Assessment Policy Change in Relation to English Language Teaching and Learning in China: A Study of Perspectives from Two Universities

Qiuxian Chen
BA (English Language Education)
MA (English Language and Literature)

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Centre for Learning Innovation
Faculty of Education
Queensland University of Technology
April 2011
Keywords

Formative assessment, high-stakes external test, English Language Learning, English language assessment, assessment policy change, sociocultural perspective, Chinese higher education, community of practice
Abstract

In response to concerns about the quality of English Language Learning (ELL) education at tertiary level, the Chinese Ministry of Education (CMoE) launched the College English Reform Program (CERP) in 2004. By means of a press release (CMoE, 2005) and a guideline document titled *College English Curriculum Requirements* (CECR) (CMoE, 2007), the CERP proposed two major changes to the College English assessment policy, which were: (1) the shift to optional status for the compulsory external test, the College English Test Band 4 (CET4); and (2) the incorporation of formative assessment into the existing summative assessment framework.

This study investigated the interactions between the College English assessment policy change, the theoretical underpinnings, and the assessment practices within two Chinese universities (one Key University and one Non-Key University). It adopted a sociocultural theoretical perspective to examine the implementation process as experienced by local actors of institutional and classroom levels.

Systematic data analysis using a constant comparative method (Merriam, 1998) revealed that contextual factors and implementation issues did not lead to significant differences in the two cases. Lack of training in assessment and the sociocultural factors such as the traditional emphasis on the product of learning and hierarchical teacher/students relationship are decisive and responsible for the limited effect of the reform.
# Table of Contents

Keywords .............................................................................................................................................. i  
Abstract ............................................................................................................................................... iii  
Table of Contents ................................................................................................................................. v  
List of Figures ......................................................................................................................................... ix  
List of Tables ......................................................................................................................................... x  
List of Tables ......................................................................................................................................... x  
List of Abbreviations ............................................................................................................................ xi  
Statement of Original Authorship ......................................................................................................... xii  
Acknowledgements ............................................................................................................................... xiii  
Glossary .................................................................................................................................................. xiv  

## CHAPTER 1: INTRODUCTION ........................................................................................................... 1  
1.1 An Assessment Policy Change in a Changing Climate ................................................................. 1  
1.2 Motivation for this Study .............................................................................................................. 3  
1.3 Key Concepts ............................................................................................................................... 3  
1.3.1 Testing and assessment ........................................................................................................ 3  
1.3.2 Alternative assessment and alternatives in assessment ...................................................... 5  
1.3.3 Backwash/washback and impact ......................................................................................... 5  
1.4 Background .................................................................................................................................. 6  
1.4.1 ELL education in an era of globalisation ............................................................................ 6  
1.4.2 EFL education in China, a political and economic agenda ................................................ 7  
1.4.3 Tertiary EFL education in a changed context .................................................................... 9  
1.5 Context ......................................................................................................................................... 12  
1.5.1 The College English assessment system .......................................................................... 12  
1.5.2 The role of CET testing system in College English education, a historical perspective ...... 13  
1.5.3 The College English Reform Program (CERP) ................................................................. 17  
1.5.4 The CERP and the new College English assessment policy ............................................ 18  
1.6 Research Gap ............................................................................................................................... 21  
1.7 Research Questions .................................................................................................................... 22  
1.8 Significance of the Research ....................................................................................................... 23  
1.9 Thesis Outline .............................................................................................................................. 24  

## CHAPTER 2: LITERATURE REVIEW ............................................................................................. 27  
2.1 Theoretical Framework ................................................................................................................ 27  
2.1.1 An interpretativist epistemology .................................................................................... 27  
2.1.2 Sociocultural theory and this study ................................................................................. 27  
2.2 Assessment Transformation ......................................................................................................... 32  
2.2.1 Changing views and paradigms ....................................................................................... 32  
2.2.2 Formative and summative assessment, a dichotomy or continuum? .............................. 34  
2.3 The Evolution of English Language Learning ......................................................................... 40  
2.3.1 Second Language Learning (SLL) and Second Language Acquisition (SLA) ................ 40  
2.3.2 Second Language Learning, a historical perspective ..................................................... 40  
2.4 Assessment in English Language Learning ............................................................................. 44  
2.4.1 English language assessment as a discipline ................................................................. 44
Assessment Policy Change in Relation to English Language Teaching and Learning in China

CHAPTER 3: RESEARCH DESIGN

3.1 Sociocultural Approach as Research Theory
   3.1.1 Sociocultural approach as a research theory
   3.1.2 Sociocultural informed research design

3.2 Case Study Approach

3.3 Selection of Cases and Participants
   3.3.1 Case definition and selection
   3.3.2 Senior administrator selection
   3.3.3 Teacher selection
   3.3.4 Student selection

3.4 Data Collection Methods and Tools Construction
   3.4.1 Interview schedule for the senior administrator
   3.4.2 Interview schedule for the teacher
   3.4.3 Interview schedule for the students
   3.4.4 Classroom observation

3.5 Piloting
   3.5.1 Teacher interview schedule
   3.5.2 Student interview schedule
   3.5.3 Classroom observation

3.6 Main Study
   3.6.1 NKU
   3.6.2 KU

3.7 Validity and Reliability Issues

3.8 Analysis
   3.8.1 Sociocultural perspective & community of practice
   3.8.2 Researcher’s self
   3.8.3 Coding & categorising using NVivo 8
   3.8.4 Ethics

CHAPTER 4: CASE STUDY ONE: KEY UNIVERSITY (KU)

4.1 Institutional Level Responses
   4.1.1 Profile
   4.1.2 Changes to the institutional assessment policy
6.5.1 Experience of process assessment .......................................................... 220
6.5.2 Formative assessment ........................................................................ 221
6.5.3 Responses to CET4 .............................................................................. 224

6.6 Discussion ............................................................................................... 226

6.7 Summary .................................................................................................. 228

CHAPTER 7: DISCUSSION .................................................................................. 231

7.1 Communities of Practice ......................................................................... 231
7.2 Negotiated Meaning of Process Assessment ............................................. 234

7.3 Factors Mediating Teachers’ Negotiation of Meaning ............................. 237
  7.3.1 Process assessment policy ..................................................................... 240
  7.3.2 Professional training in assessment ........................................................ 240
  7.3.3 Teacher’s prior experience ................................................................. 242
  7.3.4 A demand for objectivity in assessment practices .................................. 243
  7.3.5 Teacher/student relationship .............................................................. 245

7.4 Factors Mediating Students’ Negotiation of Meaning ............................. 247
  7.4.1 Students’ prior experience ................................................................. 248
  7.4.2 Focus on summative outcomes ............................................................ 249
  7.4.3 Uses and Consequences ..................................................................... 250
  7.4.4 Developing a new identity as learners .................................................... 251

7.5 Summary .................................................................................................. 252

CHAPTER 8: IMPLICATIONS, LIMITATIONS AND CONCLUSION .................. 255

8.1 An Overview of the Study ......................................................................... 255

8.2 Implications ............................................................................................... 257
  8.2.1 National level policy ........................................................................... 257
  8.2.2 Institutional level policy ....................................................................... 259
  8.2.3 Teachers and teaching practices ............................................................ 261
  8.2.4 Students and English Language Learning (ELL) practice .................... 263

8.3 Significance ............................................................................................... 265

8.4 Limitations ............................................................................................... 267

8.5 Directions for Further Research ................................................................. 267

8.6 Concluding Personal reflections ................................................................. 268

REFERENCES .................................................................................................. 271

APPENDICES ................................................................................................... 299
  Appendix A Senior administrator interview schedule ..................................... 299
  Appendix B Teacher Interview Schedule ......................................................... 300
  Appendix C Student Interview Schedule ........................................................ 303
  Appendix D Classroom Observation Schedule .............................................. 305
  Appendix E KU Teachers’ Assessment Procedures ......................................... 306
  Appendix F KU-T1 Quiz Sample .................................................................... 307
  Appendix G KU Student Quiz Paper Sample ................................................ 308
  Appendix H KU-T1’s Record Sheet ................................................................. 309
  Appendix I NKU Teachers’ Assessment Procedures ......................................... 310
  Appendix J NKU-T2’s Record Sheet ................................................................. 311
  Appendix K NKU Student Writing Assignment Sample .................................. 311
  Appendix L Ethics approval from QUT Research Ethic Unit ............................ 313
  Appendix M Information sheet for teacher interview participants .................. 314
  Appendix N Consent form for teacher interview participants .......................... 316
  Appendix O Information sheet for student participants .................................... 317
  Appendix P Consent form for student interview participants ............................ 319
List of Figures

Figure 1.1. Assessment, measurement and testing .................................................................4
Figure 1.2. Three Concentric Circles ....................................................................................6
Figure 1.3. Student population increase in Chinese higher education .................................11
Figure 2.1. Learning theory continuum ...............................................................................28
Figure 2.2. Changing conceptions of learning, curriculum and measurement .....................32
Figure 2.3. A continuum of assessment purposes and practices ..........................................37
Figure 2.4. Assessment for learning in the classroom: A typology of possibilities .............47
Figure 2.5. Components of learning-oriented assessment .....................................................47
Figure 4.1. Changes to grade distribution .............................................................................110
Figure 4.2. College English internal assessment framework ..............................................111
Figure 4.3. KU-T1’s assessment arrangements .....................................................................123
Figure 5.1. NKU College English internal assessment framework ......................................154
Figure 7.1. Dimensions of practice as the property of a community .....................................232
Figure 7.2. KU and NKU as situated in communities of practice .........................................233
Figure 7.3. Process assessment as positioned in the assessment continuum .......................237
List of Tables

