting score outcomes’ that characterised Yuki’s accounts so far of her IELTS experiences. In Extract 6.7.13, Yuki is distinguishing between her last OET score outcomes, in which she met the requirements in all skills except listening, and her self-perceptions of her ability to use this skill in her professional domain.

**Extract 6.7.13**

"I understand more medical terminology and I understand more patient situation because of my listening skills got better I think, and um I can ask to the doctor or other um medical people rather than just asking to my buddy nurse to ask somebody so I can just directly ask to the doctor. So I think I feel more confident” [Interview 2, 582-585]  

Yuki’s listening test outcome conflicts with her self-appraisals, which have become increasingly positive throughout her university studies and, in particular, throughout her clinical placements. As shown in Extract 6.7.13, comparing her most recent clinical placement experience with earlier ones, Yuki is able to identify a positive shift in her ability to deal with the communicative demands she faces in the context of her role as nurse (*I understand more medical terminology and I understand more patient situation*), which she attributes to improvements in her listening skills (*because of my listening skills got better I think*). She associates this positive self-appraisal of her listening skills with positive affective responses (*I feel more confident*) and an increasing capacity to practise
agency in her professional role (I can ask to the doctor or other um medical people rather than just asking to my buddy nurse).

As shown in Extract 6.7.14, below, Yuki’s sense of confidence and positive self-appraising is reinforced by positive feedback she receives from her nursing mentors and colleagues on her English abilities during her clinical placements. This feedback encourages Yuki to question the need for test scores as evidence of her English skills, given that she and her colleagues feel that she is already demonstrating her capacity to communicate effectively in her role as a nurse, and this questioning prompts feelings of frustration.

**Extract 6.7.14**

“I was really frustrated, because when I was talking to any nurses, they said I you know they don’t what’s the point of the English test for us, because it sometime student from other country can’t cannot speak English properly, so they have some problem about the communication there, in the clinical placement, but she um yeah, most of my buddy nurse told me your English is not too bad, so yeah, they just said oh why, I don’t know why you have to take English test again. And it’s, it gave me confidence, but it gave me frustration as well” [Interview 2, 315 - 321]

**6.7.3.3 After 9 test attempts**

After a fifth failed attempt at the OET at the end of 2012 (her 7th test attempt out of 9) in which Yuki again met the requirements on three of the four skills, this time failing to reach target scores on writing, she decided to re-attempt IELTS. While her IELTS score outcomes had also, on occasion, been inconsistent with her self-appraisals, particularly in relation to speaking, in general Yuki perceived IELTS scores to be more closely aligned with her English skills than the OET scores she achieved, as indicated above by Extracts 6.7.11 to 6.7.14.

At the beginning of 2013, Yuki was approaching the end of her temporary student visa period, and so she felt an increasing sense of urgency associated with meeting the score
requirements for permanent residency. Desperate to improve on past IELTS performances, and consistent with her view that IELTS scores were meaningfully linked to her writing ability in English, she modified her approach to the test, as shown in Extract 6.7.15, below. Extract 6.7.15 illustrates the theme ‘practising agency’, which is implicitly linked to Yuki’s understandings of score meanings, demonstrated previously. Instead of focusing on rehearsing performances using sample test questions, Yuki adopts a new strategic approach to test preparation; she sets about studying grammar and vocabulary to improve her English knowledge and skills.

Extract 6.7.15

“I just thought that English test in January is the, was the last opportunity for me so I just study really hard. I study um, normally I just practice with the um previous exam test or those kind of exam materials but this time I just do, do like more like basic stuff like grammar or you know vocabulary or… Not really related to the um test but it’s more like a general thing so yeah so I did more basic stuff” [Interview 3, 40-51]

Despite these efforts, Yuki’s IELTS score outcomes did not improve on this final attempt, as shown in Extract 6.7.16, below. Although she achieved target scores on reading and listening, she scored below 6 for writing and below 7 for speaking.

Extract 6.7.16

“Even though I studied like really basic stuff I couldn’t improve my score of the IELTS, so I just felt really disappointed because I studied really hard … …because my weakest point is writing so that’s why I studied um the grammar or vocabulary quite intensively but um yeah… ….I just cannot, how can I, I cannot like achieve higher score than 6 even though I study really hard so I feel really disappointed and then just you know like a bit frustrated and just bit kind of burnt out” [Interview 3, 83-108]

As shown in Extract 6.7.16, above, Yuki describes her test preparation as a form of remedial action, directed at the area of weakness she identified through the process of
self-appraising (my weakest point is writing) and intended to bring about improvements (so that’s why I studied um the grammar or vocabulary quite intensively). Yuki was thus at a loss concerning how to interpret her score outcomes, which she perceived to be incommensurate with the effort she had put into preparation (I just cannot, how can I, I cannot like achieve higher score than 6 even though I study really hard). These perceptions, which left her feeling unsure about how to proceed, triggered negative affective responses (I just felt really disappointed; a bit frustrated; kind of burnt out).

In a last-ditch effort to meet the requirements in time to apply to extend her stay in Australia, Yuki again attempted the OET, in the same month as her IELTS attempt, January 2013. On this occasion, she achieved her target scores or above in all four skills, which meant she was able to register as a nurse and apply for permanent residency. Asked if her score success changed how she felt about her English skill level, she responded in the negative, but described how finally realising her goals after such a prolonged effort impacted how she feels about herself and about the language (see Extract 6.7.17, below).

**Extract 6.7.17**

“I think um the overall it is not really different… but overall I think I feel more relaxed so you know I’m more like, actually how can I say, actually I’m more like um motivated to study English or you know learn more vocabulary or more like actively talking with other people. I think because I have a little confidence because of that achievement, yeah, because you know I haven’t passed English test for ages so it’s just give my confidence a little bit lower, like you know a little bit less. Even though I studied really hard you know I couldn’t achieve for a while and yeah I just, my you know lost my confidence for ages so…Yeah this time I passed the English test and then it’s just a little bit you know it’s just a little bit lift up my confidence” [Interview 3, 782-805]

As argued earlier, Yuki, by this time, had begun to question the meaningfulness of the test requirements and her score outcomes, at least on the OET if not on IELTS. After the first six test attempts, she began to disassociate the processes of self-appraising and
interpreting score outcomes in her accounts of her test experiences. As shown in Extract 6.7.17, above, she maintains this disassociation at the end of her test taking journey (I think um the overall it [English skill level] is not really different). Nonetheless, Yuki situates her ultimate success on the OET as a valued achievement; one which she associates with positive affective responses, feelings of increased motivation to improve her English knowledge and also with increased feelings of confidence in herself and her English abilities.

6.7.4 Summary of Yuki’s experiences

As demonstrated above, for most of her test-taking journey, including before her most recent efforts to gain permanent residency, Yuki believed that test scores provided meaningful and accurate information about her English skill level. As a result, her process of self-appraising was reciprocally related to interpreting score outcomes. These self- and test-perceptions influenced the type of agency Yuki practised throughout her test experiences, and reinforced her belief that immersion and practice were central to achieving her English language learning and test score goals. In Yuki’s case, the test context was embedded within her real world experiences of learning and using English, with repeat test encounters providing what she viewed as a valuable form of skill practice and as an objective source of feedback on her language development. Even as she developed confidence and a greater sense of agency in her professional context, she did not question the link between her English skill level and test score outcomes on IELTS, although, as noted, she questioned the consistency and relevance of the OET in light of emerging conflicts between her real world experiences and test score outcomes. Unexpected score outcomes prompted negative affective responses, but did not trigger any shift in Yuki’s overall strategic plan or shift in her understanding of target scores, as despite her doubts and frustrations, she accepted the authority of the test.

Figure 6.4, below, summarises the relationships between themes that characterise Yuki’s experiences.
6.8 Summary of the Four Cases

Findings presented in this chapter highlight the significance of individual subjectivities and time as variables in shaping the impacts that emerge from the use of English language test scores as selection criteria within Australia’s skilled migration policy. As demonstrated, each of the four individuals situated the test and score requirements in relation to their self-perceptions and their perceptions of their experiences learning and using English in Australia in different ways. These perceptions, shaped by individual histories and identities, influenced their processes of self-appraising, their interpretations of score outcomes and their understandings of target scores, which in turn shaped the types of agency they practised and their affective responses to their experiences.
Over time, each individual engaged in iterative processes of comparing their new and previous test and non-test experiences of using English to evaluate and revise their perceptions, decisions, and actions in an effort to establish certainty and a sense of control, with the ultimate aim of realising their score and migration goals. In so doing, each individual played a role in negotiating test consequences, shaping the different effects of the test on their own lives over time. Throughout their accounts, each individual raised questions about the appropriateness of the score requirements, the relevance of the tasks to the English demands living and working in Australia, and the consistency of the scoring procedures. These issues, in addition to the issues raised by the consequences of test use in each case as manifest in individuals’ affective and actions responses to the test and the score requirements in this policy context, are discussed in chapter 8.

In chapter 7, findings from interviews conducted with teachers and test preparation course directors are presented. As will be further discussed in chapter 8, the accounts provided by these teacher participants support many of the claims made by test takers in relation to the validity and fairness issues raised by test use in the context of Australia's skilled migration policy.
CHAPTER 7: TEACHER PERSPECTIVES

7.1 Introduction

As demonstrated in the previous chapter, each of the four migrants in this study sought out information and/or feedback from experts as part of their efforts to realize their score goals, either by taking a test preparation course or by employing IELTS examiners as private tutors. This chapter presents findings from interviews conducted with seven English language teachers involved in delivering test preparation courses to students needing IELTS scores of 7 and 8 to meet the minimum eligibility requirements to apply for permanent residency. As noted in chapter 5, three of these teachers were also IELTS examiners. The seven teachers who are the focus of this chapter represent what the migrants viewed as the expert perspective in relation to the English test score criteria within skilled migrant selection processes. As noted in chapter 5, interviews with seven teachers were designed to capture an alternative perspective on the use of language tests in the process of migrant selection, as a means of situating the experiences of the four individuals involved in the current study in relation to the experiences of a broader group of test takers seeking scores of IELTS 7 or 8 to gain permanent residency as skilled migrants, and to provide a means of interrogating the plausibility of migrant accounts as well as the credibility of the researcher’s interpretations of these accounts.

As was the case with the migrant interview data, analysis of teacher interview data focused on identifying themes and relationships between themes that characterized participants’ accounts of their perceptions, decisions and actions in relation to the test. The teachers’ accounts focused predominately on their experiences of engaging with test takers, rather than directly with the test itself, and on the impact of test use within skilled migration policy on their teaching practice and on how they perceived their roles as teachers. Five key themes were derived from the teacher interview dataset: Appraising test takers’ English use; understanding score meanings; characterising the test taker perspective; negotiating score meanings; and losing agency in the classroom. These themes parallel those derived from the migrant dataset, while taking account of the different perspective offered by teachers.
In this chapter, the seven teachers are first briefly introduced (pseudonyms are used throughout). Following this, findings are organised around key themes. For each theme, a brief definition is first provided, which is then elaborated in relation to teachers’ accounts. Within individual themes, relationships to other themes are identified and exemplified iteratively. To conclude this chapter, a summary of findings is provided.

7.2 The seven English teachers

As noted in Chapter 5 (section 3.1.2), two of the seven English teacher participants were directors of studies at the time interviews were conducted, responsible for overseeing the English language and test preparation courses offered at their institutions rather than classroom teaching. The two directors of studies, identified in this chapter as Claire and Jane, were both experienced IELTS examiners. Claire represented an institution that had introduced IELTS preparation courses aimed at those needing IELTS 7 and 8 in response to new demand associated with the introduction of these score criteria into skilled migration policy. At Jane’s institution, by contrast, the extra demand from would-be skilled migrants had been absorbed into their existing IELTS preparation courses, which were not targeting particular scored bands. Of the five remaining teachers, Paul, Michelle (also an IELTS examiner), Jenny, and Ruth worked at the same institution as Claire, and Kate worked at the same institution as Jane.
7.3 Key themes

7.3.1 Appraising test takers’ English abilities

This theme refers to teachers’ accounts of their perceptions and evaluations of migrant test takers’ English skills, including their abilities to communicate in English in general and in the contexts of their employment, and their capacity to realise their IELTS score goals.

Teachers’ appraisals of migrants taking IELTS for permanent residency were consistent with positive self-appraisals provided by Ana, Erfan, Sharon, and to some extent, Yuki, in relation to their capacity to meet the English demands of their professional domains (see Chapter 6). As shown in Table 7.1, below, four of the seven teachers explicitly expressed broadly positive appraisals of the abilities of test takers for permanent residency to communicate in English. It is important to note that other teachers (Paul and Michelle) did not make negative appraisals in relation to the abilities of these test takers to use English in real world contexts. Rather, their accounts focussed primarily on the test context and foregrounded their perceptions of the test construct, and so are shown in relation to the next theme ‘understanding score meanings’, further below.

Table 7.1. Positive appraisals of migrants’ abilities to communicate in English

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Appraisal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claire:</td>
<td>“these people are very functional, they’re working, and you’ve got no problem whatsoever communicating with them” (50-51)</td>
</tr>
<tr>
<td>Kate:</td>
<td>“[They] have really high levels of language ability” (56)</td>
</tr>
<tr>
<td>Ruth:</td>
<td>“Like I said, they function quite adequately and they certainly know, you know, you have no problem communicating with them” (94-95)</td>
</tr>
<tr>
<td>Jenny:</td>
<td>“A lot of them can function very well in English” (40)</td>
</tr>
<tr>
<td>Jane:</td>
<td>“Very fluent, very capable communicators” (206)</td>
</tr>
</tbody>
</table>
Teachers, similar to the migrant participants in this study, also made a distinction between their appraisals of this test taker group’s ability to use English in the real world and their capacity to achieve test scores of IELTS 7 and 8. In contrast to the positive appraisals teachers made in relation to their ability to use English, shown in Table 7.1, above, the consensus was that score of IELTS 7 and 8 were out of reach for many would-be permanent residents, despite their high levels of English. For example, in Extract 7.1, below, Ruth highlights typical language behaviours she associates with adult migrants from non-English speaking backgrounds in her classes, which she views as unproblematic in terms of actual communication but as presenting insurmountable obstacles to achieving IELTS 7 and 8.