Table 2.1 Two Assessment Cultures.........................................................................................................................46
Table 2.2 The Contrast Between the Advocated Principles and Reality of the Local Context.................63
Table 3.1 Link between Research Questions, Data Sources and Data Collecting Methods ..................85
Table 3.2 The Piloting.................................................................................................................................................91
Table 3.3 Demographic Information of the Students Interviewed at NKU .................................................92
Table 3.4 Demographic Information of the Focus Group Interviewees at NKU.........................................93
Table 3.5 Timeline of Data Collection Procedures..............................................................................................95
Table 4.1 A Comparison of Assessment Policy at National and Institutional Levels...............................112
Table 5.1 A Comparison of Assessment Policy at National and Institutional Levels.............................155
Table 6.1 External Test (CET4) Initiative Being Approached at the Two Local Contexts .....................228
Table 6.2 Approach to Formative Assessment in the Two Local Contexts .................................................229
Table 7.1 KU and NKU Comparison .................................................................................................................235
Table 7.2 Contextual Factors Influencing Teachers’ Assessment Practices (differences) ....................238
Table 7.3 Contextual factors that influence teachers’ assessment practice (similarities) .......................239
Table 7.4 Contextual Factors Influencing Students’ Responses to assessment (differences) .............247
Table 7.5 Contextual Factors Influencing Students’ Responses to Assessment (similarities) .............248
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>Applied Linguistics</td>
</tr>
<tr>
<td>CE</td>
<td>College English</td>
</tr>
<tr>
<td>CECR</td>
<td>College English Curriculum Requirements</td>
</tr>
<tr>
<td>CERP</td>
<td>College English Reform Program</td>
</tr>
<tr>
<td>CET</td>
<td>College English Test</td>
</tr>
<tr>
<td>CHC</td>
<td>Confucian Heritage Culture</td>
</tr>
<tr>
<td>CLT</td>
<td>Communicative Language Teaching</td>
</tr>
<tr>
<td>CMoE</td>
<td>China Ministry of Education</td>
</tr>
<tr>
<td>EEHE</td>
<td>Entrance Examination of Higher Education</td>
</tr>
<tr>
<td>ELL</td>
<td>English Language Learning</td>
</tr>
<tr>
<td>EFL</td>
<td>English as a Foreign Language</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a Second language</td>
</tr>
<tr>
<td>HEC</td>
<td>Higher Education Commission</td>
</tr>
<tr>
<td>HKCE</td>
<td>Hong Kong Certificate of Education</td>
</tr>
<tr>
<td>IELTS</td>
<td>International English Language Testing System</td>
</tr>
<tr>
<td>MCALL</td>
<td>Multimedia Computer Assisted Language Learning</td>
</tr>
<tr>
<td>MCQ</td>
<td>Multiple Choice Questions</td>
</tr>
<tr>
<td>NCETC</td>
<td>National College English Test Committee</td>
</tr>
<tr>
<td>NMET</td>
<td>National Matriculation English Test</td>
</tr>
<tr>
<td>RWT</td>
<td>Reading, Writing and Translating</td>
</tr>
<tr>
<td>SFL</td>
<td>Systemic Functional Linguistics</td>
</tr>
<tr>
<td>SILL</td>
<td>Strategy Inventory for Language Learning</td>
</tr>
<tr>
<td>SL</td>
<td>Speaking and Listening</td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
</tr>
<tr>
<td>SLL</td>
<td>Second Language Learning</td>
</tr>
<tr>
<td>TAO</td>
<td>Teaching Administrative Office</td>
</tr>
<tr>
<td>TL</td>
<td>Target Language</td>
</tr>
<tr>
<td>TEFOL</td>
<td>Test of English as a Foreign Language</td>
</tr>
<tr>
<td>WTO</td>
<td>World Trade Organization</td>
</tr>
<tr>
<td>WWII</td>
<td>World War Two</td>
</tr>
</tbody>
</table>
Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: ______________________________

Date: ______________________________
Acknowledgements

First of all, the sponsorship by China Scholarship Council and Queensland University of Technology needs to be acknowledged. Without this financial support, it will be harder and take longer time to complete this study.

The successful completion of this study is much indebted to my three supervisors: Professor Val Klenowski, Dr. Margaret Kettle and Dr. Lyn May. They have been consistently supporting me with their professional expertise, plenty of time, and encouragement. Without them, this study would not have been possible.

I should also like to express my gratitude to the academics who have provided their advice along my PhD journey, especially the reviewers at the confirmation and final oral stages: Professor John Lidstone, Dr. Cathie Doherty, Dr. Karen Dooley, and Dr. Lenore Adie. Their constructive comments attribute to the betterment of this thesis. Among many others, Professor Tom Cooper’s and Associate Professor Annette Baturo’s professional advice had helped me to clarify my thinking when I was at loss; and Dr. Briony Wainman had supported me with her expertise in academic writing.

I also wish to extend my thanks to the participating institutions, teachers and students. It was them who made data collection fruitful and enjoyable. Special thanks go to the two deans of English Language Departments. Without your cooperation and support, this study would be impossible.

Thanks also go to my PhD candidate friends: Jen, Betty, Rene, Kathy, David and Zue. We have been sharing the ups and downs, the joys and frustrations along this long journey. Without your company and encouragement, this PhD journey would be so lonely. Thanks also go to a special friend Ros, who spent a lot of time editing this thesis.

Last but not the least, I want to thank my dear husband Qilong for his understanding and my beloved daughter Nancy for being my company, comfort and my emotional support. Their love and assistance are the very impetus for me to carry on and finish this study.
Glossary

College English  Mandatory English coursework for all non-English major undergraduate students in China’s higher education institutions, also called Public English. The course lasts for two academic years, and accounts for 16 credits for the Bachelor Degree.

College English Test (CET)  A large scale, norm-referenced, standardised testing system that higher institutions in China were used nationwide to evaluate students’ English proficiency at the end of two years’ College English study, CET4 for general level and CET 6 for advanced level.

Deaf English  A literal translation from Chinese of a phenomenon that emerged from traditional grammar and structure pedagogy and curriculum that gives priority to reading and writing over speaking and listening. The term is used to refer to the fact that Chinese EFL learners, after years of English learning, have achieved only limited language proficiency, especially in listening and speaking. Deaf English is described as ‘can read but cannot speak’ fluently or understand English speakers well.

Low efficiency and ineffectiveness  A literal translation from Chinese of the problematic phenomenon that Chinese EFL learners have developed only limited language proficiency after many years of formal English study.

National Quality Higher Education Project  A project that the Chinese government launched in 2006 to improve the quality of higher education. A total of 25 billion RMB is invested at the national level.

High-stakes assessment  Assessment with results that have serious consequences for the test-takers, for example, in relation to entry into a course, university, or degree qualification.
Chapter 1: Introduction

This chapter provides an introduction to the research inquiry into the assessment policy change that the China Ministry of Education (CMoE) initiated via the College English Reform Program (CERP) (CMoE, 2004). It starts with an introduction of the assessment policy in a changing global climate (Section 1.1). Section 1.2 states my experience as the motivation for this study. Section 1.3 clarifies the definition of key concepts. The next section (Section 1.4) presents the background information that supports the rationale for such a study: the political, social and educational background of College English education in an era of globalisation. Section 1.5 provides further contextual information about the College English education and its assessment – its developments up to the present College English Curriculum Requirements (CECR), its problems and challenges. Section 1.6 identifies the research gap that this doctoral study seeks to fill and the overarching objectives that it aims to achieve. Then, the research questions (Section 1.7) are given as is the significance (Section 1.8). The last section (Section 1.9) maps out the structure of the remaining chapters of this thesis.

1.1 AN ASSESSMENT POLICY CHANGE IN A CHANGING CLIMATE

In 2004, the China Ministry of Education [CMoE] launched the CERP, a reform program in the area of English as a Foreign Language (EFL) for non-English majors in the Chinese higher education context. Initiatives to achieve better education outcomes for a sustainable future were clearly articulated. By means of a press release (CMoE, 2005) and a guideline document, the CECR (CMoE, 2007), CERP proposed two major changes to the College English assessment policy. The changes were: (1) the shift to optional status for the compulsory external test, the College English Test Band 4 (CET4); and (2) the incorporation of formative assessment into the existing summative assessment framework. With CECR implemented fully in 2007, the influence of the assessment policy change on College English teaching and learning is a significant issue.

---

1 I took an “ecumenical” view of policy, according to which policy is defined as any “vehicle or medium for carrying and transmitting a policy message” (Ozga, 2000 p.33) and includes “a speech by a Minister for Education, a press release by a senior education bureaucrat, as well as formal official policy texts” (Lingard & Ozga, 2007 p. 2).
Assessment has been practised worldwide in the educational domains for the past century, and even longer in certain contexts. It serves the educational and social purposes of selection, certification and control (Broadfoot, 2007). However, over the years the role of assessment has witnessed significant changes with regard to the purposes it serves and the functions it fulfils (Broadfoot & Black, 2004).

In the past two decades, for example, assessment has been used as a “tool” (Hamilton, 2003) or lever in many educational settings to effect pedagogical and learning practice changes. More recently, as informed by theoretical developments in learning, notably constructivist and sociocultural perspectives (Moore, 2003; Rogoff, 1999; Vygotsky, 1978), educational assessment has undergone a “conceptual shift which carries with it considerable practical consequences” (Dwyer, 1998 p. 131). Among its multiple consequences, the roles that assessment play and the purposes that assessment serves have undergone another major shift in emphasis.

This shift in emphasis, captured in the concept of “assessment for learning” (Assessment Reform Group [ARG], 1999, 2002) in England, Wales, Scotland, Australia, and New Zealand, highlights the role of assessment in “supporting and enhancing learning” (Black & Wiliam, 1998; Black, Harrison, Lee, Marshall, & Wiliam, 2003). Hence, in both research and policy a tendency is seen to seek “alternative assessment” (Balliro, 1993; Huerta-Macias, 1995) in the service of learning. This shift foregrounds assessment approaches for formative purposes (Ross, 2005). In contrast to the prevailing assessment approach for summative purposes, assessment approaches for formative purposes give more emphasis to process, learner engagement, and learning improvement rather than evaluation of product (Harlen, 2006). These shifts have become the impetus for changes in assessment practices and policies (Broadfoot & Black 2004; Broadfoot, 2007) in general education as well as in the area of ELL (Lynch, 2003; Rea-Dickins, 2004; Ross, 2005).

However, in the area of ELL, there is an acknowledged “dearth of information in regard to the impact of these innovations on teaching and learning” (Davison & Cummins, 2007 p. 415) in diverse contexts and even less in Confucian Heritage Culture based Asian contexts (Carless, 2010). Research into formative assessment related reforms and their impact is significant to ELL.
1.2 MOTIVATION FOR THIS STUDY

My personal experience as an EFL learner and EFL teacher in China motivated my interest in assessment research, especially changes in assessment for learning purposes.

I started EFL from junior high. After two high-stakes standardised tests – the Senior High Entrance Examination (SHEE) in 1987 and the national Entrance Examination to Higher Education (EEHE) in 1990, I went to college and took English as a major for my Bachelor degree. I found English-mediated instruction difficult to follow and began to reflect on the testing systems that I experienced. Six years’ English learning experience and countless tests and examinations did not result in my becoming a fluent English speaker. Passing the high stakes tests did not mean that I had communicative competence in English.

On graduation I began to teach EFL in universities for non-English major students. I was not surprised to find my students in a similar situation. I tried to help my students to pass the then mandated CET4, one of the requirements of the Bachelor degree, only to hear them complaining later to me: “Miss Chen, how come our English is so limited when we want to put it to use?” I knew the answer, but I did not know how to reply. Later, I used some alternative means to support their learning of EFL. Some of my students came to me, saying: “Miss Chen, we know you mean well for us, but the fact is we just can’t afford to ignore the test.” A feeling of powerlessness overwhelmed me. My colleagues experienced the same dilemma. We agreed with Torrance that “real change will not take place in schooling until significant change happens to assessment” (1996, p. i). We waited for over a decade for significant change in College English assessment to happen in the ongoing CERP.

1.3 KEY CONCEPTS

1.3.1 Testing and assessment

Two key concepts, testing and assessment, need to be clarified for proper use in this thesis. Testing and assessment are widely used in the literature of general education and area of ELL, sharing the common purposes of describing the making of judgements and decisions about individuals’ learning. During the past decades when psychometric testing theory prevailed, assessment in the ELL domain was primarily dominated by testing (Alderson & Banerjee, 2002). The recent theoretical
development gave rise to the pursuit of assessment approaches alternative to standardised tests (Balliro, 1993; Huerta-Macias, 1995). Therefore, assessment is also used to refer to informal and classroom-based methods that are captured under the term of *alternative assessment* (Balliro, 1993; Huerta-Macias, 1995). The ambiguity in the uses of the two terms necessitates a closer look at the distinctions between the concepts.

According to Clapham (2003), *testing* is applied more often to the construction and administration of formal or standardised tests, while *assessment* has been used as an umbrella term to cover a broad range of measurement techniques. Similarly, Lynch (2001) identifies testing with the test as an instrument “producing varieties of measurement that exist for assessment purposes,” whereas assessment is the “superordinate term” (p. 371) covering a range of procedures that includes traditional quantitative measurement, testing, and other assessment approaches (as illustrated in Figure 1.1).

According to Clapham (2003), *testing* is applied more often to the construction and administration of formal or standardised tests, while *assessment* has been used as an umbrella term to cover a broad range of measurement techniques. Similarly, Lynch (2001) identifies testing with the test as an instrument “producing varieties of measurement that exist for assessment purposes,” whereas assessment is the “superordinate term” (p. 371) covering a range of procedures that includes traditional quantitative measurement, testing, and other assessment approaches (as illustrated in Figure 1.1).

Further, Lynch (2001) argues that the distinctions between the two terms go beyond the techniques that are used, and relate to the theoretical perspectives taken. Testing, being mostly associated with formats like multiple choice and short answer questions, is usually referred to in terms of psychometric theoretical underpinnings. In contrast, assessment is informed by a non-positivist perspective and tends to make use of non-quantitative techniques for classroom-based data collection and analysis (Lynch, 2001).

In this thesis, assessment will be used to encompass both testing and other assessment approaches, whereas testing will be mainly used to refer to traditional
standardised tests. Examination is another term that appears in this thesis. It is used to refer to the traditional pencil and paper written tests that have been used for the major part of the examination history in China as well as other settings.

1.3.2 Alternative assessment and alternatives in assessment

Alternative assessment was used in contrast to traditional assessment, typically the standardised test (Huerta-Macias, 1995). However, questions were raised as to the boundaries and connotations of this term. Balliro (1993) maintains that alternative assessment relates not only to the varieties of assessment methods other than tests, but is also associated with the alternative framework:

- Ethnographic as an alternative to the psychometric tradition; literacy-associative-practices as an alternative to literacy-as-discrete-subskills … and a preference for primarily qualitative data as an alternative to reliance on quantitative data alone. (p. 559)

This view is echoed by Brown and Hudson (1998), who view the emerging dichotomy between alternative assessment and traditional assessment as being oversimplistic and potentially problematic, and so propose the term alternatives in assessment to encompass tests, portfolios, self and peer-assessment. This view is more widely accepted and hence adopted in this thesis.

1.3.3 Backwash/washback and impact

The effects of testing on learning and teaching practices are captured in the concept of backwash in general education or washback in Applied Linguistics (Alderson 2004; Messick, 1996; Bailey, 1996). They are used interchangeably with no significant difference. This study is situated in the ELL discipline, so washback is used. But backwash will be retained when citations apply.

Wall (1997) recommended making distinctions between backwash/washback and impact. The latter term refers to the effects of assessment not only on classroom practice but more broadly on the educational system and society. This study is concerned with the changes to College English assessment policy and the effects of different modes of assessment on classroom instruction and learning practices. I adopt Andrews’ (2004) broad interpretation of washback. Impact, on the other hand, covers the influence of the assessment on both “micro level of the individuals and macro level of the educational system or society” (Bachman & Palmer, 1996 p. 29).
1.4 BACKGROUND

1.4.1 ELL education in an era of globalisation

The advent of globalisation has created new political agendas for education (Dale, 1999; Ozga & Lingard, 2007; Zufiaurre, 2007). A top agenda is English language education that has been fuelled by the established status of English as an international language in academic, commercial, scientific, technological, and political domains throughout the world. Therefore, English is being learned and taught as an international language (EIL) (Warschauer, 2000) or a global language (EGL) (Crystal, 1997, 1999) in a global context. However, there is no denying that the linguistic and cultural contexts and inputs in imparting English education are diverse, which results in the differences in English language teaching and learning approaches and outcomes (Krachru, 1996). This diversity is best described in Kachru’s (1996) model of Three Concentric Circles (reproduced in Figure 1.2).

![Figure 1.2. Three Concentric Circles (Kachru, 1996, p. 137)](image)

This model is used to describe the different profiles of people using English. **Inner Circle** users are those who are native speakers and include English users in countries such as England, Canada, the United States and Australia. The **Outer Circle** countries are those in which English is used as a second language (ESL), for example, in Singapore and Hong Kong. The **Expanding Circle** countries are those in which English has no official status but is used in business, trade and so on and is taught in
schools as a foreign language (EFL), for instance, Korea, Japan and China. For those English Outer and Expanding Circle nations which are keen on gaining a competitive edge in a global forum, the status of EFL education is prominent (Ross, 2008).

China, the context of this study, is identified as one of the Expanding Circle countries. This means that English in China is a foreign language. In such a context, EFL learning is “a daunting task” (Wu, 2001) because of the lack of everyday exposure and opportunities to use the target language. Despite these difficulties, English language proficiency is considered to be vital for China’s ongoing trade, economic and political needs (Adamson & Morris, 1997; Ross, 1992). I discuss these needs in more detail below. I will use the term English Language Learning (ELL) to refer to the discipline in general terms and use English as a Second or Foreign Language (ESL/EFL) when specific context is concerned in this study.

1.4.2 EFL education in China, a political and economic agenda

The policies on EFL education in China after the establishment of the People’s Republic were largely linked to the political situation in the country (Adamson & Morris, 1997; Adamson, 2004; Ross, 1992). English gained its supremacy over Russian as the main foreign language in education after Sino-Soviet relations broke down around the 1960s (Lam, 2002). This supremacy experienced a setback in the Cultural Revolution (1966-76), when schools were closed and learning foreign languages as well as even learning were condemned. Since then, however, the policy emphasis on enhancing English in education “has continued unabated” (Lam, 2002, p. 246). With “perceived priority” (Hu, 2005) in policy, English language education in China has experienced impressive development over the past three decades (Hertling, 1996; Hu, 2005; Liu & Gong, 2000; CMoE, 2005; Wang, 2007).

The development of EFL education over the past three decades is perceived to have gone through four major phases (Q. Wang, 2007) – the Restoration Phase (1978-1985), the Rapid Development Phase (1986-1992), the Reform Phase (1993-2000) and the Innovation Phase (2001-today). These developments after the Cultural Revolution have been not only linked to the political influence but also inextricably linked to the economic and social development of the country (Hu, 2005).

The new government after the Cultural Revolution (1966-1976), led by Deng Xiaoping, perceived the vital role of English in its push for national modernization
and development (Adamson & Morris, 1997; Hu, 2005; Ross, 1992). English language education was regarded as “the barometer of modernization” (Ross, 1992, p. 239). In line with the Open Door policy in 1978, the national drive to promote English intensified (Lam, 2002). Notably, efforts were made to establish English language education in the formal education curriculum at secondary and tertiary levels. For example, in 1982, English was implemented as the main foreign language in secondary education (Zhao & Campbell, 1995; Zhou, 1995); in 1983, English became a formal requirement for admission into higher education (Q. Wang, 2007); and in 1986, College English was mandated at tertiary institutions (CMoE, 1985, 1986).

Concerted efforts from all parties concerned soon brought English language education in China into a Rapid Development Phase (1986-1992). New syllabi were issued, new textbooks published, and new pedagogies, notably Communicative Language Teaching (CLT), developed or introduced (G. Hu, 2005; Adamson, 2004; Yu, 2001). A survey conducted in 2001 by the Project Team for Research on Senior Secondary School English Language Teaching in China reveals noticeable overall improvement of students’ proficiency level. However, English language education at this phase was not without problems. Limited proficiency in speaking and listening, so called “deaf English” (refer to glossary) prevailed (Liu & Gong, 2001; Li, 1995). With the Open Door Policy opening up more opportunities for increasing international exchange in various domains, higher standards were required of English language education to meet these needs.