Extract 7.1

“They make those ingrained typical mistakes of their, you know pertinent to their culture...language, first language. In everything, in their speech of course, and they’re all ingrained problems that are very hard to eradicate because they’ve plateaued and so you can point them out to them and they will understand them but then, it’s just automatic. [R: Does it interfere with meaning?] Absolutely not, no, but I’m sure it affects their score, very much so” (Ruth, 97-105)

A shown in Extract 7.1, above, Ruth situates these non-standard language behaviours as “ingrained” and “automatic”; she feels assimilation to the norms expected on the test is unachievable for adult migrants from non-English speaking backgrounds. Kate and Jane, as shown below in Extracts 7.2 and 7.3, respectively, also felt test takers from different backgrounds were disadvantaged by score criteria, due to the different norms underlying their language use.
Extract 7.2

“Even if they have strong language skills I think the biggest problems, aside from the structure and the expectations of the examiners, are cultural. For example, writing a 250 word essay in 40 minutes is quite difficult if you’re from a culture that isn’t very direct, so then they have a tendency to write a lot more than they need to, which wastes a lot of their time... ...it’s hard for adults to think differently than the way they always have” (Kate, 56-62)

In Extract 7.2, above, Kate suggests that the cultural norms that influence some test takers’ language behaviours are incompatible with test conditions and scoring criteria, which privilege direct and concise expression. On this basis, despite their “strong language skills”, she provides a negative appraisal of their capacity to achieve scores of IELTS 8, given the difficulty involved in modifying culture-based writing behaviours (it’s hard for adults to think differently than the way they always have).

In Extract 7.3, below, while Jane positively appraises the ability of test takers from the Indian sub-continent to communicate in English (even though they might be very fluent, very capable communicators), she negatively appraises their capacity to reach scores of IELTS 8 (they don’t get those high scores they want), which she attributes specifically to IELTS scoring criteria (you have to mark to the criteria).

Extract 7.3

“You get people from the sub-continent who maybe have spoken English since they were very young but their pronunciation and their stress and their intonation, maybe their choice of lexis. It’s all those things...even though they might be very fluent, very capable communicators but when you look at the criteria, you have to mark to the criteria then they don’t get those high scores they want” (Jane, 202-207)

Extract 7.3, above, illustrates Jane’s perception that test score criteria are based on particular norms of pronunciation and prosody that privilege certain varieties of English
(and speakers) and disadvantage others. Thus, Extract 7.3 explicitly highlights a relationship between the theme ‘Appraising test takers’ English abilities’ and the next theme to be described, ‘Understanding score meanings’.

This relationship between themes was also implicit in Extracts 7.1 and 7.2, earlier. The conflict between positive appraisals of migrants’ English abilities (see Table 7.1) and negative appraisals of their capacities to reach scores of IELTS 7 and 8, exemplified by Ruth’s and Kate’s comments, suggests a perception among these teachers that the test construct includes performance attributes that are not relevant to the communication demands faced by skilled migrants in their work and life in Australia. This perception, consistent with the views expressed by the four individual migrants in the previous chapter, is also expressed explicitly by some of the teachers in relation to their understandings of score meanings, as shown below.

### 7.3.2 Understanding score meanings

‘Understanding score meanings’ refers to teachers’ accounts of their perceptions of the construct underlying IELTS scores of 7 and 8, and of their perceptions of the appropriateness of the use of such scores for skilled migrant selection purposes.

To reflect these two aspects of ‘understanding score meanings’, extracts are organised below under two subheadings: ‘Perceptions of the test construct’, which relates to teachers’ perceptions of score meanings in terms of English abilities and test performance attributes; and ‘Perceptions of test use’, which relates to teacher’s perceptions of the appropriateness of test use within skilled migration policy.
7.3.2.1 Perceptions of the test construct

Teachers described scores of IELTS 8 in terms of specific performance attributes, as shown in Table 7.2. To try to situate the English skill level associated with such scores, they compared their perceptions of the criteria they thought migrants would have difficulty fulfilling with their perceptions of the potential of ‘native’ speakers to meet the requirements, shown further below in Table 7.3.

Table 7.2. Describing IELTS scores in terms of performance attributes

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claire:</td>
<td>“you can’t get an 8 in coherence and cohesion in IELTS [writing] if you don’t have perfect paragraphing... things like prepositions, or that damned third person ‘s’, things like that, that really don’t matter um... It is impossible to get an 8 if you’re making errors there because you get that high frequency” (138-145)</td>
</tr>
<tr>
<td>Michelle:</td>
<td>“you need to know you really do need to focus on the target language. For example, me speaking to you now I would be losing marks, yeah, for not completing complete sentences, pausing, repetition, self-correction, the band score starts to go down, yeah...It’s not at all reflective of everyday conversation” (80-85)</td>
</tr>
<tr>
<td>Paul:</td>
<td>“for getting the higher band [in writing] it does seem to be more of a stylistic requirement and an argumentative, self-reflecting, the kind of skills which can only come from a long period of reading and writing and being in perhaps an academic environment, which many of the candidates are not going to have had” (183-185)</td>
</tr>
<tr>
<td>Ruth:</td>
<td>There are lots of features that are required of them which have no bearing on language proficiency” (111)</td>
</tr>
</tbody>
</table>

The examples provided in Table 7.2, above, illustrate different aspects of the test performance teachers focused on when describing their understandings of the scores associated with permanent residency, IELTS 8 in particular. As shown, Claire associates scores of IELTS 8 in writing with “perfect paragraphing” and a low frequency of errors, whereas Paul aligns such scores with a particular style of argument. Michelle viewed IELTS 8 in speaking as reflecting a level of accuracy in production that is “not at all reflective of everyday conversation”, while Marissa felt scores of IELTS 8 were associated
with skills that “*have no bearing on language proficiency*”. Also in terms of the relevance of the test construct to the English communication demands faced by migrants, Claire described the IELTS 8 criteria that migrant test takers generally failed to fulfil as “*things that really don’t matter*” and Paul associated the development of the skills needed to achieve such scores with experiences that “*many of the candidates are not going to have had*”.

Consistent with their questioning of the relevance of the test construct to the English language demands faced by migrants general, teachers viewed scores of IELTS 8 as unattainable not only for migrants but also for many existing ‘native speakers’ of English (Table 7.3). This indicates they are interpreting the score requirements to represent an English skill level over and above that possessed by existing Australian citizens and residents.

**Table 7.3. Comparing IELTS 8 to English ‘native’ speaker ability level**

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>“I don’t know what the official thing is, you know there would be native speakers who wouldn’t necessarily get an 8” (46-47)</td>
</tr>
<tr>
<td>Claire</td>
<td>“Many of our native speakers wouldn’t do very well in the test” (138-139)</td>
</tr>
<tr>
<td>Michelle</td>
<td>“It’s not an easy test. To be honest with you, if you asked the guy next door or anyone off the street, most people wouldn’t do very well,” (61-62)</td>
</tr>
<tr>
<td>Paul</td>
<td>“I think for many students who would not be particularly strong in writing, native speakers, then 8 is, would be a difficult task to achieve” (166-167)</td>
</tr>
<tr>
<td>Ruth</td>
<td>“I think the working against the clock feature would be very difficult for any even native speaker” (108-109)</td>
</tr>
</tbody>
</table>

### 7.3.2.2 Perceptions of test use

As expected, a perception that the test construct included performance attributes that are unachievable and also peripheral to successful communication on the part of
migrants was associated with negative perceptions of test use in skilled migration policy. For example, Jane and Claire, the two directors of studies, expressed strong criticism of the use of IELTS in the process of determining rights to permanent residency. As demonstrated earlier within the theme ‘Appraising test taker’s English abilities’, they each associated the test construct with an exclusionary function. Jane identified norms of pronunciation and prosody that underlie score meanings as an obstacle to score success for English speakers from the sub-continent, while Claire saw the need to produce error free language and perfect paragraphing as limiting the chances of success for most test takers from non-English speaking backgrounds. In Extracts 7.4 and 7.5, below, these perceptions of exclusion are associated with claims of unfairness by both teachers.

Extract 7.4 demonstrates that while Jane perceives IELTS to be a good test in terms of its technical qualities, she questions the gatekeeping function served by scores in skilled migration policy. She views the score requirements as unnecessarily high (there are lots of jobs where people don’t need that), which, at best, signals to her that government expectations about what migrants should be able to achieve are unrealistic, or worse, that test use represents a deliberate effort to exclude particular groups (are we trying to screen out a whole lot of people?). These perceptions prompt Jane to question her own role in the testing process, as shown below.

Extract 7.4

“I just started to question whether I wanted to be part of that system that seems to me unfair because obviously in certain jobs you want to have people with that language proficiency which equates to a 7 or an 8 but there are lots of jobs where people don’t need that... ...Do we actually...are we trying to screen out a whole lot of people? I don’t know, so...” (Jane, 21-32)

In Extract 7.5, below, Claire’s criticism of the score requirements is focused primarily on the writing section of the test, which she sees as representing the most significant obstacle to test takers. As shown in Extract 7.5, while Claire seems to accept the legitimacy of government efforts to ensure incoming migrants possess adequate English
skills, she also feels that the score requirements are unnecessarily restrictive (it is unrealistic... I really feel like the hoops are unnecessary). According to her perceptions, the criteria associated with high IELTS scores, particularly for writing, overemphasise accuracy to the extent of being discriminatory (it also really prioritises or favours people from an English-speaking background for immigration purposes).

Extract 7.5

“I just think it is unrealistic and unfair... ...I think the test itself I think the construct particularly in the writing is, should be rethought, because it’s being used for immigration purposes and I think it’s holding out, it’s making, it’s, it’s costing people a lot of money to try and jump through these hoops and I really feel like the hoops are unnecessary... ***

...I understand a government wanting to be sure that its citizens are going to be sort of taking on a functional level of the language but you know it also really prioritises or favours people from an English-speaking background for immigration purposes so I... it’s really unfair in that way” (Claire, 143-148; 173-176)

Similarly, Kate, who perceived some test takers to be hindered by cultural factors that conflicted with the test construct (see Extract 7.2, earlier), supported the need for English proficiency criteria but objected to what she perceived to be discriminatory effects of the IELTS tasks and scoring criteria, as shown in Extract 7.6, below.

Extract 7.6

“I don’t have a problem I suppose the language requirement, I have a problem with the cultural bias that I think exists in the way we assess the requirement” (Kate, 98-99)

Other teachers, despite perceiving scores of IELTS 8 to be above the level of English needed to enable effective communication in English, were less critical of the use of IELTS scores in migrant selection processes. Michelle, for example, remained
ambivalent, taking the view that in any context there will be limitations associated with test use but that on the whole, whether it be for immigration or academic purposes, IELTS yields useful information about test takers (Extract 7.7, below).

**Extract 7.7**

“In terms of trying to ask is it a useful immigration tool, in a lot of ways yes, it is and in a lot of ways no it isn’t. Just as well for its own purpose, it could be good for academic purposes and a lot of ways it would fall short, it just due to the nature of the test, how much time people have to prepare, the actual test itself. For what it does cover and what it does examine it’s really quite a good measure of a lot of different things” (Michelle, 136-139)

Jenny, despite feeling that the migrant test takers she encountered were already able to function effectively in English, supported the use of IELTS within skilled migration policy, as she felt it created an impetus for migrants to prioritise reaching even higher English skill levels (Extract 7.8, below).

**Extract 7.8**

“I can see the merit in there being a language test in that it enforces that people have some kind of functional English, and I think that’s a good thing... The test is quite high level so as soon as people had functional English they would be absorbed by the demands of their life and I doubt very much if they would they would, they would pursue a higher level of ability in English, so...” (Jenny, 124-137)

As shown in Extract 7.9, below, Paul privileges the perspective of existing Australian citizens and residents in his evaluation of test use, suggesting that the test is functioning to ensure their right to be able to communicate in English with migrants from different language backgrounds. Paul held a similar view to Jenny in that he also saw the test as functioning to motivate continued English development, as shown below.
Extract 7.9

“Those who are citizens or permanent residents already, must be able to expect a level of communication that doesn’t cause them any serious frustration. I think the test does play, it plays a key role. I mean I guess another way of posing the question is ‘if the test was removed and the requirements removed, would people strive to become as proficient in English as they would if the test weren’t there?’ and I have to say just on my experience, they probably wouldn’t” (Paul, 114-122)

In Extracts 7.8 and 7.9, above, Jenny and Paul, respectively, make assumptions about the migrant perspective on the value of developing English skills in order to evaluate the appropriateness of test use for migration purposes. Jenny claims that “as soon as people had functional English they would be absorbed by the demands of their life and I doubt very much if they would they would, they would pursue a higher level of ability in English”, and Paul, as already noted, makes a similar claim that in the absence of the test and score requirements, migrants would lack motivation to continue developing their English skills. These extracts thus illustrate a relationship between ‘Understanding score meanings’ and the next theme, described below, ‘Characterising the test taker perspective’.

7.3.3 Characterising the test taker perspective

In considering the effect of the score requirements for permanent residency on their teaching practice, teachers located test taker perceptions of and responses to the inclusion of IELTS 8 in migrant selection criteria as the source of a tension they noticed emerging in their test preparation classes. This tension is examined in relation to the next theme, ‘Negotiating score meanings’. The current theme, ‘Characterising the test taker perspective’ refers to teachers’ accounts of their assumptions or observations about migrant test takers’ perceptions, feelings and actions in relation to the test and score requirements.
As already noted, the inclusion of scores of IELTS 8 into skilled migrant selection criteria led to the creation of new test preparation courses, including at the English college where Claire worked as Director of Studies. Claire noted that the introduction of a course explicitly aimed at the scores needed for permanent residency led to a significant surge in student demand:

**Extract 7.10**

“We ran IELTS classes called IELTS Saturday or IELTS evening and we had two or three students, as soon as we put IELTS 8 onto it, it was incredible, we were just flooded with people” (44-45).

Even though at Jane and Kate’s institution no new courses were introduced, they also felt that the changes to the English criteria in skilled migration policy had led to a significant shift in the composition of their classes. As Kate remarked:

**Extract 7.11**

“Before you’d have maybe one or two students who need a 7.5 or an 8 and now you have about half of the class” (9-10).

As shown in Table 7.4, below, teachers perceived the demand from this new group of test takers to be driven by a perception that scores of IELTS 8 were associated with a particular, idealised test performance, rather than a level of English proficiency *per se*. Teachers felt that test takers expected them, as test experts, to have access to the supposed formula for this ideal performance and to be able to deliver it in their classes.
Paul: “there is a tendency just to focus on ‘how can I pass the test?’; ‘what’s the magic formula?’” (124-125)

Kate: “they just immediately want you to pour into their head a perfect essay formula” (80-82)

Jenny: “I started to realize that they felt like there was a hurdle in front of them, like through the eye of a needle, they didn’t know what was the magic formula that made you small enough to pass through and get out to the other side” (90-93)

Ruth: “they’re looking for magical cures and tips that are going to give them that extra point” (11-14)

Teachers also remarked on the impact they perceived the test and score requirements to have on the test takers for permanent residency in their classes, as shown in Table 7.5, below. As shown, they describe the emotional effects of the test on this group in negative terms (depressed, sullen; desperation; stressed, frustration), and associate the pressure to achieve high scores with negative attitudes towards language learning (they actually become almost hostile to the idea).

Table 7.5. Characterising test taker affective responses

<table>
<thead>
<tr>
<th>Jenny:</th>
<th>“They are often very depressed; they seem sullen”’ (20-21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kate:</td>
<td>“there is a feeling, that sense of desperation” (5)</td>
</tr>
<tr>
<td>Paul:</td>
<td>“there is so much pressure on them to get this score in a high band that they actually become almost hostile to the idea of learning the language” (136-137)</td>
</tr>
<tr>
<td>Ruth:</td>
<td>“They’re very stressed... The frustration is extremely obvious” (10-14)</td>
</tr>
</tbody>
</table>

Paul’s observation, shown in Table 7.5, although seeming to conflict with his assertion that the score requirements are likely to function to motivate migrants to further develop their English skills (Extract 7.9, above), reflects a view that the test is playing a coercive role in creating a necessary impetus for further English skill development, regardless of test taker attitudes towards the need for of value of such development.
In terms of characterising test takers’ action responses to the score requirements, all teachers remarked on a phenomenon of repeat test taking amongst this group, as shown in Table 7.6, below.

**Table 7.6. Characterising test taker action**

<table>
<thead>
<tr>
<th>Name</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane</td>
<td>“I know sometimes people will come in and book 5 tests because they are just hedging their bets and sometimes they get the score they need” (211-212)</td>
</tr>
<tr>
<td>Kate</td>
<td>“There’s more people who need such high scores and I have more repeat test takers” (3-4)</td>
</tr>
<tr>
<td>Paul</td>
<td>“I have seen a lot of those students have taken the test in some cases several times if not more so you know, clearly there is a presence of the significant impact that the score will have on them” (21-23)</td>
</tr>
<tr>
<td>Ruth</td>
<td>“they will book three at a time. It’s a bit like a lottery” (11)</td>
</tr>
</tbody>
</table>

Extract 7.12, below, illustrates the relationship between ‘Characterising the test taker perspective’ and the next theme, described further below, ‘Negotiating score meanings’. In Extract 7.12, below, Ruth explains why she thinks test takers engage in repeat test taking. Her account sheds further light on how teachers viewed test takers’ perceptions of the test and understandings of IELTS 7 and 8, and on the reasons behind their resistance to teachers’ efforts to teach English skills in test preparation classes. The underlined part of the extract represents ‘Negotiating score meanings’, as will be further explained under the associated theme heading below.
Extract 7.12

“They don’t see the tests as being the same. I think they approach the test and think that some of them [the tests] are more difficult than others. They seem to think that it’s, that it’s the content that’s more relevant than the actual skill and it doesn’t matter how often I point out to them that really it’s, you know, their sophistication in their language delivery that matters and not their ideas they still come back with ‘but I knew nothing about that topic’” (Ruth, 18-22)

As illustrated in Extract 7.12, according to teachers, test takers perceive topic familiarity as a factor that impacts the quality of their test performance and thus represents a source of score variation. This prompts repeated test attempts, as test takers try to increase their chances of encountering favourable test materials. As shown in the previous chapter, this view of the role of topic is consistent with the perspective of the four migrant participants in this study. Ruth, however, makes a distinction between language skills and content that eliminates topic as a source of score variation. She sees the test as a measure of the former, not the latter, and so views the perspective adopted by test takers, rather than any lack of topic familiarity on their part, as the obstacle to score success.

As discussed further in relation to the next theme, ‘Negotiating score meanings’, the last part of Extract 7.12, underlined above, highlights a tension and conflict over score meanings between test takers and teachers, which has become salient to teachers since the introduction of new score criteria into skilled migrant selection procedures. This sense of conflicting perspectives prompts resistance from teachers, as explained further below, as they feel that a ‘magic formula’ view of scores undermines the purpose of the test and their roles as teachers.
7.3.4 Negotiating score meanings

The theme ‘negotiating score meanings’ refers to teachers’ accounts of their attempts to influence and shift test taker perceptions of the test and scores to better align with their perceptions of their roles as teachers.

As demonstrated previously, even though teachers questioned the appropriateness of the English proficiency level and the relevance of the performance attributes they associated with IELTS 8, they nonetheless viewed the test as a measure, albeit a limited and overly prescriptive one, of language learning outcomes. Accordingly, they perceived their primary role as teachers to be to help students increase and refine their English skills in order to at least improve their chances of improving their score outcomes.

Teachers thus resisted test taker interpretations of score meanings as reflecting a ‘magic formula’. As shown above, in Extract 7.12, this was because such a perception conflicts with teachers’ understandings of score meanings. Extract 7.13, below, illustrates that resistance also stemmed from a perception that a test taker belief in an ideal performance formula undermined the value of language learning that teachers sought to promote in their classrooms.

Extract 7.13

“You try to make the students realize that there is an overall learning... that there isn’t a list of things that make you a mythical IELTS 8 and that I can’t supply that knowledge even if I really want to and that for me there’s still merit in learning things along the way to round out a full and... a fuller understanding of speaking another language or being able to communicate in another language. So they are not interested in that, generally speaking”

(Jenny, 34-38)

As shown in Extract 7.13, above, Jenny contested test takers interpretations of IELTS 8 as divorced from English language learning. She felt unable to fulfil test takers’
expectation to be provided with an IELTS 8 formula (I can’t supply that knowledge even if I really want to), and this expectation conflicted with her commitment to the value of “overall learning”. Jenny attempts to resolve this conflict by re-establishing a link between scores and language learning, thereby encouraging students to resituate English skill development as part of their test preparation efforts. As shown in the last line of Extract 7.13, as was also the case for Ruth (Extract 7.12, above), she feels that her efforts are rejected (they are not interested in that, generally speaking).

7.3.5 Losing agency in the classroom

This theme refers to teachers’ accounts of losing authority and control of the classroom teaching and learning agenda, in the face of demands from test takers for permanent residency.

As demonstrated in relation to the previous theme, teachers perceived themselves to be locked in a struggle over score meanings with test takers. They felt that the view of scores as disconnected from English language learning outcomes, which they associated with migrant test takers, conflicted with their own understandings of scores and created demands on them as teachers that they were unable to fulfil. As shown in Extract 7.14, below, Jenny perceived that her inability to provide the formula that test takers sought led to a loss of authority in the classroom.

**Extract 7.14**

“A lot of them can function very well in English so all they want is the formula to pass the test, and they just want you to know what the formula is and if you look like you don’t know that they lose confidence in you” (Jenny, 40-41)

In Extract 7.15, below, Ruth suggests that the tension over score meanings emerges as a result of the pressure test takers feel in response to the score requirements for permanent residency, which she feels is likely to prevent them from focusing on the intrinsic value of language learning, regardless of whether or not they actually attribute value to it. Despite conflict she experiences in relation to her own perception that
language learning leads to improved score outcomes, she feels pressured to shift the focus of her classes to accommodate the demands of test takers, as shown below.

Extract 7.15

“They just want a magical cure, they’re not really here to ... they haven’t got, they don’t see languages are having intrinsic value, they might but the pressure is too great, so... if I start waffling on about the language and its intrinsic value I have to stop myself because they just want to... they come back with and what if, you know, they’re very focused on, they know exactly what they want to do” (Ruth, 87-91)

Extracts 7.14 and 7.15, above, thus illustrate that test takers for permanent residency, through their numbers and their sense of urgency, overcome teachers’ efforts to promote their own understandings of scores within teaching agendas. Instead, teachers find themselves faced with a series of paradoxical dilemmas – teaching on courses that promise scores they perceive to be unattainable, pressured to reveal a magic formula they believe doesn’t exist, and feeling a responsibility to deliver positive learning outcomes for a group of students who do not see themselves as learners and who do not, in this policy space, see any value in pursuing language learning outcomes.

7.4 Summary

As demonstrated in the current chapter, the perspectives offered by the seven teacher participants on the test and score requirements is consistent in many ways with those expressed by the migrant participants in this study, particularly in relation to how they appraise test takers’ English use in relation to the score requirements and how they perceive test use in the context of skilled migration policy, in terms of fairness. Teachers described migrant test takers as possessing a high level of English ability, but viewed scores of IELTS 8 as nonetheless unattainable. They saw the test as functioning to discriminate against migrants from non-English speaking backgrounds, particularly
those from dissimilar cultural backgrounds, and they questioned the relevance of the score requirements to the English demands faced by migrants in Australia. These concerns prompted some teachers to question the fairness of test use. Others, by contrast, saw the high score requirements as a legitimate means of encouraging migrants with already functional English skills to continue their language development.

Teachers’ perceptions of the test taker perspective were also consistent with the perspectives offered by the four migrant participants in this study, described in the previous chapter. They described test takers’ understanding of scores in terms of an elusive ‘magic’ formula. They observed feelings of desperation and frustration on the part of test takers in response to the test and score requirements, and they identified a phenomenon of repeat taking, which they associated with a perception that score outcomes were contingent on topic and chance, to some degree at least.

Furthermore, interrelationships between test taker and teacher perceptions were identified in the teacher interview data. Teachers’ perceptions about test takers and about score meanings shaped their perceptions of the appropriateness of the score requirements. Teachers’ perceptions of the test taker perspective challenged their views of the relationship between tests and learning, and also undermined their capacity to fulfil their teaching responsibilities, which include facilitating language learning, as they saw it. Furthermore, their experiences were characterised by conflict; over the meaning of scores and between the impossibility of the requirements and the need to justify their own role in providing courses which target people seeking to meet the requirements.
CHAPTER 8: THE DYNAMICS OF TEST IMPACT

8.1 Introduction

The use of English language test scores in Australia’s skilled migrant selection processes highlights the complex and potentially conflicting functions served by tests as instruments of immigration policy. Shifting score requirements within the Points Test for skilled migration have resulted in uncertain migration outcomes for many international graduates of Australian universities, including the four migrant participants in the current study. Since changes introduced in 2011 and 2012, the scores needed by migrants to become eligible to transition from temporary to permanent resident status have been contingent on the number of points they accumulate across other, non-language-related criteria, as well as on the number and profile of other permanent visa applicants. While some migrants need scores of IELTS 6, others, such as the four individuals in this study, need scores of IELTS 7 or IELTS 8, regardless of the fact that they have been living in Australia for several years and are already in employment. As pointed out in Chapter 2, the contingent nature of the score requirements undermines any link between test scores and the level of English language required to access skilled work in Australia as it might be imagined by language testers.

In Chapters 3 and 4, the premise underlying existing theoretical frameworks - that intended test purpose is defined and specified by test developers, in consensus with score users, as a precursor to evaluations of validity and fairness claims - was argued to be incommensurate with the nature of Australia’s skilled migration policy and other similar policy contexts in which tests operate. Immigration policy spaces are characterised by multiple and conflicting perspectives and intentions, many of which can only be surmised rather than specified. Moreover, test purpose can be further obscured by complex and dynamic interactions between test scores and other aspects of policy, as is the case in the Australian skilled migration context. Furthermore, as argued in these chapters, the view of test takers adopted in language testing theory, in terms of abstracted and decontextualized components of knowledge and skills, precludes any examination of the complex interrelationships between test takers’
thoughts, feelings and actions, and the roles these likely play in shaping score meanings and consequences in high stakes policy domains.

Findings from the current study provide further support for these arguments, as developed throughout Chapters 2, 3 and 4 of this thesis, in two main ways. Firstly, the experiences of the four migrant participants in the study, described in Chapter 6, highlight the importance of taking account of the role that individual subjectivities and agency play in influencing the consequences associated with the uses of tests in high stakes policy contexts. Five key themes characterised these experiences: understanding target scores, self-appraising, interpreting score outcomes, practising agency and responding affectively. Taken together, these constitute a model of the dynamics of test impact that foregrounds the experiences of test takers over the intentions of test developers and users. It is argued here that this model, presented in Figure 8.1 and discussed in section 8.2, below, provides a way of shedding the condition of determinability that has hindered efforts in the field to account for the social and political dimensions of test use in immigration policy contexts.

Secondly, as will be highlighted in this chapter, an examination of the experiences of test takers in this policy domain – a focus on social realizations rather than ideal intentions, in Sen’s (2009) terms - revealed several validity and fairness issues that undermine the legitimacy of test use in this policy domain. As will be shown, these include conventional validity issues, as well as issues related to substantive fairness in Kane’s (2010) terms. A comparison with findings presented in Chapter 7 will demonstrate that test taker accounts of these issues were largely supported by accounts from the teacher stakeholder group. To conclude the chapter, the justice implications of the test consequences that emerged across the migrant cases will be considered, in terms consistent with notions of justice put forward by Sen (2009) and McNamara and Ryan (2011).
8.2 A model of the dynamics of test impact

As described in Chapter 6, each of the four individuals engaged in iterative and interrelated cycles of comparing, interpreting and evaluating new and previous experiences of learning and using English, and of preparing for and taking English tests as they sought to make sense of their test experiences for permanent residency. These cycles were characterised by five key themes: understanding target scores, self-appraising of English abilities, interpreting score outcomes, practising agency and responding affectively.

Across the four cases, understandings of target scores were derived through interactions between test perceptions and the process of self-appraising. On the one hand, test perceptions were based on evaluations of the test construct, in particular, the extent to which test tasks and the score requirements were thought to be representative of the language demands faced by migrants living and working in Australia. On the other hand, they were based on beliefs about test purpose, which depended on the extent to which individuals perceived the test to be serving political intentions rather than a measurement function. Self-appraising of English abilities involved evaluating experiences of learning and using English across study, work and social contexts, and individuals made comparisons between these evaluations and their perceptions of their test experiences as part of situating themselves and their English usage in relation to their understandings of target scores. These interacting processes shaped and were shaped by interpretations of score outcomes across repeat test encounters, which provided a form of iterative feedback on self-appraisals and understandings of target scores, and prompted different affective and action responses over time as individuals attempted to exercise some degree of control over their test and migration outcomes. Throughout their engagement in this iterative cycle of interacting processes across repeat test encounters, individuals experienced, negotiated and generated test impacts in dynamic and unpredictable ways, as discussed in relation to each migrant case, further below. These interacting processes, as mentioned in the introduction, together constitute a model of the dynamics of test impact, illustrated in Figure 8.1.
The model of the dynamics of test impact, shown in figure 8.1, above, provides an overview of the complex ways that individual migrants involved in this study integrated test perceptions and self-perceptions into decision-making and action processes as they negotiated test and policy constraints to realise their migration intentions. A comparison of the consequences that emerged in individual cases, presented below, further highlights the importance of incorporating test taker subjectivities and agency into conceptualisations of test impact. As shown below, the test functioned in different ways...
in each case, producing diverse and unpredictable impacts on migrants’ sense of self and belonging, and on their engagement with work, study and the wider community.