These requirements were further elevated by the advent of globalisation in the 1990s. The Chinese leadership recognised the role of English in its national competitiveness. Hence, strategic plans were made and implemented for the purpose of “quality education” to meet these challenges and changes (CMoE, 1998). As such, English language education in China witnessed another phase of development. On one hand, English education was expanded to earlier years. For instance, English language education began to be implemented in the primary schools throughout the country in 2001 (CMoE, 2001). On the other hand, a wave of curriculum innovations and reforms were staged at varied levels to improve the quality of English language education: the Basic Requirements for Primary School English (2001) and the New National English Curriculum Standards that unified primary and secondary English
curricula into one continuous entity in 2003, and the CECR (2007) at the tertiary level.

China’s accession to the World Trade Organization (WTO) and its successful bid for the 2008 Olympic Games in 2001 made English language education even more pressing throughout the country (Jin & Cortazzi, 2002; Pang, Zhou & Fu, 2002). The significance of English was also seen in the nation’s competitiveness in the global market and its recognised role in world affairs (Pang, Zhou & Fu, 2002). Moreover, the 2008 Olympic Games ushered in an opportunity for the country to showcase to the world. In addition, with the perceived priority given to English language education, prestige has been conferred on those with good English proficiency (Jin & Cortazzi, 2002). English proficiency is used as a gate-keeper in educational as well as other social domains (J. Wang, 2007; Jin & Cortazzi, 2002). It is regarded as a passport to academic, professional, and social advancement, as “the key to success” (Cheng, 2008) and is a capital for the betterment of personal well-being (G. Hu, 2005). These social, political, and economic factors have contributed to a rise in the prominent status of English language education throughout the country and an escalating demand for better quality English language education (Jin & Cortazzi, 2002). Thus EFL education in China has become a political and economic agenda, and a national and individual commitment (Jin & Cortazzi, 2002; Q. Wang, 2007).

This study is primarily concerned with EFL education for non-English majors in the Chinese higher education context – the so called College English or Public English. The following section will examine English language education at this level for this particular group of students, addressing its significance, problems, and the proposed solutions in the ongoing CERP.

1.4.3 Tertiary EFL education in a changed context

As elaborated in Section 1.3.2, EFL education in China emerged as a priority soon after the Cultural Revolution (1966-1976) (Adamson & Morris, 1997; Adamson, 2004; Ross, 1992). To accelerate the modernisation process under the Open Door policy, initiatives were taken in the mid 1980s to offer mandated EFL education in higher education institutions (CMoE, 1986). The emergence of College English has been a source of both interest and concern to policy-makers, researchers, and practitioners.
**College English and its objectives**

College English was introduced for non-English major undergraduate students in the mid 1980s. Currently, it is taught to more than 99% of the target group (J. Wang, 2007). The practice has been regulated through the nationally unified curriculum syllabus – initially, the *College English Teaching Syllabus* (1985/1986, 1999) and more recently, the CECR (CMoE, 2007). According to these policies, College English constitutes a compulsory course unit for four academic terms or two academic years (usually the first two years), and accounts for 16 points or 10% of the total credit points required for a Bachelor degree. Despite the shift in emphasis of specific language skills (Cai, 2007), the objective of College English is consistent in cultivating talents to meet the demands of national and societal development (CMoE, 1986, 1999, 2007).

**The significance of College English**

College English education remains a focal topic and receives a great deal of policy attention (Q. Wang, 2007; Hu, 2005). From the political perspective, the Chinese government considers higher education an important contributing force to the nation’s strength, and its quality a key index of the nation’s competitiveness in the global context of the knowledge economy (CMoE, 2002; Mok, 2005). Also, the Chinese leadership is convinced that proficiency in English is an essential component of quality education, as is clearly stated in the CERP (CMoE, 2004).

The significance of College English education can also be seen from social and economic perspectives. As the elite rung of the educational ladder in China, higher education carries expectations of students achieving high levels of English proficiency. And, as the final context of the long process of education, higher education is faced with challenges from societal demands and changed contexts (Boud, 2007). The challenges from societal demands mainly come from the increasingly global and competitive employment markets (Han, Dai & Yang, 2004; D. Wang, 2007). However, the mismatch between employers’ needs and graduates’ actual limited English proficiency levels is well documented in the literature (Shao & Huang, 1997; Zhao, 2001; D. Wang, 2007). Therefore, to improve the overall College English education outcomes is a concern for policy-makers, researchers, and practitioners (Z. Hu, 2002, Li, 1995, D. Wang; 2007).
**College English in a changed context**

In addition to societal demands, College English education in China has to face challenges arising from a changed context. The first major change relates to enrolment expansion. After restructuring the higher education system and merging universities into large units in the mid 1990s, the initiative was taken to expand the scale of higher education enrolment (CMoE, 1996; Mok, 2005). During the years between 1998 and 2006, the gross enrolment rate increased from 8-11% to 22% in the Chinese higher education sector, which resulted in the student population more than tripling (as shown in Figure 1.3). The increase continued until it slowed down in 2010 (Zhou, 2007; CMoE, 2006a).

![Figure 1.3. Student population increase in Chinese higher education (Source: Tengxun News).](image)

With such an exponential increase in number, the quality of higher education has become a troubling issue and a major concern for policy-makers and the public alike (CMoE, 2002). Therefore, in 2002, the Chinese government took another major move with the *Higher Education Quality and Reform Project* (HEQRP) to improve higher education’s overall quality (CMoE, 2002). Whereas the first move put great pressure on the College English education with the dramatically increased student population, the second move (the HEQRP) involved College English as a major part (CMoE, 2002, 2006b) of the program.

**Problems**

In sharp contrast to the policy orientation and a general desire for quality, “low efficiency and effectiveness” (literally translated from Chinese mandarin 费时低效 fèishídíxiào, means spending a lot of time, yet achieving little effect) (Dai, 2001; Jing, 1999, 2000) has been a “thorny issue that has long been identified, yet to be solved” (Hu, 2005 p. 5). The efforts and resources individually and nationally invested in formal schooling or informal training have far from paid off (Niu, 2001; Wu, 2001).
The reasons contributing to this phenomenon are multiple. Lack of target language learning contexts, insufficient numbers of qualified teaching staff, limited resources and large class sizes have all been identified (Hu, 2002; Jing, 1999, 2000). A major reason by consensus, however, is the CET4/6, the external testing system that College English has adopted (Han, 2002; Han et al., 2004; Niu, 2001; Zhao, 2001, D. Wang, 2007). The CET4/6 is considered to have had a major impact on College English language teaching and learning in China. I will discuss key aspects of the test in the context section below.

1.5 CONTEXT

1.5.1 The College English assessment system

The inception of the CET testing system in 1987 was a strategic move that was taken to reinforce the effective implementation of the 1986/1987 CET Syllabi (CET Syllabus Committee, 1987; Jin & Yang, 2006; Yang, 2003). The CET testing system embraces the standardised testing format, and is a six band test battery (Cheng, 2008; Yang, 2003). But when the CET is mentioned, people are usually referring to the CET Band 4 (CET4) and Band 6 (CET6) in particular, because only these two are nationally administered and certificated on passing. Band one to three and five (CET 1-3 and 5) are achievement tests administered at collegiate level at the end of each semester to assess students’ learning. That makes the College English Testing system a two-faceted design – the first facet, achievement test (CET 1-3 and 5) for credit point related internal assessment, and the second, proficiency test (CET4 and 6) for external assessment. Both are standardised testing formats with only minor variations. With this design, assessment for College English is virtually synonymous to testing (Tang, 2005b; J. Wang, 2007).

The CET4/6 is administered nationwide, twice a year, by the National College English Testing Committee (NCETC) on behalf of the Higher Education Commission (HEC) of the CMoE. According to the College English Test Syllabus for Non-English Major undergraduates (CET Syllabus Committee, 1999, 2005), the CET4/6 aims at providing an “accurate, fair and objective” measurement of the English proficiency level of non-English major undergraduates as related to the set norm, while at the same time providing feedback information for teachers to improve English teaching at the higher education institutions. The objectives are to enhance
the implementation of the College English Curriculum Syllabus and ensure the qualitative and quantitative requirements stipulated in the syllabi are met. Motivating English learning and teaching in the higher education sector is also important (CET Syllabus Committee, 1999, 2005; Jin & Yang, 2006; Yang, 2003). By means of certification and control, this nationally administered testing system has exerted significant impact on College English education as well as the Chinese society during its 20 years’ implementation (CMoE, 2005; J. Wang, 2007, Yang, 2003; Jin, 2005).

1.5.2 The role of CET testing system in College English education, a historical perspective

Since the mid 1980s, the College English has had three guideline documents, the *College English Teaching Syllabus* (1985/1986), the *College English Teaching Syllabus* (1999) and the *CECR* (2007), each accompanied by a testing syllabus. Over the two decades, College English has witnessed noticeable development (Jin & Yang 2006), which can be classified into three phases: rapid development, innovation, and reform. However, the three phases have also witnessed the role of the College English assessment system changing as a response to a shift of College English curriculum requirements, the developments of English language learning and testing theories and technology, and the changes arising from Chinese political and economic contexts.


In 1985 and then in 1986, the former State Education Commission (the current China Ministry of Education) issued respectively the first *National CET Syllabus for the Science and Technology Students* and for the *Liberal Arts Students* (CMoE, 1985/1986). It was hoped that with higher English proficiency standards, China could have more technological and scientific exchange with the outside world and catch up with the developed nations as soon as possible (CMoE, 1985/1986; Hu, 2005). The emphasis of these two syllabuses was on reading ability because reading was the primary means of communication with the outside world at that time (Cai, 2002). To support and enhance effective implementation of the syllabus, the psychometrics-grounded CET testing system was developed and adopted to secure an “objective, scientific and fair” measure of whether undergraduates met the
requirements prescribed in the syllabus (CET Syllabus Committee, 1986, p.1). It was also used as a stimulus to motivate English teaching and learning in higher education sectors (CET Syllabus Committee, 1986; Yang & Weir, 2001). The CET format of the time was comprised of paper-and-pencil tests only. The test method of multiple-choice-question (MCQ) stands for 85%, which left 15% to passage writing. Test scores were based on a 100-point scale, and 60% constituted a passing grade. Being criteria-related and norm-referenced (Yang, 2003), the reported score conveyed two pieces of information: first, whether a test-taker had met the requirements of the College English Curriculum Syllabus; and second, the percentile position of the test-taker in the normed group, which consisted of over 10,000 undergraduates from six top universities in China (CET Syllabus Committee, 1999; Jin, 2005a; Yang & Weir, 2001). Certificates were issued by the HEC to students who passed the test.

During this period, the standardised CET was used as the exclusive assessment instrument of College English education in most universities. When the CET4 pass rate became a criterion to evaluate and rank higher education institutions, the institutions in turn linked teachers’ merit rewards and promotion with the CET4 pass rate, and also, pegged the CET4 score to the students’ academic degree and graduation (Gu, 2005; J. Wang, 2007). The CET4 certificate was imposed in many universities as a prerequisite for the Bachelor degree qualification. These high stakes, attached to the CET on one hand, did help to reinforce motivation as well as pressure to learn English, and hence accelerated a rapid development of tertiary EFL education. On the other hand, it directed the students’, teachers’ and the institutions’ attention to the test itself instead of College English teaching and learning (Yang, 2003). This has resulted in a marked test-orientation, in which teachers teach to the test, and students learn for the test (J. Wang, 2007). Under such circumstances, the limitations of the CET as a standardised testing format (that is mainly composed of multiple-choice questions) became evident. It began to exert negative washback on College English teaching and learning practices (Tang, 2005a).

The latter years of this phase saw constant modifications made in both form and content of the CET, in an attempt to steer away from the negative washback. For example, in 1990, the writing section was separated from the main test paper; in 1996, new items such as short-answer questions and translation were added as alternatives; and then in 1998, a minimum score requirement was set for writing (J.
Wang, 2007). However, the intended washback effects were not quite obvious. Criticisms became louder and more severe (Jing, 1999, 2000). Li Lanqing (1996), the then Vice Premier Minister, asserted at a conference that College English education was “low in both efficiency and effectiveness” (p. 9), and urged more to be done to promote effective English learning. This fuelled a round of rigorous CET innovation and prompted the issue of a new National College English Syllabus in 1999.

Overall, the CET in this phase fulfilled its designed objectives of enhancing effective implementation of the syllabus, though it unintentionally introduced some negative washback. In contrast with the previous stage, students’ English standards were noticeably higher (CMoE, 2005; J. Wang, 2007). Therefore, the CET certificate gained increasing recognition not only in educational domains but also in society (Jin & Yang, 2006).

The CET-SET and the 1999 College English Curriculum Syllabus at the innovation phase

As stated earlier, the 1999 National College English Teaching Syllabus was issued in response to the problematic phenomenon of “low efficiency and effectiveness”. It was also intended to be the means to strengthen the use of communicative language teaching approach (CLT) (Yu, 2001). Highlighting “communicative competence”, this syllabus gave equal weight to speaking and reading (CMoE, 1999) to better serve the purpose of meeting the needs of China’s Reform and Open Door Policy.

A major move on the part of College English assessment, as manifested in the CET Syllabus (1999), was the inclusion of an oral section, so called CET-SET (Spoken English Test). Positive washback effects to College English teaching and learning in the classroom were intended (Jin, 2000a; Yang, 1999). The CET-SET designers asserted that since it was a direct assessment of test-takers’ interactional competence, CET-SET should be valid as long as it was properly designed (Yang, 1999). Two timely studies conducted by Jin (2000a, 2000b) claimed that the CET-SET was welcomed by both students and teachers. Based on the survey, the research claimed to see a tendency of positive changes in College English teaching – stating the institutions began to pay more attention to improving students’ communicative competence and students were becoming more active in oral activities in class (Jin, 2000b).
Unfortunately, owing to human and financial resource limitations, the oral section was accessible only for students who gained a score of 80 or above in the CET4 or 75 in the CET6. The effects on low-achieving students might be even less. Given the disparity in teacher and student quality, uneven distribution of resources and facilities, little or no substantial changes occurred in colleges and universities of less developed or underdeveloped regions. In fact, in 2005, a large scale survey (involving 9000 students from over 50 universities in 20 provinces around the country) indicated little positive influence even on CET-SET test-takers’ (usually high-achieving students) learning motivation and practice, since, as the interviewees claimed almost all of them can pass CET-SET (W. Yu, 2005). A corpus-based investigation (He & Dai, 2006) into the validity of the CET-SET group discussion confirmed this claim, revealing that the topics failed to challenge or attract the interest of as many as 60.2% of the test-takers. Also, because test-takers tended to interpret the CET-SET as an examination rather than as an authentic communication, they regarded the examiner rather than the group members as target interlocutors. That resulted in a low range of communicative functions used by candidates and hence was a threat to the validity of the test.

Limited as the CET-SET’s positive washback effects are, with the inclusion of CET-SET, the CET testing system was established. So was its impact on society, where the CET4/6 certificate was increasingly used as a gate-keeping device for access to higher degree education, general employment, and even residential certification in big cities (Gu, 2005; Jin, 2005b; J. Wang, 2007). This gave rise to a sense of competition and crisis that pervaded the higher education institutions and led to more severe test-orientation. The College English curriculum was reported to be narrowed to give way to coaching materials; grammar-translation was once again used predominantly in classrooms to explain mock test papers; students spent most of their time memorising vocabulary and doing mock tests rather than developing communicative competence; and developing test-wise skills was common practice (Gu, 2005; Jia & Yang, 2005; J. Wang, 2007).

The situation was exacerbated by a higher education enrolment expansion starting in 1998. Insufficient recruitment of teaching staff, coupled with facility shortages, resulted in big classes. Classes of 40 to 50 were commonplace; classes of 100 to 200 were not rare in some provincial higher education institutions (Jin &
Increased competition in employment encouraged the already severe CET certificate “addiction” (Zhang, 2005). The negative impact was seen in other aspects as well. The statistics from a study by Zhang (2005) were daunting: 75% of the participating students spent an average of three hours per day on English; 80% of the students thought English as the course that took most of their time, yet yielded least effective results; and 95% identified English as the most examination-oriented of all the courses. The study concluded that CET has impacted negatively not only on College English education, but also on other disciplinary subject courses that are more important to students as qualified university graduates. Another investigation by Shu (2004) identified the CET as a “nightmare” to some low-achievers that has had detrimental effects on the psychological well-being of those students.

The above evidence suggests that the College English education saw more problems than progress during this period. The negative washback effects of the CET, mainly out of its overuse of multiple-choice and short-answer questions (Niu, 2001) have outweighed its positive impact (Han, 2002). Also the misuses and overuse of the CET test results have led to serious distortion of College English instruction and learning practices in and out of the classroom (Jia & Yang, 2005; Niu, 2001). Professor Sun Shaozheng condemned the nationally administered CET as a “sword that destroys our College English education” (cited in Zhang, 2005, p. 100). Calls to abandon College English and the CET are heard (Cai, 2004, 2005; Zhang, 2005), saying that since EFL education in primary school has been popularised in 2001, EFL in higher education should be applied for special purposes, the so-called English for Special Purpose (ESP), rather than general English. These heated debates on the now well-perceived negative effects of the CET, on the necessity of the nationally administered English test, as well as on the necessity of College English, ushered in the reform phase of College English education with the ongoing CERP.

1.5.3 The College English Reform Program (CERP)

The CERP was launched in 2004 as a major part of the National Higher Education Quality and Reform Project (2002). The CERP took multiple actions, among which were introducing a new CET testing system (CMoE, 2005) and promulgating CECR (CMoE, 2004, 2007).

Two pilot studies were conducted with the CECR (CMoE, 2004 for trial implementation), the first involving eight universities in 2004, and the second in 180
universities in 2006. Then in 2007, the CECR (revised) was put into full implementation. With emphasis on speaking and listening in an integrated communicative competence background, it aims to further raise the standards of English proficiency at the tertiary level, so that graduates can command sufficient English communicative competency to communicate in their academic and professional careers (CMoE, 2007).

The CECR promulgated by the CMoE serves as the guideline for College English reform at the institutional level. The 2007 CECR is distinguished from previous syllabi in several ways. First, three different levels of requirements with specifications (general, advanced, and more advanced) are prescribed instead of one unified syllabus to meet varied interregional and intercollegiate needs; second, student-centeredness and autonomous learning are highlighted; third, multi-media and internet learning is advocated; last, an assessment model that incorporates formative components into a summative assessment framework, and seeks to engage students is proposed (CMoE, 2007). Hu (2004) identified individualisation, collaboration, modularisation and hypertextualisation as the innovative foci of the CECR. The initiatives are that new theories of learning and assessment, and more advanced technology, could bring about substantial changes to tertiary EFL education in China.

1.5.4 The CERP and the new College English assessment policy

The CERP made two changes to College English assessment: (1) a reformed CET4/6 testing system and its status for external assessment; and (2) a College English assessment model prescribed in the CECR for internal assessment (CMoE, 2007). The first change was publicised via a press release and an accompanying testing syllabus, the second was issued through a guideline document. The changes are detailed below.

The CMoE issued a press release (2005) especially for the reformed CET4/6. According to the conference proceedings and the College English Testing Syllabus (2005), substantial changes occurred to the CET4/6. First, the high stakes attached to CET are to be alleviated by discouraging pegging the CET4/6 certificate to an academic degree certification and employment; second, the CET4/6 certification practices are to be abandoned. Instead, transcripts of results and test-takers’ performance on each section of the test but no indication of pass or failure are issued;
third, a 710-point scoring system is to be adopted to replace the 100-point system; fourth, the CET4/6 format is to change with more non-multiple choice questions (35%-40% instead of 15-20%), more listening allocation (35% instead of 20%) and authentic listening materials (CET Syllabus Committee, 2005; Jin, 2005a). With all these substantial changes, positive washback is intended (Jin, 2005a). While acknowledging the changes in multiple respects, this study is particularly interested in the shift of the external test from largely compulsory to optional, and its influence on classroom practices in relation to the formative assessment initiative.

The excerpt below is from CECR (revised version, 2007, pp. 5-6, translated by the researcher from Chinese); italics are added for emphasis:

Assessment is an important part of College English curriculum. Comprehensive, objective, scientific and accurate measurement is crucial for objectives to be fulfilled. Assessment is important evidence for teachers to get feedback information, improve teaching and management, and ensure good teaching quality. It is also effective means for students to adjust learning strategies, better their learning approaches, improve their learning proficiency, and achieve satisfactory learning outcomes. Assessment of College English learning should include both formative assessment and summative assessment.