8.3 Individual subjectivities, agency and test consequences

As established in Chapter 3, empirical investigations of consequences in the field of language testing have focused primarily on educational settings, with studies investigating the impact of language tests on teaching and learning outcomes (e.g., Alderson & Wall, 1993; Cheng, 2005; Cheng et al., 2004; Hamp-Lyons, 1997; Wall & Anderson, 1993). Within such studies, as argued in Chapter 3 (3.2.2), test consequences are conceptualised in dichotomous terms as either positive effects on teaching and learning, or negative effects on teaching and learning (Chen, 2008; Wall, 2013). However, as demonstrated by existing research into use of tests in educational policy contexts (see Chapter 3, 3.4), teacher subjectivities and agency play a significant role in shaping score meaning when tests enter the public domain, generating diverse and conflicting consequences, which cannot be categorised in these dichotomous terms (Hardy, 2013; Evans & Hornberger, 2005; Menken, 2008, 2009; McNamara, 2011; Shohamy et al., 2010; Van Avermaet & Pulinx, 2010). In addition, a limited number of studies investigating individual experiences of negotiating the language requirements for citizenship have highlighted the contingent nature of consequences (Cooke, 2009; Khan, 2013).

As demonstrated in Chapter 6 and illustrated below, findings from the current study show that test takers' thoughts, feelings and actions, which are rarely examined in studies of test impact, similarly played a key role in generating the consequences associated with testing practices in the high stakes context of Australia's skilled migration policy. In this context, the way in which the test actually functioned in each of the four cases depended, in part, on each individual's perception of test purpose. These perceptions shaped the effects of the test on their sense of self and on their actions, which, in turn, generated broader consequences, as shown below.
While Ana, Erfan and Sharon emphasised the potential for the test to function as a tool of exclusion, as will be discussed further below in relation to the validity and fairness issues associated with test use in this policy context, all four individuals also took the view that tests are primarily intended to function as measurement instruments, providing external, standardised evaluations of test takers’ English abilities. This is consistent with the policy rationales typically used to support the use of standardised testing practices, particularly in educational policy domains (Shohamy, 2001, 2006), and aligns with the view of tests that underlies much of the validation work undertaken by language testers, as set out in Chapter 3 (3.2). However, there were differences between individuals concerning the extent to which they perceived IELTS to provide an accurate and meaningful evaluation of their English skills and usage, which shaped test impact in diverse and unpredictable ways in each case.

Ana, for example, contested the appropriateness of the score level required for permanent residency, but nonetheless accepted that IELTS provided a meaningful measure of test takers’ English levels. This perception of the test was associated with conflicting impacts in her case. One the one hand, it allowed Ana to situate her score goals for permanent residency as achievable and to position herself as agent, able to exercise some control over her test and migration outcomes. As reflected in her accounts of her experiences learning and using English since arriving in Australia, Ana assumed agency in relation to her successes in overcoming English language related obstacles, and attributed success in language learning and in achieving legitimacy as a functional English user to her own hard work and tenacity. A view of the test as a meaningful measure of her English skill level enabled her to position herself in a similar way in relation to the score requirements - as a successful learner capable of making further improvements to her English skills through effort and perseverance. Thus, her perception of the role of the test enabled her to maintain a sense of agency, and to mitigate the impact of the fears associated with the uncertainty of her migration status (I’m afraid. I’m shaking. I’m very worried. But life goes on, and you have to do your best).

On the other hand, as evidenced by dramatic fluctuations in her self-appraisals and affective responses across test encounters, Ana’s belief in the authority of the test as an
evaluation of her English skills also left her self-perceptions vulnerable to test 'failures', triggering self-doubt and thereby simultaneously exacerbating the uncertainty of her situation. As demonstrated in Chapter 6, through her test experiences, she came to perceive a difference between her use of English and normalized forms of English usage, which prompted her to question her sense of accomplishment in becoming an English speaker and to doubt the sense of belonging she had derived from her successes in using English in study, social and work contexts (I have this feeling that everything is wrong with me).

As noted in Chapter 3 (3.3), Shohamy (2001) argues that language tests provide governments with a means of creating and reinforcing categories of belonging and exclusion, by classifying individuals according to their ability to assimilate to language behaviours that align with socially and politically dominant, typically monolingual, ‘norms’. While the consequences felt by Ana appear to be consistent with this argument, Shohamy, tends to emphasise the exclusionary and punitive role of tests. However, Ana’s acceptance of the authority of the test also meant that positive feedback she received in relation to her test experiences, including her ultimate test success, enabled her to establish a renewed and even reinforced sense of legitimacy and belonging (At the end of the process I understood, finally, that I had enough English... I love Australia, I love English). Her shifting and contingent self-perceptions thus highlight the indeterminable nature of the consequences associated with testing practices; the test in Ana’s case both offers inclusion and threatens exclusion, with any act of classification remaining unstable and incomplete (McNamara, 2012b).

In Erfan’s case, his test experiences appeared to have no impact on his self-perceptions. From the outset, he dismissed the legitimacy of the test as a measure of his English skills on the basis of its lack of relevance to the English demands he faced in his life and work in Australia, and the arbitrariness he perceived in the score requirements. However, in resisting the authority of the test, Erfan was unable to situate himself as agent in the way that Ana did. Whereas she was able to mitigate the impacts of uncertainty by assuming some level of control, Erfan appeared to feel the weight of the irrationality of the policy and the uncertainty created by the score requirements more acutely (I have
no way to escape... It’s not very good, it’s very difficult and you know the whole, my life depends on it).

Furthermore, Erfan’s rejection of the authority of the test meant that he was unable to derive any sense of legitimacy from his ultimate test score success. As demonstrated in Chapter 6, Erfan adopted Australian-born ‘native’ English speakers as a normative reference point for his self-appraising. He felt trapped in a cycle of social exclusion because he situated himself and his English skills in a deficit position compared to these more legitimate English speakers. Comparing the cases of Erfan and Ana, in conducting self-appraisals they each oriented to similar ‘orders of indexicality’ in Blommaert’s (2010) terms - that is, similar notions of what counts as language and who has access to it. However, they held opposing views of the test in relation to these norms, which was associated with different test impacts in each case. In terms of what counts as language, Erfan perceived the English usage of Australian-born ‘native’ speakers to constitute 'legitimate' English, and only those who achieved high levels of assimilation through integration with this in-group of speakers could access this legitimacy. In Erfan's view, the speaking test provided no measure of English in these terms, which he evidenced by comparing his success on the test with his lack of success in establishing social connections with this group. For Erfan, positive test experiences thus could not mitigate his vulnerability to judgment and potential exclusion from the Australian-born community, despite his newly acquired eligibility for permanent residency status.

Ana also associated gaining legitimacy as an English speaker with achieving integration with Australian-born English speakers. Prior to commencing her test efforts, she felt she had already achieved this. However, when she experienced inconsistencies between her test outcomes and her lived experiences of using English, she privileges the authority of the test as an embodiment of the normative framework against which to evaluate her status as an English speaker. Thus, negative test experiences prompt Ana to suspect that her Australian-born friends and colleagues have been accommodating her deficiencies, rather than accepting her as an equally legitimate English user. As Blommaert (2010) suggests, the normative frameworks that individuals adopt as a reference for evaluating and positioning themselves within social hierarchies are located within shifting social
spaces, and are thus dynamic rather than stable. This is exemplified by Ana's shift in orientation, away from her lived experiences towards the test as a point of reference, and her shifting interpretations of score meanings across repeat test encounters. Although she accepts the authority of the test, as she moves through different expert tutors, she continues to interrogate and destabilise the normative framework the test embodies, accepting and dismissing different readings of it as she seeks to pin down the elusive norms to which she feels she needs to assimilate in order to succeed on the test. For Ana, as suggested earlier, the threat of exclusion and the promise of inclusion embodied by the test remained tentative and open to contestation.

The test in Erfan's case, besides not mitigating his vulnerability to native speaker judgment, also reinforced his sense of social exclusion. As noted in Chapter 6, he decided to forego participating in activities that would have enabled him to meet and engage with members of the local community so that he could dedicate his time and financial resources to his test efforts, despite suffering from loneliness and feeling a sense of desperation concerning his social isolation in Australia. In addition, the test transformed his employment goals. Because he understood a lack of construct relevance to mean that his everyday English usage did not provide any form of test preparation, he decided that if he was unable to negotiate with his employer to reduce his work hours to free up time for practicing IELTS tasks, he would leave his job, which was in his area of expertise, and "work four hours in McDonalds or something like that". In her examination of migrant experiences of temporariness in Australia, Robertson (2016) also found that the English test requirements transformed the ways in which migrants engaged with work, pushing individuals to renegotiate their goals in unanticipated directions. The experience of Robertson's participant, Liu, mentioned in Chapter 2, was similar to Erfan's experience - Liu also felt compelled to sacrifice a job in her profession for casual employment in hospitality to create more time for IELTS preparation.

In Sharon's case, as was the case with Erfan, her experiences of the test had no impact on her self-perceptions. Like Erfan, Sharon emphasised the artificiality of the test construct. She felt that IELTS was not representative of the English demands she faced in the contexts of her graduate studies and employment, and that the writing behaviours
needed to succeed on the test actually conflicted with the behaviours valued in these contexts, where careful reflection is valued over speed. However, while Erfan, like Ana and Yuki, identified as an English language learner to some extent, and so accepted the appropriateness of subjecting his English skills to external judgment, Sharon identified as a bilingual English speaker and rejected the need for any such evaluation in her case. She had grown up using English in Sri Lanka, and had also demonstrated her ability to achieve a high standard of academic writing throughout her tertiary studies in Australia. Moreover, she had moved on to a full-time research position in a university, which also required high level writing skills, soon after her graduation. She thus rejected the appropriateness of English testing in her case, and dismissed the legitimacy of IELTS scores as an evaluative measure of her writing ability. Instead, she perceived that the test played an exclusively political role in her migration trajectory, obstructing access to permanent residency and serving no other purpose.

Unable to attribute any meaningful purpose to her test preparation activities, Sharon quickly became frustrated by her inability to produce the artificial writing performance she felt was needed to achieve IELTS 8, and sought legal advice in an effort to circumvent the requirement to reach these scores. She succeeded in doing this and, as mentioned in Chapter 6, she applied for a state-sponsored permanent visa on the basis of her existing IELTS scores. However, with scores of IELTS 7 instead of 8, her Points Test total just met the minimum requirements for eligibility for this alternative visa category, which, as her lawyer advised her, meant that her application was not likely to be prioritised for selection. As a result, she remained in limbo for a further 12 months after submitting her visa application. Thus, while Sharon was able to exercise agency to shift the score requirements in her favour, she was unable to completely mitigate the uncertainty of her situation. As shown in Chapter 6, her unresolved visa status undermined her sense of belonging in Australia, disrupting her efforts to properly realise social integration; she resigned herself to a "nomadic lifestyle", remaining unsettled and living in temporary student accommodation long after finishing her studies. She directly attributed her inability to shed this transient 'student' identity and to establish a more "homely environment" to uncertainty over her migration status ("I didn’t quite get to doing that because I wasn’t quite sure whether I can stay here for good").
Finally, in Yuki’s case, migration intentions intersected with constraints imposed by the test and policy to prompt her to renegotiate work and study goals in dynamic ways. As demonstrated in Chapter 6, Yuki constructed a complex and convoluted migration pathway in order to continue living in Australia, engaging in a process of moving through different temporary visas as a tactic to enable her to side-step the English score requirements that she was unable to meet at various points in her trajectory. Robertson (2016) notes that visa ‘churning’ has emerged as a consequence of the Australian government's emphasis over the past decade on temporary rather than permanent forms of migration, and that little is yet known about the social impact of these "under researched and politically 'hidden'" (p. 54) migration pathways. Consistent with Robertson’s (2016) finding, mentioned earlier, that migrants renegotiate their goals depending on the policy and test constraints they face, Yuki’s complex migration pathway also meant that she was forced to postpone her goal of gaining skilled employment as a nurse and to construct new, intermediary work and study goals according to the constraints of her different visa conditions - she engaged in casual hospitality work and decided to undertake a second nursing degree. While this goal-shifting can be seen as an effect of the coercive power of the test in this policy domain, the process of visa churning also represents a subversion of the social categories designated by policy and assigned by the test - Yuki transgresses the boundary between categories of temporary and permanent, constructing instead a kind of ‘temporary permanence’ which serves as a form of unintentional resistance to the authority of the test and policy, enabling her to continue pursuing her overarching goals.

In summary, as shown above, individual subjectivities and agency intersected with the constraints of the test and score requirements in this policy context to generate diverse and shifting consequences in each of the four migrant cases, highlighting the indeterminable nature of test purpose and test impact in this policy domain. The test functioned in different ways in each case, contingent on individuals’ decisions and actions throughout their migration trajectories, which derived from interactions between dynamic test perceptions, self-perceptions, goals, and intentions. As detailed above, migrants responded to the test score requirements in ways that variously
undermined their sense of self as English speakers, transformed their broader social and employment goals, and disrupted their social and labour market integration. The perceptions and experiences of these four migrants, including the ways in which they felt and shaped test consequences, detailed above, raise several questions concerning the validity and fairness of test use in this policy context, as will now be discussed below. As shown below, teacher accounts of their experiences of test takers seeking permanent skilled migrant status and the use of IELTS in this policy domain (see Chapter 7) were largely consistent with migrant accounts in relation to these issues.

8.4 Validity and fairness issues

As shown in the sections above, in each of the four migrant cases, individuals’ understandings of target scores derived, in part, from perceptions of test purpose. Each individual accepted that English tests, in general, are designed to serve as measurement instruments, intended to provide evaluations of test takers’ English language knowledge and skills. As demonstrated in the current section, migrants and also the teachers in this study situated this evaluative purpose in relation to their perceptions of what would constitute a legitimate policy aim - to ensure that those selected for skilled migration were capable of achieving integration into the Australian labour force. Thus, their judgments of the legitimacy of IELTS as a tool to measure English proficiency were contextualised, oriented to perceptions of the extent to which the score requirements aligned with test takers’ abilities to function adequately in English in their professional domains.

As demonstrated in Chapter 6, migrant perceptions of whether or not IELTS was fit for this purpose varied, case by case and over time. Perceptions that the test was not functioning as an effective measurement instrument prompted suspicions that the use of scores served a primarily arbitrary political purpose in this policy context, which was unrelated to migrants’ actual English skills, thereby raising issues of fairness. Some of the teachers, as demonstrated in Chapter 7, were similarly oriented to the latter purpose, drawing a distinction between IELTS as a measure of English proficiency and
the arbitrary use of scores for skilled migrant selection purposes, and thereby questioning the legitimacy of test score-based decision-making in this policy domain.

Questions raised by migrants and teachers over the appropriateness of the use of IELTS in Australia's skilled migration policy are broadly categorised in this section according to whether they relate to the efficacy of the test as a measurement instrument, in which case they are discussed under the heading 'Conventional validity issues', or to broader perceptions of the political purposes served by the test, discussed under the heading 'Fairness issues'. As will be shown, conventional validity issues stem from conflicts between migrant test takers' abilities to use English effectively in real world contexts and their inability to reach the scores needed to transition from temporary to permanent resident status, which suggest the test construct lacks relevance to the English demands faced by migrants living and working in Australia. In addition, the phenomenon of repeat test taking, which migrant and teacher accounts suggest is characteristic of the transition from temporary to permanent skilled migrant status, revealed score inconsistencies across test sittings that further undermine validity in any conventional sense, especially given the high stakes involved in this policy domain. As discussed under the heading 'Fairness Issues', these conventional validity issues raise associated questions of substantive fairness (Kane, 2010) or 'justice' in McNamara and Ryan's (2011) terms, which relate to the reasonableness and appropriateness of interpretations and uses of test scores within Australia's skilled migration policy.

8.4.1 Conventional validity issues

Migrant experiences and teacher accounts of the use of IELTS scores as part of skilled migrant selection criteria highlight several potential threats to the validity of the test as a measurement instrument in this policy domain. This is perhaps not surprising, since, as mentioned in Chapter 3, IELTS and other well-established international standardised tests are now widely used for a variety of purposes, including the screening and selection of migrants, which are beyond the purposes for which the tests were originally designed. As pointed out in Chapter 3, IELTS was originally designed to assess the English language abilities of test takers wishing to enter university (IELTS Academic) or technical or
business colleges (IELTS General) (Ingram, 2004) and the appropriateness of its use as a tool for skilled migrant selection remains open to question.

All validation frameworks require that test tasks elicit the same language abilities as tasks in relevant real world domains, and also that scores reflect performance quality in these real world domains (Bachman & Palmer, 2010; Kane, 2006, 2012, 2013; Xi, 2008). In the current study, threats to validity emerged from inconsistencies between migrants’ test experiences and their experiences of using English in their employment situations and daily lives, which prompted perceptions on the part of migrants and teachers that test constructs were irrelevant to the English demands faced by skilled migrants in Australia, in terms of task design, scoring criteria and the score levels required for skilled migration. A further threat to validity emerged from perceptions and evidence of score inconsistencies across test encounters, which were associated with the phenomenon of repeat test taking. As argued in Chapter 4, any link between particular test scores and the level of English language required to access skilled work in Australia is impossible to sustain, due to the way English test scores interact with other, non-language related criteria within Australia’s skilled migration policy (see Chapter 2, 2.4). Thus, as emphasised in Chapters 3 and 4, any conventional validation argument is impossible to construct, given that all frameworks require that the intended meaning of scores is defined in specific terms (Bachman & Palmer, 1996, 2010; Chapelle et al. 2010; Kane, 2006, 2012, 2013; Xi, 2008). Importantly, however, perceptions of the legitimacy of the evaluative function of the test were central in influencing how migrants in this study responded to their test experiences, and thereby played a key role in shaping test impact in this policy domain.

As shown in Chapter 6 and above, migrants differed in how they accounted for inconsistencies between their test and real world experiences depending on their perceptions of test purpose, which shaped interpretations of and responses to score outcomes and influenced the consequences that emerged in each migrant case. Ana and Yuki, for example, tended to accept the authority of the test as an evaluative tool, regardless of any conflict between score outcomes and real world successes using English. In Ana’s case, tension between lived experiences and test experiences
prompted shifts in her sense of legitimacy as an English speaker, as already discussed. For Yuki, her IELTS score outcomes aligned well with her self-perceived English learning outcomes for most of her test-taking journey, and where inconsistencies emerged, she reconciled them by assuming the standard of English expected on the test was higher than in real world contexts \((I \ think \ it’s \ more \ strict, \ I \ guess, \ yeah \ that’s \ why \ I \ can’t \ get \ much \ score)\). As detailed further below, in Ana's case, validity issues became salient as she engaged expert help to negotiate the uncertainty of target score meanings, while in Yuki's case, as in all four cases, repeat test taking revealed arbitrary score inconsistencies.

In contrast to Ana and Yuki, Erfan and Sharon both dismissed the legitimacy of the test as a measure of their English skills, attributing their inability to readily achieve the score requirements for permanent residency to the irrelevance of the test construct, which they viewed as artificial and unrelated to the English demands they faced in their work and daily lives. Erfan, for example, highlighted a lack of alignment between the writing demands he faced on IELTS and the tasks expected of him in his profession, whether it be in English or in his first language \((I \ can \ write \ technical \ report \ very \ well \ but \ I \ can’t \ write \ an \ essay \ about \ a \ social \ thing)\). Sharon became aware that the test demanded a specific style of writing through the IELTS preparation class she took after her first unsuccessful test attempt, and she noted that this style was different to the type of writing she produced during her graduate studies and in her professional domain \((I \ noticed \ that ... \ their \ essay \ writing \ style \ is \ different, \ like \ it’s \ very \ specific)\). Furthermore, she felt the writing behaviours needed to succeed on the test conflicted with the behaviours valued in the contexts of her graduate studies and employment, as noted earlier.

Regardless of whether they accepted or rejected the legitimacy of IELTS as a measure of the English abilities relevant to living and working in Australia, all four migrants claimed that the score requirements for permanent residency were too restrictive, perceiving IELTS 8, in particular, to represent an unnecessarily high level of English that was unrealistic to expect of migrants from different language backgrounds. Ana, for example, understood IELTS 6, needed to access a temporary post-study work visa, to indicate an adequate level of English for skilled employment. This perception was
reinforced through her successful experiences communicating in English in her work context, which prompted her to question the purpose behind the introduction of requirements for IELTS 7 or 8 to gain a permanent visa (the thing is how they change the score so drastically...from 6 to 7 or 8... I am functional. I work in an Australian company. And if you ask them, I think, they have no problems with me and my communication).

Accounts provided by the teacher participants were consistent with migrant perceptions that these scores were too restrictive. As demonstrated in Chapter 7, teachers characterised the migrant test takers in their IELTS preparation classes as “very fluent, very capable communicators” with “really high levels of language ability”. Teachers similarly felt that the English proficiency levels associated with IELTS 7 and especially IELTS 8 were above and beyond the level required by skilled migrants to function well in English, also noting that despite being highly proficient in English and already in employment, many of the migrants in their classes were unlikely to achieve such scores. They felt this was particularly the case in writing, which was also the subtest that all four migrants in the current study felt presented the most significant obstacle.

While the four migrants in the current study, as noted above, varied in how they understood their lack of immediate score success, teachers generally attributed the unattainability of these high IELTS scores to the irrelevance of scoring criteria, which they felt emphasised aspects of English usage that had little impact on migrants’ abilities to communicate effectively in professional and other domains. For example, scores of IELTS 8 in writing were associated by Claire, one of the Directors of Studies and an experienced IELTS examiner, with “perfect paragraphing” and a low frequency of errors in the use of low salience features such as third person ‘s’, which she described as “things that really don’t matter”. In relation to speaking, one of the teachers, Michelle, who was also an IELTS examiner, associated IELTS 8 with speech that was free of “pausing, repetition, self-correction”, which she felt represented a level of correctness that was “not at all reflective of everyday conversation”. Teachers also perceived that IELTS scoring criteria were not reflective of real world English usage even among English ‘native’ speakers, many of whom they felt would be unable to achieve the IELTS scores
needed for permanent skilled migration, due to the artificial nature of the performance requirements.

Migrant and teacher accounts thus suggest that scores of IELTS 7 and 8 do not align with the English abilities needed by migrants to communicate effectively in their lives and work in Australia, and furthermore, that scores of IELTS 8 are unattainable and therefore exclusionary, as further discussed below under the heading 'Fairness and Justice Issues'. As shown in Chapter 7, teachers responded to the potentially exclusionary role of the score requirements by attempting to encourage migrant test takers to further develop their English skills to enable them to cope with the unfamiliar and high level demands they encountered on the test. However, they found that migrant test takers resisted such efforts because they were convinced that score success depended on accessing a “magic formula” rather than making substantive improvements to their English level. While teachers tended to characterise this perception as misguided (You try to make the students realize that there is an overall learning... that there isn’t a list of things that make you a mythical IELTS 8), Ana’s experience suggests that it has a rational basis; as shown in Chapter 6, Ana’s third tutor attributed her prior lack of score success to a failure to deliver the expected IELTS performance formula, rather than a deficit in Ana’s English skills, as she had been assuming (... he told me something very interesting: ‘Ana, you’re a good writer but just tell them what they want to see just for the test and you’ll be alright’). She went on to achieve her score goals by abandoning efforts to improve her English skills and focusing instead on producing the formula that this tutor, who was also an IELTS examiner, provided her.

In addition to the potential irrelevance of the test construct, arbitrary score variations represented a further threat to the validity of the use of IELTS scores as a measure of migrants’ English proficiency. As mentioned above, the phenomenon of repeat test taking, which characterised the trajectories of the four individuals in the current study and was identified by teacher participants as a central feature of migrant test taker behaviour, gave rise to some apparently random shifts in scores. As shown in Chapter 6, in Ana, Erfan and Yuki’s cases, scores on individual skills fluctuated above and below
individuals’ target scores across test encounters within relatively short periods of time. In Ana’s case, for example, her listening scores changed from IELTS 8 in October 2012 to IELTS 6 in December to IELTS 7.5 in May 2013. Over the same period her reading scores similarly fluctuated from IELTS 6.5 to 6 to 7. Erfan scored IELTS 8 in reading in February 2013 and IELTS 7 two months later. Conversely, Yuki scored IELTS 6.5 in reading in July 2012 and then IELTS 8 for this skill two months later. The issue of score inconsistencies was not limited to IELTS. Yuki’s scores shifted across each of her last five OET encounters, which took place within a six-month period; on each attempt, she failed to reach the score requirements on only one of the four skills (speaking, writing, listening and reading), a different one each time.

As shown in Chapter 7, teacher participants associated repeat test taking with a belief on the part of migrant test takers that score success was contingent, at least partly, on chance (they will book three at a time. It’s a bit like a lottery). As demonstrated in Chapter 6, Erfan viewed score outcomes as contingent on chance from the outset, which motivated him to adopt repeat test taking as a key strategy for achieving his score goals. Ana became aware that chance was a factor in determining scores prior to her final test attempt, based on information from the third tutor (he told me ‘sometimes it’s luck Ana...because they work under a lot of pressure) and from a story she heard from another test taker, who told her that his writing scores went from IELTS 6 to IELTS 7 from one week to the next (he took the test, for example, this Saturday and he got 6 and he took the test the next weekend and he got 7 in writing).

Although Sharon was quick to abandon the notion of taking the test repeatedly, she had also heard suggestions of arbitrary inconsistencies in score outcomes. As shown in Chapter 6, teachers of the IELTS preparation course she took prior to her second test attempt advised students that they would achieve higher scores if they returned to their home countries to take the test because in Australia, examiners orient to ‘native’ speaker norms in evaluating performances. Sharon also heard that this was the case from friends who had returned to Sri Lanka to take IELTS, achieving success there after failing to reach the scores they needed in Australia. Yuki’s experience also suggests that a location-based scoring bias may exist, as she had been unable to achieve the IELTS 6.5
needed for university entry while she was in Australia but reached these scores when she returned to Japan and took the test.

As suggested above, evidence of a lack of relevance of tasks and scoring criteria and of inconsistencies in score outcomes across different test encounters, which emerged through an examination of individual experiences of the test, threaten to undermine the validity of the use of IELTS as a measure of the English abilities needed by migrants to function effectively at work and in their lives in Australia. Perceptions of these validity issues were further associated with claims of unfairness, as set out below.

8.4.2 Fairness issues

As mentioned in the subsection above, migrant and teacher participants felt that scores of IELTS 7 and 8 were unnecessarily high, representing a level of proficiency above that required by migrants to meet the English communication demands of living and working in Australia. This triggered perceptions that the test served an arbitrary political function, which prompted claims of unfairness. As discussed below, these claims of unfairness derived from a conventional validity issue (construct irrelevance) but focused primarily on the cultural and linguistic norms underlying test constructs and the potential for the use of test scores, given these norms, to deliver exclusion as a consequence.

In the absence of a legitimate evaluative function, Erfan, for example, felt the test operated as "a very fine filter", with the score requirements serving primarily to regulate the number of applicants and to exclude migrants from undesirable backgrounds. As noted earlier, Erfan perceived the English usage of Australian-born 'native' speakers to constitute 'legitimate' English, which he sought to acquire but felt was only accessible through assimilation. While he perceived the speaking test to be unrepresentative of 'legitimate' English, he felt that 'native' speaker norms were embodied in the writing test construct, which represented a source of unfairness, according to his perceptions (I shouldn’t be expected to be an English writer, it’s not fair to expect me to have the same sort of ah skills as a native speaker). Sharon similarly associated scores of IELTS 8 with
'native' English norms, and felt that it was unreasonable to expect English as a Second Language speakers to achieve the same level of performance as ‘native’ speakers (English is not the mother tongue of most of us so it’s not fair to think us of being of very, very good high standard as maybe a native speaker would). Erfan’s position, supported to some extent by Sharon’s perception of unfairness, echoes concerns raised by Berg (2011), who, as mentioned in Chapter 2, claims that the test requirements within Australia’s skilled migration policy reinforce a message that homogeneity is preferred over diversity and that belonging depends on assimilation to dominant cultural and linguistic norms.

Teachers also felt that test constructs embodied dominant Standard English norms, although they distinguished these from typical ‘native’ English speaker behaviours, perceiving the norms underlying scoring criteria for IELTS 8, in particular, to be highly prescriptive and artificial, atypical of English usage even among ‘native’ speakers, many of whom they felt would be unable to achieve the score requirements, as mentioned earlier. Nonetheless, some teachers maintained a view consistent with the perspective adopted by De Jong et al. (2009), discussed in Chapter 4 (4.2), perceiving that the high test score requirements served to encourage migrants who were already functional in English to further improve their English skills. These teachers thereby felt the test requirements were beneficial for migrants, which they believed legitimised the use of scores in migrant selection processes.