Formative assessment is the procedural and developmental assessment conducted during the process of teaching and learning. The purpose is to ensure that what is going on during the teaching and learning process is aligned with teaching objectives. Specifically, it is a means to adapt various assessment approaches and a means to follow up the teaching and learning process, and to provide timely feedback information so as to enhance students’ overall development. Formative assessment is propitious to monitor effectively students’ self-directed learning process. It is especially important for computer and classroom-based teaching modes. Formative assessment includes self-assessment, peer-assessment, teachers’ assessment of student learning, and the administration’s assessment of students’ learning. Formative assessment can adopt approaches such as in- and out-of-classroom activity records, online self-directed learning records, portfolios, interviews and conferencing. It is used to observe, evaluate and monitor the learning process for the purpose of enhancing effective learning.
Summative assessment is evaluation conducted at the end of a certain period of study. It includes final term examinations and proficiency tests. This kind of assessment is used mainly to evaluate students’ comprehensive ability in using English – including reading, writing and translating as well as listening and speaking.

On meeting the general, advanced and more advanced requirements, institutions can organise examinations individually, or participate in intercollegiate, regional or national examinations to assess summatively the College English teaching and learning. Whatever, students’ practical communicative ability, especially listening and speaking, should be put at the centre of assessment.

This excerpt reveals several points about the reform directions. That is, assessment approaches for formative purposes (such as portfolios and conferencing) are advocated to go hand-in-hand with the summative test. Specifically, learner involvement in assessment procedures (by way of self-assessment and peer-assessment) is encouraged along with teacher’s assessment; both the assessment process and the product are emphasised; and the assessment is to serve both purposes of learning enhancement and accountability. In short, this is an attempt to integrate formative assessment components into a summative assessment framework.

With the new CET testing system and this formative assessment plus summative assessment model in the CECR, the College English testing and assessment system is partially changed in nature. The initiative goes beyond seeking alternative and better ways of assessment. Additionally, the assessment framework in this program, framed by external summative assessment washback and classroom-based formative assessment promotion, assumed a new role – to measure as well as to promote College English learning. It is anticipated that a positive impact can be brought about to improve the quality of College English education (CMoE, 2007). This underlying assumption echoes what Black and Wiliam (2001) stated in their portrayal of assessment practice. That is, by means of assessment, advances in psychology and social learning and technology, the currently undesired situations of education can be transformed.
1.6 RESEARCH GAP

Since the trial implementation of the CECR in 2004, a varied weighting (ranging from 10% to 60%) of the College English assessment is reported to have given way to formative assessment in an increasing number of higher educational institutions (Tang, 2006). The official report (HEC, 2006) on the two rounds of pilot implementation presented largely positive findings in terms of active involvement and morale increase on the part of both teachers and students. The report also indicated optimistic prospect for its development. However, concern remains as to (1) the extent to which the official purpose of the CERP and the available theories articulate with each other; (2) how this reform, based on elements that originated outside of China, aligns with the local context; (3) how this top-down policy will be interpreted and implemented; (4) the extent to which this integrated assessment model will positively affect students’ EFL learning; and above all, (5) how the latest assessment theories can be socially and culturally situated to bring optimal results to EFL education for a sustainable future in the Chinese higher educational context. These are all areas that need to be explored.

The CERP began its full implementation in 2007 with the CECR (revised version). The assessment model on which it is based borrowed new elements and concepts from Western countries and the general educational discipline. It is common with curriculum changes for issues to arise, particularly when theories and concepts are borrowed from different contexts, and conflicts and problems are to be expected in the process of implementation (Biggs 1996a; Fullan, 2004; Hu, 2002). As a consequence, the changes to the assessment policy open up many possibilities for exploration both theoretically and practically.

Research conducted in China to date, as published in Chinese academic journals, consists mainly of empirical experiments aimed at exploring the feasibility and application of formative assessment (Cao, Zhang, & Zhou, 2004; Tang, 2006; Zhou & Qin, 2006; Xue, 2006). There have been very few studies conducted that focus on the theoretical alignment of this integrated assessment model, its enactment in the classroom contexts and its impact from the teachers’ and students’ perspectives. This is the primary area of research interest for this study.

The study will seek to fill the gap by:
examining the theoretical developments of learning and assessment as they relate to ELL and assessment, and the application of these theoretical positions to the Chinese higher education context;

• analysing the uptake and enactment of the changed assessment policy at institutional and classroom levels; and

• identifying principles and conditions for assessment to support ELL in policy and practice.

In this way, the study aims to serve objectives at both macro and micro levels. At the micro level, the research aims to examine the impact that the assessment policy change might have on College English teaching and learning. This objective is to be achieved by investigating and analysing teachers’ and learners’ understandings about and responses to the newly introduced assessment approaches. It has an informative value for College English assessment policy and practice in the service of EFL learning. At the macro level, the research aims to examine the alignment of the intended role of the current assessment policy with its enactment. This objective has a critical and evaluative orientation to determine the extent to which the policy claims are met. It needs to be noted that of the two policy changes, the formative assessment initiative is foregrounded, whereas the external test initiative is placed on a background position. That is, the focus of this study is the tension arising from the implementation of formative assessment in an environment where external tests are still in use (Breadly, 2007).

1.7 RESEARCH QUESTIONS

This research will address the following questions:

1. How is the College English assessment policy change interpreted at an institutional level?

2. To what extent does the assessment practice that is implemented reflect the policy intent?
   a. What are teachers’ understandings of the assessment policy change?
   b. How do teachers relate their understandings to assessment practice?
   c. How do students view the approaches implemented to assess their English language learning?
d. How do students respond to the approaches implemented to assess their English language learning?

As Gipps (1999) noted, in a social learning context, the assessment process itself is “inherently both a social and cultural learning situation” (p. 375). With these questions addressed, this study seeks to address the question of whether the changed external test and the incorporation of formative assessment components in the College English assessment framework as advocated in the CERP (CMoE, 2004) has supported learning in this particular social and cultural context and how assessment can better serve the purpose of effective College English learning.

1.8 SIGNIFICANCE OF THE RESEARCH

The significance of this research relates to the topic, the methodology, the perspectives taken, and the potential implications of the findings. To begin, assessment has always been a focal topic in the educational domain (Torrance, 1996; Broadfoot, 2007), and in the applied linguistics field in particular (Clapham, 2003). With the conceptual shift (Gipps, 1994) in learning and assessment, formative assessment assumes a vital position in educational innovations and reforms (Murphy, 2007). That is the case with the CERP in Chinese EFL higher education contexts. In this reform, the role of assessment as a measurement and promoter to learning is of primary concern to all stakeholders. Furthermore, because there have been very few studies on the impact of formative assessment-related reform on teaching and learning (Carless, 2008; Davison & Leung, 2009; Keppell & Carless, 2006; Rea-Dickins, 2004, 2007), the research topic is valuable in its potential to bring insights from the Chinese context into the international domain.

This research will explore the implementation of the College English assessment policy change in two university cases (one Key and one Non-key in accordance with the Chinese university classification) in the Chinese higher education context. The case study methodology allows a deep and thorough inquiry into the topic (Stake, 1996; Yin, 2003) and provides valuable insights (Flyvbjerg, 2006). Conducting two cases will also allow comparative analysis, facilitating more in-depth discussion into the research problem (Yin, 2003).

This study is also distinctive in the perspectives it has taken. Specifically, it takes a general educational perspective rather than a linguistic one. It elicits its
theoretical support from the domain of educational assessment, and aligns these interests with those in the applied linguistics discipline. With such a perspective, this study aims to look far and deep. Also a sociocultural perspective (Murphy, 2008; van Lier, 2004; Rogoff, 2003) is adopted. That means a keen eye is kept on the interactions between the assessment principles and the sociocultural context of the Chinese higher education, and the relationship involved. It is believed that this is the first study that has adopted this theoretical perspective to investigate the assessment reform in the Chinese context.

The findings of this study are expected to contribute to new understandings about English language assessment and learning. The findings are also expected to have implications for assessment policy and practice in Chinese College English education. The research undertakes to use the findings to theorise on the practice of “how to” achieve efficiency and effectiveness in College English language teaching and learning via assessment. The findings may have implications and ramifications for other universities and other levels of EFL education in China too. In addition, this study is situated in China, the origin of Confucian Heritage Culture (CHC). The implications of this study may provide insights regarding similar concerns faced by Asian countries with similar backgrounds and contexts. The definition of CHC and the key tenets of CHC related to this study will be elaborated in the literature review (Section 2.5).

1.9 THESIS OUTLINE

This thesis comprises eight chapters. The present Chapter, Chapter 1: Introduction, contextualises the proposed research within its background and particular context – EFL for non-English majors in the Chinese higher education. It also defines the key concepts for proper use, states the objectives to be achieved and the significance of the study, and then identifies the research questions to be addressed. Chapter 2: Literature Review starts with the theoretical framework, which informs the layout of the literature review and the research design. Then, a critical review and synthesis of extant literature will sketch the researcher’s understandings about the theoretical development of assessment and ELL. Further, a perusal is conducted into the Chinese culture of learning and the assessment tradition in China to investigate the feasibility of implementing the innovation from foreign origin into a local context. This is followed by a discussion of the identified research gap from a
national as well as international research perspective. Chapter 3: Methodology firstly states the sociocultural approach that informs the case study methodology. This is followed by a description of the specific design – the boundaries of the cases, specific methods to be used, the participants to be involved, data collection and analysis methods, instruments, process, and the ethical considerations. Chapters 4 and 5 document the results of the two case studies. The descriptions aim to capture the influence of the new assessment policy and practice on College English learning from the perspectives of teachers and the students and the interactions of both. The description will also provide an insight into the complexity involved in putting the theory and policy into practice. Chapter 6 makes a cross-case analysis of the two single cases. The research questions are addressed in this comparative analysis. Major findings elicited point to a conclusion about the impact of the CERP assessment policy change. Chapter 7: Discussion explicates from a sociocultural perspective what affords and constrains the translation of the assessment policy change within classroom practices. The concluding Chapter 8 draws on the findings of the previous analysis and discussion, and elaborates upon the implications for assessment policy and practice in its particular context.
Chapter 2: Literature Review

This chapter starts with the theoretical framework (Section 2.1), which states the sociocultural perspective adopted to examine the research problem. Section 2.2 describes the transformations that have occurred in the educational assessment domain during the past two decades. The resultant misalignment between assessment and learning and the tensions between assessment purposes and assessment practices are regarded as the major impetus for policy changes in these educational domains. Sections 2.3 and 2.4 present theoretical development in the fields of English Language Learning (ELL) and language assessment. The focus is on the theoretical alignment between ELL and language assessment. Sections 2.5 and 2.6 contextualize the research focus in its particular learning culture and assessment context. The multidimensional barriers and affordances to the reform implementation are made explicit. Section 2.7 presents the research findings regarding teachers and English assessment practices. Section 2.8 identifies the factors that influence students’ views and responses to assessment. Lastly, Section 2.9 presents a summary of the literature reviewed.