By contrast, other teachers felt that the use of IELTS 8 as part of migrant selection criteria was exclusionary, as these scores were perceived to be unattainable for highly proficient English learners and for speakers of different varieties of English, regardless of any efforts they might make, due to the characterisation of the grammatical and pronunciation features characteristic of the English usage of these groups as ‘errors’. Perceptions that the test privileges dominant Standard English norms and situates other English usages in a deficit position are consistent with arguments in the literature that language testing in immigration and citizenship contexts privileges monolingualism, de-values minority languages and perpetuates the exclusion of minority speakers (e.g. Blackledge, 2009; Horner, 2009; Shohamy, 2009; Van Avermaet, 2009).
Ruth, for example, perceived non-standard features to be “ingrained” and “automatic” in the English usage of adult migrants from non-English speaking backgrounds. She felt that although these features do not hinder communication, they make IELTS 7 and especially 8 unattainable. Kate highlighted a potential bias in the IELTS construct that privileged cultural norms in English speaking societies, where directness is valued in written expression, which she felt served to exclude test takers from different backgrounds. Jane similarly emphasised the exclusionary function of IELTS scoring criteria, which she felt were based on particular norms of pronunciation and prosody that privilege certain varieties of English (and speakers) and disadvantage others (You get people from the sub-continent who maybe have spoken English since they were very young ... ...even though they might be very fluent, very capable communicators ... you have to mark to the criteria then they don’t get those high scores they want). Finally, Claire felt that the score requirements were unnecessarily restrictive (it is unrealistic...I really feel like the hoops are unnecessary) and that the IELTS construct served an exclusionary function in this policy context (it also really prioritises or favours people from an English-speaking background for immigration purposes).

Perceptions on the part of migrants and teachers that the score requirements were exclusionary, representing a level of English that is unrealistic to expect of migrants who have learned English as adults or who speak other varieties of English, undermine what Kane (2010) terms ‘substantive fairness’, or ‘justice’ in McNamara and Ryan's (2011) terms. As noted in Chapter 4 (4.3), Kane (2010) stipulates that substantive fairness relates to the way that a test actually functions in a given context, in terms of whether test takers are treated equally or equitably, and the extent to which the ‘rules of the game’ are considered to be reasonable and appropriate. Although Kane does not problematize notions of ‘reasonable’ and ‘appropriate’, as McNamara and Ryan (2011) argue, these are value-laden notions. As highlighted by the existence of diverse perceptions, particularly among teachers in the current study, any fairness claims thus remain open to contestation. However, as argued below, the actual consequences that emerged from test use in this context, as experienced and felt by the four migrants involved in the current study, call the appropriateness of the test score requirements
into question in this policy domain, regardless of whether or not the test is considered to be valid or fair as a measurement instrument.

8.5 Justice as consequences

Sen's (2009) conceptualisation of justice in terms of actual social realizations rather than idealised notions of what constitutes fairness, detailed in Chapter 4 (4.4) offers a way forward in accounting for the diverse perspectives, conflicting interests, and disputes over values that characterise test use in immigration policy contexts. Consistent with the notion of justice put forth by Sen (2009), it is argued in this thesis that in order to explain and account for the issues raised by the uses of language tests in immigration and other policy contexts, the perceptions and experiences of those subjected to testing regimes, rather than the intentions of test developers, must constitute the focal point around which the appropriateness of testing practices is evaluated.

In the current study, as demonstrated in Chapter 6 and highlighted in section 8.3 of this chapter, individuals’ test perceptions, including those related to validity and fairness, described in section 8.4, played a role in shaping self-perceptions and individual agency as each migrant sought to realise their migration intentions. Their responses to their test experiences, in turn, played a role in mediating and generating test impacts in various ways, influencing their sense of legitimacy as English speakers, disrupting their social integration and sense of belonging, and transforming their engagements with work and study in ways that conflicted with personal goals and also, arguably, the aims of Australia’s skilled migration policy. As shown in Chapter 6, each migrant further perceived the prolonged periods of uncertainty over their migration status, which were a direct result of the difficulties they faced in achieving the English test scores they needed to transform from temporary to permanent status, to impact negatively on their emotional and psychological well-being. As noted by McNamara and Ryan (2011), the question of whether a test should exist in a given policy context is rarely asked by test developers, but as highlighted in this chapter, the consequences that emerged in each of the four migrant cases undermine the appropriateness of the use of English test scores for migrant selection purposes, especially given that all four were already
demonstrating their capacity to meet the English demands of living and working in Australia.

8.6 Conclusion

This chapter provided a discussion of the findings of the study, as presented in Chapters 6 and 7. As demonstrated in this chapter, the experiences of the four migrants and the accounts provided by the seven teachers involved in this study evidence the need for a more nuanced understanding of test impact, which foregrounds the importance of individual subjectivities and agency in shaping both the meanings of test scores and the consequences that emerge from the uses of tests as policy instruments. A model of the dynamics of test impact was proposed, which privileges the perceptions and experiences of test takers over the intentions of test developers and users. It was argued that this offers a means of shedding the condition of determinability that has limited efforts in the field of language testing to evaluate in any meaningful way the appropriateness of the use of language test scores for regulating access to permanent residency and citizenship rights. An in-depth examination of the experiences of four individuals seeking access to such rights revealed complex interactions between individual self-perceptions and test perceptions, and between individual intentions and the constraints imposed by the test and the policy, that taken together were involved in generating test consequences. Furthermore, a focus on the test taker perspective in this study brought to light several threats to validity and fairness that undermine the integrity of the tests in question and, moreover, highlight the impossibility of maintaining any sort of artificial distinction between the measurement properties of tests and their social and political uses.
CHAPTER 9: CONCLUSION

9.1 Introduction

This thesis explored the dynamics of test impact by exploring how four individuals, all former international graduates of Australian institutions, perceived and responded to the English test requirements that regulate the transition from temporary to permanent visa status within Australia’s skilled migration policy. The experiences of these four individuals were examined over time as they engaged in repeat test attempts as part of their efforts to gain permanent residency, to ascertain if and how their perceptions shifted as they interacted with the test, and how their perceptions of and interactions with the test shaped their lives, influenced their actions and decisions, and produced consequences. The perspectives of seven teachers of IELTS preparation courses were also sought, as a means of situating the experiences of the four individuals involved in the current study in relation to the experiences of a broader group of test takers seeking scores of IELTS 7 or 8 to gain permanent residency as skilled migrants, and also to interrogate the plausibility of the fairness and validity issues that emerged from the four migrants’ accounts of their test experiences.

In this chapter, a brief overview of the theoretical challenges the thesis aimed to address is first provided. This is followed by a summary of the main findings. The chapter concludes with a discussion of the implications of the study for theorising the social and political dimensions of testing practices, the study limitations, and areas in need of further research.

9.2 Overview of Theoretical Challenges

As demonstrated in this thesis, the use of language tests as tools of immigration policy raises critical questions concerning the adequacy and appropriateness of current theoretical conceptualisations of test use, test consequences, and test validity in the field of language testing. Within existing theory, these notions rest on the premise that test purpose and associated score meanings exist as coherent, fixed, and stable entities.
Test developers and test users are assumed to be in consensus concerning how test scores should be understood and used, and when tests and test scores are used as intended, the premise underlying evaluations of consequences is that positive effects on individuals and societies should outweigh or at least justify any negative effects. However, the purpose of tests and testing practices, when situated within policy, are contingent on broader social, political and economic priorities, and score meanings, the values underpinning them, and the test and policy impacts that emerge, are often multiple, dynamic, and potentially conflicting.

In this thesis, the assumptions underlying validation frameworks in language testing were thus argued to be incommensurate with the immigration policy spaces within which language tests operate. Moreover, it was argued that the conceptualisation of test takers within existing validity frameworks in the field of language testing, as abstracted and decontextualized components of knowledge and skills, ignores the complex interrelationships between test takers thoughts, feelings and actions, and the role played by individual subjectivities and agency in shaping score meanings and consequences. As has been demonstrated, in order to account for the role tests play in the lives of individuals and in societies more broadly, tests within these contexts must be reconceptualised as politically situated and contingent sites of struggle, in which individual perceptions and intentions intersect with the constraints imposed by tests and policies to produce consequences that are fluid, multidirectional, and unpredictable.

9.3 Summary of main findings

The experiences of the four migrant participants in this study evidenced the need to reconceptualise tests and test takers in order to account for test impact in the context of Australia’s skilled migration policy. The findings of this study showed that individual subjectivities and agency played a key role in shaping score meanings and the consequences that emerged over time from the use of test scores in this policy domain. A model of the dynamics of test impact was proposed to describe the iterative and interrelated processes of comparing and evaluating that characterised the way the four
individuals in this study derived score meanings and situated themselves in relation to
the score requirements for permanent residency. These processes included self-
appraising of English abilities, understanding target scores, and interpreting score
outcomes, which interacted over time to prompt various affective and action responses
as individuals negotiated the constraints and uncertainties associated with being
subjected to a testing regime in this policy context.

As demonstrated in Chapters 6 and 8, the test functioned in different ways in each case,
contingent on individuals' thoughts, feelings, and actions. In Ana’s case, for example,
her test experiences variously undermined and reinforced her sense of legitimacy as an
English speaker, as her acceptance of the authority of the test meant that it
simultaneously offered inclusion and threatened exclusion. Erfan, by contrast, rejected
the legitimacy of the test as a measure of his English skills. While this meant that his self-
perceptions were protected from test failures, he remained vulnerable to the ‘native’
speaker standard against which he evaluated his English usage, and he was unable to
derive a sense of legitimacy from his test success in the way that Ana did. Furthermore,
the test reinforced Erfan’s sense of social exclusion and hindered his engagement with
skilled employment. His perception that the test lacked relevance to the real world
English demands he faced prompted him to abandon efforts to participate in social
activities and led him to negotiate with his employer to reduce his work hours, as he felt
he needed to dedicate all of his time to rehearsing the artificial performance he
perceived was required to achieve score success. Sharon, also rejecting the authority of
the test, maintained her well-established sense of self as an English speaker.
Nonetheless, she felt the uncertainty over her migration outcome created by the English
score requirements hindered her social integration, as she felt forced to maintain a
“nomadic lifestyle”. Yuki dealt with her difficulties in meeting the score requirements
over time by constructing a complex and convoluted migration pathway, which required
her to renegotiate her study and employment goals at various points but also enabled
her to resist the threat of exclusion embodied in the test and policy, and to continue
pursuing her overarching English learning, social integration, and employment goals.
The ways in which each of the four migrants interacted with the test also raised questions about the validity and fairness of the use of English test scores for skilled migrant selection purposes in Australia. As established in Chapter 8, accounts provided by the seven teacher participants largely supported the claims made by test takers in relation to these issues, and further suggested that the experiences of the four individual migrants in this study were reflective of the experiences of the broader group of test takers seeking scores of IELTS 7 and 8 for permanent residency that they encountered in their IELTS preparations classes.

A sense of conflict between migrant test takers’ abilities to use English effectively in real world contexts and their inability to reach the scores needed to transition from temporary to permanent resident status, which emerged from teacher and migrant participant accounts, brought several validity and fairness issues to light, as described in Chapter 8. For example, test constructs were perceived to be irrelevant to the English demands faced by skilled migrants in Australia, in terms of task design, scoring criteria and the score levels required for skilled migration. In addition, score inconsistencies across test encounters, associated with the phenomenon of repeat test taking, undermined the integrity of the tests in question as measurement instruments in this high stakes policy domain. Perceptions that tests lacked relevance and suitability as a measure of migrants’ English skills derived from an awareness on the part of the four individuals, through their own experiences, and on the part of teachers, that even migrants who had been living and working in Australia for several years were unable to readily achieve the score requirements. This awareness prompted claims of unfairness, as Erfan and Sharon, as well as many of the teacher participants, felt that the test construct embodied cultural and linguistic norms that served to exclude migrants from non-English speaking backgrounds, regardless of their ability to communicate effectively in English.
9.4 Implications for language testing

Findings in this study supported the argument made in this thesis that the perceptions and experiences of those subjected to testing regimes, rather than the intentions of test developers, must constitute the focal point around which the appropriateness of testing practices in policy domains is evaluated. The impacts that emerged in the context of test use in Australia’s skilled migration policy, combined with the validity and fairness issues raised by the lived experiences of the four migrants in this study, undermine the appropriateness of the use of English test scores as selection criteria for skilled migration in Australia, and raise doubts about the quality of the tests in question.

By investigating the processes that underlie interpretations of score meanings and the emergence of test consequences from the perspective of test takers, this thesis provides a framework to account for the role of individual subjectivities and agency in generating test impacts in high stakes immigration policy contexts, and offers insights into the complex and multifaceted ways language tests can function within these policy domains. In so doing, the thesis offers a basis for evaluating the appropriateness of the use of language tests in policy spaces that accounts for the indeterminable nature of test purposes, score meanings, and consequences in contested political spaces, and thereby contributes to furthering theoretical accounts of the social and political dimensions of language testing practices.