2.1 THEORETICAL FRAMEWORK

2.1.1 An interpretativist epistemology

The study adopts a sociocultural perspective. This perspective views knowledge as being actively constructed rather than passively transmitted, and hence falls into a constructivist paradigm (Moore, 2003). A constructivist paradigm regards learning and knowledge as generated through meaning construction in the social world rather than objective reflections of an external world (Piaget, 1975; Rogoff, 1999; Vygotsky, 1978; Wertsch et al., 1995). This is the epistemological position that I take to look at the issue under research.

2.1.2 Sociocultural theory and this study

emphasises the roles of “social interaction and cultural context in learning and identity formation” (Gipps, 1999, p. 362). It aims to “explicate the relationships between mental functioning, on the one hand, and the cultural, institutional, and historical situations in which this functioning occurs, on the other” (Wertsch et al., 1995, p. 3). Murphy et al. (2008) place this theory at the extreme right end of the learning theories (see Figure 2.1). They argued that learning viewed as “appropriation through participation” is a further development of social constructivism, which asserts learning happens by adaptation in social interaction and is located in an agentive and local view of mind.

By “situating learning as an aspect of interrelated historical, cultural, institutional and communicative process” (Renshaw, 1998, p. 83), sociocultural theory of learning has profound implications for teaching, learning and education as a whole (Tharp & Gallimore, 1988). This perspective is foundational to the theoretical framework of this study. Lantolf (2007) summarises sociocultural perspectives of learning into three fundamental propositions:

1. Human mental activity is always and everywhere mediated;
2. Mediation develops through internalization of socially constructed activity;
3. Instruction, development, and assessment are inseparable processes linked in a dialectical unity. (p. 693)

The following section will present each of these three propositions and examine their respective implications for this study.
Learning is mediated

Learning, in Vygotsky’s (1978) view, is more than an internal and independent psychological development of cognition as Piaget (1975) claimed in his cognitive discovery. Rather, learning is mediated through and determined by sociocultural experience and/or an “expert other”. The latter is captured within the concept of Zone of Proximal Development (ZPD) (Vygotsky, 1978). This concept focuses on what a learner can do with the assistance of a more capable other and what can be achieved alone in the future. As such, it highlights a futurist perspective to learning (Moore, 2003). During the learning process, the children or learners must gradually develop “conscious mastery” over their learning rather than simply reciting which may have little meaning for them (Vygotsky, Davidov, & Silverman, 1997). This concept is also understood as “autonomy”, “self-regulation”, “metacognition”, or “scaffolding” in metaphor terms (Lave & Wenger, 1991). It implicitly foregrounds a role change for teachers from instructor to facilitator or scaffolder, and more importantly, for learners from passive information recipients to autonomous knowledge constructors. With these role changes, a dialogical rather than monological teacher-student relationship is established (Moore, 2003).

These understandings of learning relate to assessment in that conventional summative assessment falls short of what is needed for a judgement of learners’ development (Moore, 2003). Therefore, other forms of assessment such as formative assessment which focuses more on the learning process are foregrounded and have contributed to assessment changes to promote reforms (Broadfoot, 2007; Moore, 2003) such as the College English Reform Program (CERP) in China.

Learning is situated and appropriated

Vygotsky (1978), in his seminal book Mind in Society, emphasises the importance of social involvement in learning: “human learning presupposes a specific social nature” (p. 88). Bruner (1996) adds a cultural aspect to this point, suggesting acknowledgement of the influence of the social and cultural context on learning. Thus, learning is a process of social participation and happens through “negotiation of meaning” within particular “communities of practice”, in which learners construct their identities (Lave & Wenger, 1991; Wenger, 1998, 2000). That is, learning is “situated, active and interactive” (Moore, 2003 p. 11). This engagement in actions and interactions is also termed as “legitimate peripheral
participation” (Lave & Wenger, 1991). The interactions are not limited to those between participants. More importantly, the interactions include those between learners and the social and cultural context of learning, and the historically held ideology and cultural values about learning (Rogoff, 2003; Wenger, 1998).

Therefore, from a sociocultural perspective, learning is not seen as traditionally residing within the individual *per se*, rather, learning is “embedded in and constituted by social, cultural and historical processes” (Renshaw, 1998 p. 85). Murphy et al. (2008) reiterate this sentiment: “sociocultural approaches to learning emphasise the socially negotiated and embedded nature of meaning-making and how learners learn to use the cognitive tools of their cultural community through participation in social activity” (p. ix). In other words, sociocultural approaches to learning take into account the constituted nature of the context and the learners rather than isolating learning from the specific context. This point is expressed by Lave and Wenger (1991) as:

> That perspective meant that there is no activity that is not situated. It implied emphasis on comprehensive understanding involving the whole person rather than “receiving” a body of factual knowledge about the world; on activity in and with the world; and on the view that agent, activity, and the world mutually constitute each other. (p. 33)

Based on these insights, Rogoff (2003) observes that cultural tools of thinking such as values, models of social relations and practices within a shared cultural community influence the individual’s perceptions and constructions of behaviours. This theoretical insight serves as the rationale for me to situate the research problem not only into the College English learning and assessment context but also into the Chinese cultural community of practice. I present the historically inherited Chinese learning culture (Section 2.5) and examination tradition (Section 2.6) as key components of the Chinese cultural community.

**Instruction, learning and assessment are inseparable processes linked in a dialectical unity**

Learning theories rarely present relevant theories or even statements about how learning should be assessed within their models (James, 2006). Nevertheless, learning needs to incorporate compatible assessment; otherwise, effective teaching and learning practices will be undermined (James, 2006). The sociocultural
approaches to learning provide two new perspectives for understanding instruction and assessment practices, and their interplay with learning (Lantolf, 2007). First, assessment, just like learning, is an “interactive, dynamic and collaborative process” rather than a static product (Gipps, 1999, p. 375). Second, assessment is “an integral part of learning and instruction and is embedded in the social and cultural context of the classroom” (Gipps, 1999, p. 376).

These new perceptions have engendered notable changes in assessment in two ways. First, they have resulted in “considerable changes in the practice and philosophy of assessment” (Gipps, 1999, p. 366) in terms of design, administration and use (Pellegrino, Baxter & Glaser, 1999). The focus of assessment has shifted towards a broader assessment of learning, greater engagement of teachers and learners, and an emphasis on enhanced learning for the individual (Broadfoot, 2007; Gipps, 1999). A concern, however, is when research-informed assessment principles precede current practices; misalignment increases the complexity of the intended policy changes (Pellegrino et al., 1999; Willis, 1993). To identify the degree of alignment between current English language assessment practices and the advocated assessment principles in the CERP reform is a focus of investigation for this study.

Second, these perceptions necessitate a linkage to be established between the two areas: theory of learning and pedagogy, and theory of assessment (Pellegrino et al., 1999), which are often separated and sometimes inconsistent with each other (Willis, 1993). For the purpose of this study, I investigate the linkage between English language learning (ELL) and English language assessment, the two areas with which this study is particularly concerned in Sections 2.3 and 2.4. The understanding is that the disconnection between these two areas and the misalignment between theoretical development and practice are two major causes that have contributed to the complexity of policy implementation (Adie, 2010). Pellegrino et al. (1999) maintains that implementing reforms in one area without consistent changes in the other makes it difficult to achieve the intended results. Willis (1993) strongly supports this point, saying that the rhetoric of curriculum reform for constructive learning is meaningless unless assessment reflects complementary theoretical principles. These theoretical insights of the assessment changes in education are evidenced by the washback studies in the discipline of EFL in various settings (Cheng, 1998, 2005; Qi, 2005, 2007; Wall, 1993). They inform
my understanding of the complexity involved in implementing initiatives such as the CERP policy.

2.2 ASSESSMENT TRANSFORMATION

2.2.1 Changing views and paradigms

In the 1980s and 1990s, educational assessment underwent a series of transformations (Dwyer, 1998). The transformations firstly related to the underlying paradigm. Specifically, “the orthodox, psychometric testing framework that highlighted the key tenets of social efficiency curricula, behaviourist learning theories and scientific measurement, gave way to a contrasting social constructivist conceptual framework that blends key ideas from cognitive, constructivist, and sociocultural theories” (Shepard, 2000, p. 5). Shepard (2000) presents an historical overview of the changing conceptions of curriculum, learning theory, and measurement to explain the current incompatibility between new views of instruction and traditional views of testing (reproduced in Figure 2.2). This “paradigm shift” (Gipps, 1994) in assessment is regarded as a response to the establishment of the constructivist paradigm in learning theory. The new constructivist assessment framework embodies an “antithesis” in many cases (Shepard, 2000) to the old behaviourist framework in addressing the questions of what assessment is conducted by whom, when, and for what purposes (Broadfoot, 2007). These aspects will be addressed below.

![Image of Figure 2.2: Changing conceptions of learning, curriculum and measurement (Shepard, 2000, p. 5).](image)

**Figure 2.2.** Changing conceptions of learning, curriculum and measurement (Shepard, 2000, p. 5).

**What is assessment?**

The behaviourist view assumes learning is objective knowledge that can be transferred and attained bit by bit through stimulus-response on an individual basis.
Learning is sequenced in hierarchical small steps and can be measured scientifically (Hull, 1943; Skinner, 1938, 1954). This view of learning leads to the view of tests as “isomorphic with learning” (Shepard, 2000, p. 5). Tests are used rightfully to ensure learning aims are achieved before moving on to the next step. Thus assessment from the behaviourist perspective relies largely on objective testing to ensure grasp of constituent elements. Its content emphasises rote memorisation and its format is “locked in a particular and outdated conception of subject matter” (Shepard, 2000, p. 5).