9.5 Limitations and future research

A scarcity of empirical research investigating the perceptions and behaviours of test takers seeking permanent residency and citizenship rights has meant that progress in accounting for the now widespread use of testing practices in immigration policy contexts has remained limited. In developing an empirical approach to account for the impacts of the use of language test scores from the perspective of test takers, this study offers a way forward in conceptualising the complex and multifaceted roles tests play as tools of immigration policy, as mentioned above. Although a focus on four single cases
is an obvious limitation of the current study, this was mitigated to some extent by the inclusion of the perspectives of seven English teachers. Nonetheless, further in-depth investigations into how individual stakeholders interact with and respond to language tests in immigration policy domains, and the impacts associated with their test experiences, are needed across a range of national contexts in order to build a robust knowledge base to better inform test design and validation processes, as well as to develop meaningful stakeholder engagement and assessment literacy initiatives.
REFERENCES


APPENDIX 1: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (ANA)

220. K: Yeah, yeah. I didn’t realise that your sp- on a um if you’re on a student visa your partner could work full time, that’s good
221. A: Yeah er
222. K: So he didn’t, did he have to do any kind of English test back then to start working, or, did he ever do any test?
223. A: No, no. As many Colombians, because ((sigh)) no one...how can I explain to you? When we were kids
224. K: Yep
225. A: our parents had no vision about the world so they didn’t worry about our English, so we never studied English when we were kids @@
226. K: OK, yep, fair enough @@
227. A: But they did their best
228. K: Yeah, yeah
229. A: they didn’t know how useful is English, how useful, useful English become in our life in the future
230. K: Yeah, yeah, yeah
231. A: So, because we didn’t speak English I started working as a cleaner.
232. K: Uh ha. When you arrived?
233. A: Yep
234. K: Is this for the first six months before you started studying?
235. A: Yeah. And after one year working as a cleaner I said enough, and I think my English is, is functional enough
236. K: Yeah
237. A: and it’s time to me to move on
238. K: Mm
239. A: So I quit, but, but because of, because he supported, supported me, you know what I mean
240. K: Yeah, yeah, yeah
241. A: He told me OK you can quit and try to find something else
242. K: Yeah
243. A: All right. At the same time that your husband was?
244. K: OK, then you did that before you started at Swinburne?
245. A: Yeah, yep
246. K: So, did you need to um...? You would have needed an IELTS score to be admitted at Swinburne?
247. A: No, because, no because Swinburne, because everything is a business here, if you study English with them, if you pass their test you can apply for their course.
248. K: Right. So that’s what you did for 20 weeks?
249. A: Yeah, but in Melbourne Uni it’s different
250. K: No OK, yeah, yeah
251. A: So I was afraid of my English level
252. K: Yep
253. A: so I decide to study. I wasn’t ready. Even when I started my Uni, I wasn’t ready. I don’t know how I did...
254. K: Yeah, yeah. But you passed?
255. A: I never failed any, any subject
256. K: Mm
257. A: and I love Jamie and I...He has worked very hard and my semester, no my year, cost $60,000
258. K: Yeah right

264
A: So I, ah, I, I, I
K: Yeah
A: how can I, I couldn’t disappoint him, you know?
K: Yeah, yeah, that’s a lot
A: But I used to cry, after classes I used to cry
K: Yeah?
A: because I couldn’t understand a word what they say to me
K: Mm
A: But your brain is wonderful, our brains or wonderful. At the beginning, I couldn’t
understand anything, but after 3 months or 4 months, everything becomes easier
K: Yeah?
A: you know.
K: And then you could understand?
A: Yep, and I passed, and I learned a lot of things.
K: Yeah, that’s great.
A: No, it’s Graduate visa. After finishing your studies in Australia
K: OK right yep
A: OK? So you have to apply for a
K: Graduate visa
A: graduate visa, but some people call it temporary residency
K: Right
A: but the real name is graduate visa.
K: OK
A: OK? So you have to show to them that you are functional, so you, you are asked to get 6.
K: 6 on IELTS?
A: In every single band.
K: So you did that
A: Yeah
K: when you finished your…?
A: Yeah, I got 6.5 in every single band.
K: Oh yeah you told me that last time
A: Yeah
K: so when did you sit the test?
A: September 2011
K: OK, OK so you sat the test before you finished, before you graduated the Masters?
A: Yeah, yeah, yes, yes, yes
K: OK
A: I was starting my, no I was in the middle of the semester
K: Mm hm
A: Yeah.
K: Do you guys spend much time socialising with other people when you’re not at work?
A: Ah yes.
K: You do?
A: Yes, weekends
K: What kind of things do you do?
A: Ah we have, we are 6 basically. We are 3 couples
K: OK
A: But they are from Colombia as well.
K: Everyone is from Colombia?
A: Yeah. And I, and I speak a lot with Neil and Bridget, they are my neighbours
K: Mm hm, they’re...
A: Ah, my neighbours are so lovely.
K: Are they from- are they Australian?
A: Yeah. Super Australian@@@
K: Oh are they?
A: The funny thing is my, my English, my, my friends
K: Yeah
A: are Aussie but I consider them my friends
K: Yeah of course yeah
A: Er my neighbour
K: Yep
A: he's very Aussie. And Sara is super Aussie. And Jill, from work, she's super super
K: Aussie, and Dave as well
K: What, what makes them-?
A: the way they, the way they speak to me, their slang.
K: Oh yeah right, yep. So do you find it easy to understand them?
A: Yes!
K: Yeah
A: This is a different thing than the listening IELTS test @@
K: OK @@
A: Even if they use slang with me, I can understand everything. We went out on Friday with
Jill and Dave
K: Yep
A: and we were drinking and..
K: So these are the guys from your work?
A: Yeah, yeah. They are the closest ones
K: Mm hm
A: And we were drinking beer and we couldn't feel the difference between, you know...
K: Yeah, yeah, yeah. That's good
A: And so...@@@

K: Yeah. So, do you think that it's fair to have a test?
A: Fair?
K: Yeah... A language test?
A: It's fair?
K: Do you think so?
A: Yes and no
K: OK
A: I am on the bench, on the, on the fence
K: Yeah
A: Sitting, sit, I'm sit,
K: Sitting on the fence?
A: Sitting on the fence
K: Yep
A: Because this is a, this is an English country, everyone speaks English
K: Mm hm
A: and the government has to measure that in one or another way, huh?
K: OK. So that it's fair in that way?
A: In that way
K: Yeah
A: But the, the thing is how they change the score so drastically
K: Right
A: from 6 to 7 or 8...
K: Yeah
A: Is that the way they, they, they jump from one score
K: Yep
A: to...
K: Mm
A: And I am functional. I work in an Australian company
K: Yeah
A: And if you ask them, I think, they have no problems with me and my communication
K: OK right
A: I haven’t heard, I haven’t heard any complaint about me
K: Mm yeah
A: I haven’t heard any complaint about me from my customers
K: Mm hm
A: because I’m in contact with them every day
K: OK
A: So
K: Yeah
A: I pay taxes, like you
K: Yeah
A: I haven’t committed any crime here. I consider myself a decent person and my husband is very decent
K: Yep, yeah
A: We were trying just, we are trying just to build a better future
K: Yeah

1015. K: Yeah, yeah, yeah. Do you, do you feel optimistic about um…what score, you need a 7 on
1016. IELTS right?
1017. A: Yep, 7.
1018. K: Yep, so...
1019. A: Because I am, I am getting points for my work- work experience.
1020. K: Yeah. Do you feel positive about your chances to get a 7?
1021. A: I, I have to.
1022. K: Yeah
1023. A: It’s, because you...Everything what happens into your life, because you create it.
1024. K: Yep
1025. A: If I think I am not able or I am not capable
1026. K: Mm
1027. A: I won’t get it.
1028. K: Yeah, yeah, yeah
1029. A: I have to brush, wash my brain
1030. K: Yep
1031. A: every morning you can do it, you will improve, you can do better. What else can I do?
1032. @@@
1033. K: Yeah, I mean this is true. But are you doing much study or?
1034. A: Yes
1035. K: Are you doing preparation?
1036. A: Yes
1037. K: What kinds of things are you doing?
1038. A: OK, I have several IELTS, IELTS books
1039. K: Mm hm
1040. A: Cambridge, I just bought 8
1041. K: OK
1042. A: They’re, from 1 to 8. I won a bid on eBay, the 8 books for $110.
1043. K: OK. Congratulations @@

1266. K: Do you have to check often the website? Because I think things change quite regularly.
1267. A: Yes
1268. K: Yeah
1269. A: My husband is in charge of checking immigration page
1270. K: Mm hm
1271. A: We have tasks. We have allocated tasks
1272. K: OK right
1273. A: So he has to
K: Sharing responsibilities
A: And, and, and I don't do it because I am not able to
K: Yep
A: it's because I am afraid of the changes
K: Yeah
A: So I can't handle it
K: Yeah
A: Every time that I open that page I feel...
K: Yeah
A: Ah so he's more, he's stronger than I.
K: OK
A: So he [XXX] @@
K: Yeah so he can take the bad news and then break it to you gently?
A: Yeah @@
K: Yeah for sure
A: Yeah
K: Do you feel affected by the uncertainty?
A: Of course.
K: Yep, Is it something that weighs on you or?
A: Yes
K: Yeah
A: It's um, because [name redacted] depends on my results
K: Mm
A: and we have built a life here
K: Yeah
A: and now my residency depends on that exam...
K: Yeah, it's pretty...
A: Um but if you want to get good score you have to apply all your skills
K: Mm hm
A: it's not about English skills
K: OK
A: It's about attitude, you have to be very positive
K: Yep
A: you have to practice
K: Yeah
A: but the, but you have to think, you must start thinking that you are able to pass the exam
K: Yeah
A: You have to encourage yourself, you have to be your own cheer, cheer leader.
K: @@ that's cute
A: @@ I..
K: Yeah, yeah
A: But I'm afraid of
K: You are?
A: I'm shaking. Yes, I'm shaking
K: Mm
A: I'm very worried
K: Yeah OK
A: But life goes on, and, and you have to...
K: Yeah
A: You have to do your best.
K: Do you, have you booked a date for the test?
A: Mm no, I filled the form out today
APPENDIX 2: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (ERFAN)

89. K: Yeah. You were saying that you feel like you're alone here, do you mean you don't have a partner here or you have, it's hard to meet people in general or to make friends in general.
90. E: Yeah, I guess so. Actually it's, yeah it also depends, for me I believe, actually I feel, I felt two hurdle for meeting friends, meeting other people and making friends.
91. K: Mm hm
92. E: The very first one for me was financial situation because me having friends costs you more than not having friends, obviously. So, and the other part was obviously the language barrier as well.
93. K: OK so the language barrier...I mean, it's fairly easy for us to communicate here, have you improved, do you think, since you've been here?
94. E: Ah, yeah but yeah, obviously
95. K: Yep, OK
96. E: but I still don't consider it to be sufficient to meet whenever I want and get whatever job I want.
97. K: OK
98. E: I still feel pressure and not comfortable in conversation.
99. K: Right, right
100. E: Yeah
101. K: Do you mean you don't feel confident about approaching people that you don't know?
102. E: Um, yeah.
103. K: Is that what you mean? Just instigating, starting up conversations?
104. E: Yeah. Yeah. And actually because since I didn't have any opportunity to practice speaking
105. K: Mm hm
106. E: So yeah, so all you're hearing is just improvised @@@
107. K: Has not been rehearsed previously @@@
108. E: @@@ Yeah
280. K: OK. So, just a bit about the language related to this. What about, so you scored 6 on IELTS before you started and then you did this 10 week course
281. E: Yeah
282. K: Did that make you feel equipped?
283. E: No. Not at all. Apart from the fun that I had in the class it was absolutely useless.
284. K: So the 10 week course you didn't achieve...
285. E: No, nothing
286. K: You didn't feel like you went forward. But what about the test itself? Did you feel like it's a good measure of whether or not you're ready to start university?
287. E: Um...actually it's not a very bad test but I think that the thing about IELTS is it's varies from one exam to another exam, it varies dramatically. I mean, in one exam you feel very, very confident and very on top of the material
288. K: Mm hm
289. E: and you feel very relaxed and the next exam you don't feel nothing like that.
290. K: Mm
291. E: I don't know, and that happened in, will happen in every section, it's not just a problem for the writing or reading.
292. K: Mm
293. E: It depends on the subject of the reading and the, I don't know.
294. K: Yeah
300. E: To me it's not a very standard test. Yeah.
301. K: OK, and how many times have you done the test now?
302. E: Twice.
303. K: When was the last time you did it?
304. E: Ah I did it in February, 2011
K: Uh ha, last year
E: and I got 6.5.

K: What's the score that you need?
E: Actually I am aiming for 8 but I'm OK with 7
K: Yep
E: actually I think I’m OK with 7.
K: All right. So why are you aiming for 8?
E: Because to get 7 at least @@@
K: @@@ You don’t have any specific goal attached to an 8 though?
E: No, not specifically 8, yeah. Because of my work, work experience I think I’m gonna be, I’m able to manage with 7
K: Right
E: but it’s going to be very tight.
K: So you feel like an 8 is going to be safer?
E: Yeah.
K: Yeah
E: Much more
K: How do you aim for an 8? What do you have to do?
E: Enrolling in that class was I think was a very first approach
K: Mm hm
E: Apart from that… I think 8 is very unrealistic target because it’s very high, the difference between 7 and 8 is very high. I bet a lot of ordinary Australian people are not going to make that score in writing and reading yeah specifically, especially.
K: OK
E: Yeah. They might be able to get 9 in the listening and reading..er speaking, yeah, but yeah, particularly in writing, yeah. I don’t know why they require such a big, such a high band. It’s unrealistic.
K: Do you think there is any reason behind it?
E: Ah, I don’t know, I just see it as a very fine filter. I don’t want to think about it in this way I don’t want to think about it in this way but because of, it gets in my mental way
K: Yep
E: yeah, but yeah, to me it works, it’s just like a filter because yeah.
K: So what’s, what’s the filter? What’s being filtered?
E: The number of people, the number of applicants. Because apparently they think they can’t say that we don’t like people I don’t know, older than 40 or from these countries or whatever
K: Mm hm
E: so even though they might, they may potentially say in the future @@
K: Yeah, who knows @@
K: OK, so what kind of people then do you think…If it’s a filter to let fewer people through, OK, what are the sorts of people who will get this score?
E: Hmm, I don’t know the people which had a lot of…Me, for example, I didn’t have any exposure to English, I live in a country, an extremely isolated country which even today having satellite receiver is forbidden
K: Right
E: and there is and there isn’t any international visitor in sight, and because of the difference in the currency people can’t easily travel
K: OK
E: so myself didn’t have any exposure to English at all before coming here.
K: Mm hm
E: So, people who have similar situation like in Australia, like what it is in Australia back in their home has very, has biggest opportunities, yeah.
K: Mm
E: Yeah
K: When are you planning to do the test?
E: About...by the end of October. I haven’t, I haven’t registered yet but I’m going to.
K: Oh yeah, that’s what you’re thinking. Do you expect, I mean, this might be a stupid question
as well, but do you think that you'll pass the test, you'll get the score you need next time that
you sit it?
E: No, no
K: No?
E: 100% not.
K: You feel sure that you won't?
E: Yeah.
K: So what will you do after that?
E: Um, register again
K: OK
E: Every month. It's the only way to, it's the only thing I can do.
K: Right. Do you get advice from anyone about this, or do you hear what's going on from people
in your course or...
E: No, no
K: OK

K: So, just to understand the process that you have to go through...To put in that expression of
interest do you need to have this IELTS score?
E: Yeah, yeah, you have to have all the documents ready.
K: So you need to have your application
E: Yeah
K: all the points you need prepared?
E: Yeah, yeah
K: What do you think about that?
E: About what?
K: About the fact that you have to do the test, achieve these things and then put the
application, put in the expression.
E: Actually yeah, actually I don't want to think about it
K: Yeah
E: It is a thing that I have to do and I have no way to escape so I don't want to mess my mind
around with that
K: Yeah
E: I don't want to blame anyone
K: No, all right.
E: It's not very good, it's very difficult and you know the whole my life depends on it
K: Yeah
E: I have plan for doing PhD back in 2006 but I still haven't been able to achieve that so this
point by itself is a pain.
K: Yep
E: Yeah
K: Is it something that you feel anxious about often?
E: Yeah.
K: Do you think about it much?
E: Yeah, of course. Yeah, yeah, why wouldn't I?
K: Yeah @
E: @ yeah
K: Do you have anyone to talk to? Do you talk to your sister or anyone?
E: No.
K: What's your sister's situation? Is she...?
E: She's just graduated her Master, yeah actually.
K: Is she also planning to try to get permanent residency in Australia?
E: Yeah, actually, yeah actually she started medicine back in Iran but here they didn't accept her
degree and she is preparing for another test for ah for in order to be able to, how can I say, to
practice medicine.
APPENDIX 3: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (SHARON)

61. S: Yeah, so um I've done this um preparation course
62. K: Uh ha where did you-
63. S: for the IELTS here and
64. K: Where did you do the preparation course?
65. S: That was at [college name redacted].
66. K: OK
67. S: Um and I noticed that the reason that their essay writing style is different
68. K: Mm hm
69. S: like it's very specific but then I think-
70. K: Different to what you've learnt in school?
71. S: Yes, yes.
72. K: Well, you've been to university here as well anyway
73. S: Yeah, yep, so um and I've done a thesis and all that from my honours so
74. K: Yeah, so when did you, so you last did the test w- after you'd finished your studies?
75. S: Yes, yes
76. K: So that must be recently?
77. S: That was recently
78. K: yeah
79. S: Yeah, this year
80. K: Oh OK, this year
81. S: Yeah
82. K: Uh ha, and so when, you did you do the preparation course right before you sat the test again?
83. S: I- no, after
84. K: After
85. S: because I want to redo it again
86. K: Yep
87. S: so that I can get 8 in every band.
88. K: Mm hm
89. S: Yeah

112. K: so you were born in Sri Lanka?
113. S: Yes
114. K: And had you spent all your life there before you came here?
115. S: Yes, yes
116. K: Yep. So did you go to s- did you ha-, get taught English in school or speak English in
117. school
118. or...?
119. S: U:m, we spoke English at home since I was very little
120. K: OK
121. S: and then I went into an English speaking school
122. K: right
123. S: um though I studied in my mother tongue
124. K: Mm hm
125. S: but I was, I did English literature and all that sort of thing
126. K: At school?
127. S: Um so I was taught English from a very young age
128. K: Yep
129. S: and then I did elocution, like speech and drama
130. K: Mm hm
131. S: as well, so...yep
132. K: This was all in Sri Lanka?
133. S: Mm hm
134. K: Yep, OK and so what other languages do you speak?
S: Um I speak Sinhala, that's my mother tongue
K: Yep
S: but yeah, I think I'm more better in my grammar in English than in Sinhala @@@
K: Oh right, really @@@
S: Yeah
K: Do you speak Sinhala much?
S: Um, yes occasionally when I speak back home to family
K: OK
S: Yeah with my grandparents
K: So to your parents, or...
S: Um, mostly with my grandparents, but, um once in a while I do speak in Sinhala as well to
my parents, yeah
K: To your parents.

K: Is there, would there be any way for you to avoid doing the test? Is there ano-, another way to get enough points?
S: No
K: It's just necessary?
S: Yeah
K: Yeah
S: Mm
K: even working full time?
S: Yeah
K: If you work full time for two years here
S: Yeah
K: Mm, that's really tough I reckon
S: Yeah it is @
K: Yeah, and it
S: it's pretty frustrating when I mean, yeah I can sort of communicate the whole point is to
be able to communicate in English properly
K: Yeah
S: and I don't have a communication problem as such, so
K: and is, English is your, would you say that it's your first language, or bi, bilingual, or a
second language, if you started spea-
S: Um
K: speaking it with your parents before you started school
S: Yeah it al, it's almost the same like I use English more often than I would use my mother
tongue, so
K: Mm
S: Yeah
K: Yeah, it's interesting

K: And what do you think about um what do you think just about the fact
that you need to do an, a language test like this in order to get your visa?
S: Mm I think it's a good thing to do
K: Mm hm
S: because that way, I mean I've got a few friends, they're from um Asian countries
K: Yep
S: they've been studying here, I mean, they've finished their degree but unless you know
them very well, it's very difficult to converse with them simply because their English is not too
good at all
K: Right
S: So, it's a good way of gauging somebody's ability to communicate, but what I think is you
shouldn't have a very high score, because after all, um, English is not the mother tongue of
most of us
K: Mm hm
S: so it's not fair to think us of being of very, very good high standard as maybe a native
speaker would
K: Mm hm
S: so um, and plus when I went for the preparation class, the lecturers, the teachers rather,
S: um who was, who were taking the classes were telling us they themselves have an issue with
S: the test because um of the time pressure
K: Right
S: like, you have to do everything very very quickly
K: Yeah
S: so and even like they were saying even if they were going to sit for it I don't think they'll get
S: a very high score either simply because of that fact
K: The teachers said that?
S: Yeah
K: Yeah, OK
APPENDIX 4: SAMPLE INTERVIEW 1 TRANSCRIPT SEGMENTS (YUKI)

1. K: What I just want to find out is anything really that you've experienced you know in relation to coming here and taking tests and what your goals are and things like that, so maybe we can start off just telling me about your background, like where you're from and when you came to Australia and why you decided to come and things like that, so just talk.

2. Y: It's about 3 years ago ah now it's about 4 now, 4 years ago. Um, before 4 years ago I was doing nursing in Japan, and then um because I wanted to be a nurse in other countries, so I went to England and then came back to Melbourne actually. Yeah and then um, so at the time I was studying medical terminology to be a nurse here, like a- other English test, but that was really hard so I was taking um IELTS as well, and because at the time my English wasn't really good, so I couldn't speak properly, so yeah, that's why I couldn't pass and then so I kept taking English test and then studying about those things. Um..

3. K: Is this since you arrived in Australia or did you do this in England as well?

4. Y: Yeah, since. Ah, no, I didn't study English test in England so. Yeah, it started since I came here.

5. K: OK, so 4 years ago?

6. Y: Yeah, about 4 years ago. And then I decided, because I couldn't pass English test for registration as a nurse, so um I decided to go to uni. So, um I changed my goal to get into the university, not going to register as a nurse. So the first goal become going to uni, er but the IELTS still quite hard for me. While I was studying in the same school as my friend and they already gave up to go to university or be nurse here.

7. K: So, do you mean in school here?

8. Y: Yeah, in language school.

9. K: Oh so you were in…What school were you at, do you remember?

10. Y:  Um, that's [college name redacted] and I yeah...

11. K: So you were just studying English?

12. Y: Yeah, so I was studying just English for about 2 years, yeah, and yeah, about 2 years, and then I just tried to get into university.

13. K: OK

14. Y: Yeah

15. K: What course did you apply for?

16. Y: Um, that was nursing, that is nursing.

17. K: OK And so the first time you took IELTS, was that back in Japan?

18. Y: Yep

19. K: What sort of...how many times have you taken...When was the first time you took IELTS?

20. Y: That was quite a long time ago, I think. It's also, it's like, yeah 3 years or 3 and a half years ago.

21. K: So, the first time you took it was after you arrived in Melbourne?

22. Y: No, before. Ah, sorry, no after, yeah after Melbourne, after studying medical terminology for one year.

23. K: OK And so the first time you took IELTS, was that back in Japan?

24. Y: Yep

25. K: Do you remember what your scores were then, roughly?

26. Y: Ah, I remember overall was like 6 or something like that, but because my writing is not really good so I um, but I was I being in, you know, speaking English speaking country, so my speaking wasn't too bad, and then listening either, so um yeah but writing just reduced my overall score.

27. K: OK right OK yeah, so what was your goal when you took the test in Japan?

28. Y: At the time? Because I was thinking to go to uni so that was a 6.5, yeah

29. K: OK, so then what did you do?

30. Y: Then, um, so I came back to Melbourne as a working holiday, and I was thinking to go to uni while my working holiday is in um, how do you say? Ah, useful? How do you say? It's like a...

31. K: OK, so while it was still valid?

32. M: Valid, yeah, that word, I was thinking that word. Yeah, so yeah but because I couldn't reach the 6.5 so I was just spending time studying English and then just working full time.

33. K: Where were you working while you were here then?
Y: I was working... At the time I was working in a Japanese restaurant.
K: In a restaurant, oh OK. How long did you work there?
Y: Um, I think about half a year.
K: And during that year did you take the test again while you were in Melbourne, the IELTS test?
Y: Yeah, yeah
K: How many times did you take it?
Y: Ah nearly 10 times, I think
K: Oh, in the one year?
Y: Yeah, I took so many times just, just, because I'm... I was studying um IELTS practice, test
practice, but I thought if I just have a little bit of break, I just don't remember or I'm not really get,
I'm not really familiar with the format of the exam, so I just thought it's a good idea to take
continuously, if, until I can reach the certain level I need, so yeah.
K: OK so what did you... Why did you think that was a good idea? What did you think you'd get out of
it, by doing it over and over and over like that in a short space of time?
Y: Because, I getting familiar with um the test
K: What scores have you got in the past in speaking and reading?
Y: Um I got speaking was 6.5 at the time as well, and speaking was 7.0, yeah, because I think
I'm using English in daily life here, so speaking sometime you know... it depends on the topic of the
exam though but um yeah speaking is not really huge problem for me, yeah.
K: Yep, so why do you think you got a lower score on speaking last time?
Y: I think because the topic was in um IT, and I'm not really IT person @@@ I have no idea about
what I can talk to the um examiner, so I was quite um not really active speaking so, yeah, so that's
why I think um, because our conversation doesn't expand if you are not sure about the topic, so you
know because um I, I understand if your speaking is good score you should speak more like um
debate type of speaking. So yeah if you are sure about the topic you can talk a lot and you can even
debate about you know the opposite opinion of that, but if you're not sure about the area that the
topic is, so...
K: So do you just get one topic in the speaking test?
Y: Um no also I think ah PE physical education, like a...
K: That was the other topic?
Y: Yeah that was the other topic. I think that was the topic for debate. Because I'm not really active
person and I'm not really playing sports or something like that so um... I was talking as much as
possible, as I can you know um show my ability of speaking but I couldn't expand much about topic
because I yeah just didn't have you know...
K: OK, so you don't think that you need to have to pass all the four sections in the one time?
Y: Yeah, cos it depends on the topic, and yeah, like the same as the OET, it really depends on the
topic and it depends on your condition, like your physical condition as well. Like this time I got cold
before the test and then I feel a bit not really prepared but because I couldn't change, I couldn't
cancel it, because so you have to take it, so it depends on the situation.
K: Yeah, so for you personally now right, do you feel like you're ready, like you have the language,
the English you need to be able to work as a nurse here?

Y: Um, I need a bit of ah practice a bit more, yeah, especially speaking, because I’m somet-, I still stuck. I’ve been speaking English about nearly 4 years now, over, I still stuck what I want to say to them or some unfamiliar situation, I can’t explain it, so yeah, I need it, I need practice but, yeah.

K: Yeah. Do you think, so is having to do the test helping you improve your English?

Y: Yeah IELTS is quite helping I think because it’s more general topics so you can talk-, you can- how can I say, get used to talking Engl- speaking English in general situation, not like a medical situation, so you can, how can I say, more? You can know how to speak with a patient, not other nurses or not doctors or not medical staff, so that helps me I think, yeah.

K: OK. What about the OET?

Y: OET...Um, speaking is maybe helpful, because um it’s more like a real situation in a hospital so it’s just explain to something to the patient or um persuading him or her to be doing something in a situation so that’s helpful but I’m not sure about the um writing, because I haven’t seen other nurses writing about a referral letter, maybe they typing, so you don’t need to be very strict about the spelling, you know.

K: Do you get to have a social life as well?

Y: Um sometimes I need it but it’s really hard to have one, and so, yeah when I was doing um, preparing for OET during my clinical placement that was really hard and then after OET test I got sick

K: Yeah right

Y: because um, yeah I woke up at 6, no 5.30 for clinical placement and then came home and studied for OET and then went be- went to bed at midnight, and then woke up again 5 o’clock and doing placement and then working all weekend, so yeah.

K: Yep, yeah. Sounds great

Y: It’s very interesting. And also I had two assignments at the time, so yeah I did assignment as well so we didn’t have time, so yeah we didn’t, I didn’t have time for myself. Even going to like a shop, to get food or something, was really hard to get the time, so...

K: Yeah it doesn’t sound like much fun at all actually

Y: So I just decided to do not doing too much for me, because after that I can’t do anything, I feel just sick and then just lasts long time and healing slowly, you know, takes so long time to heal.

K: Yeah I know what you mean

Y: So I just though, good idea not doing too much.

K: Yep, so no more tests for a month. It doesn’t seem like much time off really.

Y: No, but I have a yeah assignments so I have to do that, but I think it’s just part of the like practice for English test, because I still use English for you know assignment, or school, or something.

K: So you feel like that gives you practice for the test?

Y: Yeah, so I’m thinking like that way. Because even just the normal daily usage of English just, is still part of practice of my English test so you know.

K: Yep, so I think, so just one more thing OK and then I’ll let you go I don’t know what time it is but...

Y: Yep

K: Um I think, so you’re doing all this practice and all this preparation, do you feel like your scores are improving, to reflect that?

Y: Um, um...

K: Is it, is something happening, is something changing in the test taking? Are you getting a better score or do you feel like it’s easier?

Y: Um, not getting better score though, but it really depends on topic, so it’s kind of luck, but um I feel more relaxed when I take, when I’m taking that exam, so you know I don’t get easily nervous or something like, so that’s the benefit for me I think. Because if you are very nervous you just miss some, especially listening if you’re really nervous you just miss some words to feel and then just panic and then just other question is blank or something like that so I don’t want to get those you know situations, yeah. That is the benefit for me to keep doing it

K: Yeah keep doing the test practice

Y: Yeah but it costs a lot, so.

K: When are you graduating? The end of this year?

M: Yep, the end of this year.
Author/s: 
Frost, Kellie

Title: 
Test impact as dynamic process: individual experiences of the English test requirements for permanent skilled migration in Australia

Date: 
2017

Persistent Link: 
http://hdl.handle.net/11343/194277

File Description: 
Test impact as dynamic process: individual experiences of the English test requirements for permanent skilled migration in Australia

Terms and Conditions: 
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.