In contrast, constructivism views learning as active mental construction and sense-making through social interactions (Bruner, 1996; Piaget, 1975; Rogoff, 1999; Vygotsky & Cole, 1978), which leads to a different view of assessment. Assessment from this perspective is seen as a continuous ongoing practice in the classroom rather than an important static event occurring at the end of the course; a qualitative process rather than a quantitative product; integrated with and central to learning rather than separated from it (Gipps, 1999; Lantolf, 2000). This realisation shifts the agenda of assessment to a consideration of how assessment can be “a constructive part of the process of learning” (Wiliam, 2001, p. 768). Thus the nature and the role of assessment are reconfigured.

**Assessment by whom?**

A behaviourist view of learning positions learners as passive knowledge recipients limited by intelligence and hence as “the assessed” (Hull, 1943). In contrast, constructivist theories position learners as active meaning-constructors, who play an active and central role in their own and others’ skills and competence building (Nuthall, 1997) within “communities of shared practice” (Elwood & Klenowski, 2002) or a “figured world” (Holland et al., 1998) or an “activity system” (Engeström, 1987). Since assessment is an integrated part of the learning process, learners should be actively and constructively involved in assessment processes as well. Through involvement, learners develop self-confidence, identity and a capacity to monitor and manage their own learning (Broadfoot, 2007). The capacity, labelled as “meta-cognitive skills” or “meta-cognition” is regarded as more important for the success of learners in the future, as it embodies learners’ gains in terms of skills and the ability to “learn to learn” (Black et al., 2003; Broadfoot, 2007). Learner engagement therefore is seen as an assessment agent and, as a result, is a priority in
formative assessment-related reforms. The role of the teacher changes from an authoritative knowledge transmitter in behaviourist theory to a scaffold in a constructivist’s perspective. The teacher’s role as agent of assessment has also changed from an authority to a negotiator (Broadfoot, 2007). This will be reviewed in Section 2.7.

**Assessment for what purposes?**

The traditional form of assessment, typically summative standardised tests, aims to summarise what is learned (Wiliam, 2001). It is regarded as problematic for “assessing only a small part of the learning claimed to be attained” (Wiliam, 2001, p. 165). The assessment from the constructivist perspective emphasises the process of learning and view learning as for the future. It highlights the formative purpose of assessment to enhance and support learning (Ross, 2005). The shift of emphasis in assessment purpose carries important implications for assessment policy and practices. Notably, assessment for learning improvement and enhancement purposes becomes in part the impetus for educational reform initiated for better education attainment (Broadfoot, 1992; Pellegrino et al., 1999). In all, educational assessment showcases an overall changing tendency from behaviourist to constructivist, from quantitative to qualitative, from knowledge to skills and understandings, from products to processes, and from measurement to improvement (Klenowski, 2009; Wiliam, 2001).

Central to all these considerations is the emphasis on formative purposes of assessment and the tensions between formative and summative functions of assessment (Biggs, 1998; Broadfoot, 2004; Wiliam, 1996). In the following section, I will look closely at the relationship between these two assessment approaches, examining the purposes and the functions they perform in relation to learning.

**2.2.2 Formative and summative assessment, a dichotomy or continuum?**

The relationship between formative assessment and summative assessment had been a controversial topic of the past decade. It is now agreed that it is a matter of decision. Harlen (2005) has noted:

All assessment in the context of education involves making decisions about what is relevant evidence for a particular purpose, how to collect the
evidence, how to interpret it and how to communicate it to intended users.

(p. 207)

Such decisions follow the purposes of conducting the assessment. The purpose
of assessment is critical in that it “determines how students react and how teachers
and schools respond in terms of curriculum content, form and organisation”
(Goldstein, 1993, p. 33). Newton (2007) suggests the need to clarify the purposes of
educational assessment before further discussion about decision-making. Taking this
point as my cue, I now present the definitions of the two terms before reviewing the
debate over the relationship between formative and summative assessment.

The definitions of formative assessment and summative assessment were
originally related to the timing (during the process or at the end) and
purpose/function (to help learning or to summarise what has been learned) aspects of
assessment (Scriven, 1967). Bloom et al. (1961) expanded the meaning and defined
summative evaluation (summative assessment) as “the assessments that are given at
the end of a certain period of teaching/learning and are designed to judge students’
learning for the purpose of grading, certification, evaluation of progress or to judge
the effectiveness of a curriculum” (p. 117). Formative evaluation was defined as a
contrastive type of evaluation.

Now the distinction concerns the nature of the assessment event and the nature
of the assessment purpose, that is, the use to which assessment results are put
(Newton, 2007). Harlen’s (2006) expression is more explicit: “what lies at the heart
of the distinction between formative and summative assessment is not the timing and
frequency, but the use of evidence, who uses it, how and for what purposes” (p. 209).
Wiliam and Black (1996) argue to distinguish the two assessment functions by
meanings and consequences they brought about and defined formative assessment as
to “aid learning” and summative assessment for “review, transfer and certification as
well as accountability to the public” (Black & Wiliam, 1998, p. 34). Based on
research, the Assessment Reform Group (ARG) (2002) proposed ten principles to
guide classroom assessment practices. The principles foreground, among other things,
learning and the learner’s central position, constructive and sensitive feedback,
sharing of goals and criteria, and development of learners’ reflective and self-
managing capacity.
The tensions over formative and summative assessment were firstly explored in literature reviews of the impact of classroom assessment on student learning (Black & Wiliam, 1998; Crook, 1988). Notably, Black and Wiliam (1998), based on a meta-analysis of evidence from the research literature on assessment and classroom learning, claim that formative assessment provides better support for learning. Their work has triggered a favourable view of and interest in the promising effects of formative assessment (Taras, 2005). This interest has coincided with literature reviews on the impact of standardised testing (Hamp-Lyons, 2007; Solórzano, 2008; Au, 2007). A negative view of summative assessment and the impact of high-stakes standardised testing are apparent in these reviews. While upholding formative assessment, summative assessment in the form of high-stakes standardised testing, is being put into “a position open to abuse” (Broadfoot, 2000, p. i).

Researchers are divided in their understandings about the relationship between formative assessment and summative assessment and the developing tendency of assessment theory and practice. Some tend to regard the two assessment functions as a dichotomy that cannot be reconciled. Goldstein (1993), for example, argues that formative assessment, to achieve the purpose of learning improvement, encourages learners to expose their weakness for subsequent correction. Contradictorily, summative assessment, which aims to maximise their scores for the best possible results, can sometimes lead learners to cover up their weaknesses. These arguments warrant caution if the same evidence is used for conflicting purposes simultaneously. The potential harm of ignoring these problems relates to an assessment system and even to the intellectual integrity (Goldstein, 1993). This stance has other supporters (Pellegrino, Chudowsky, & Glaser, 2001; Teasdale & Leung, 2000). Teasdale and Leung (2000), for example, regard formative assessment and summative assessment as “incommensurate discourse” conflicting at every facet: theoretical underpinnings, purposes, approaches and procedures. Therefore, the assessment evidence should not be misused or cross-used. On this point, Pellegrino et al. (2001) go further. They warn that if a single assessment seeks to perform multiple purposes, it is likely that the assessment purpose will be compromised, and the more purposes for which the single assessment is used, the more likely it is that compromise will occur. These theorists tend to see an either/or relationship between formative and summative assessment. Their perspective, however, is not universally accepted.
Some scholars do not see a sharp discontinuity between the purposes of assessment for learning and assessment to measure learning (Harlen, 2006; Harlen & James, 1997; Wiliam, 1996). Rather, they perceive formative and summative assessments as sitting at the two ends of the assessment continuum (Harlen, 2006) or a spectrum (Wiliam, 1996), with many intermediates in-between (see Figure 2.3) in terms of both purposes and practices. This view of the relationship tends to favour flexibility and practicality in assessment practices for different purposes.

![Figure 2.3. A continuum of assessment purposes and practices (adapted from Harlen, 2006, p. 105).](image)

Still some scholars see no point in making the distinction between the formative and summative functions of assessment (Kennedy et al., 2006). So called good assessment (Boud, 1991; Jolly & Grant, 1997) is an attempt to blur the boundary between summative and formative assessment. There seems to be no settled consensus regarding the question of whether formative assessment can be used in place of summative assessment or whether the two can be harmonised in one assessment model (Taras, 2005).

However it should be understood that the range of assessment purposes is not limited to formative and summative purposes. Broadfoot (2007) stated that the assessment purposes constitute a continuum ranging from “understanding the complex process of learning at one extreme to that of providing for the ongoing monitoring of quality at the other through teacher appraisal, institutional inspection and national evaluation” (p. 14). Hence, the debate over the purposes of assessment is mainly focused on the prioritising of a certain purpose among the wide choices rather than a choice matter of one or the other (Broadfoot, 2007).

In light of these debates and tensions associated with formative and summative assessment, Shepard (2000) asserts that the extreme favour of one purpose over another is temporary. It will be a “merged, middle-grounded theory that will eventually be accepted as common wisdom and carried into practice” (Shepard, 2000, p. 6). I support this position and believe the multiple purposes of assessment will be reconciled in some way in the future although the process might be hard to implement. Wiliam (1995, 1996, 2001) argues for a need for combination,
aggregation and reconciliation between formative and summative functions of assessment and proposes the combination to be done on evidential and consequential bases.

Integrating the emergent assessment for learning initiative with the traditional accountability objective, that is, combining the strengths of both assessment functions to bring about a synergy seems most appealing to policy makers (Goldstein, 1993). This will at the same time cater to the interests of the stakeholders involved and the society at large (Carless, 2008). In reality, it is admitted at classroom level and promoted at policy level in some European (Black & Wiliam, 2005), African and Asian countries (Carless, 2008; Chen, 2008; Davison & Cummins, 2007; Ross, 2005). However, the question remains: is the integration of the two assessment purposes thereby possible?

Theoretical support came from Sfard (1998), who, on explicating the metaphors underlying the two theoretical paradigms, warns that choosing just one might face dangers. She argues for the need to unify the two learning theories. Harlen (2006) provides an affirmative answer to this question from a more practical perspective. She claims it is possible to use the same evidence for both formative and summative purposes provided that a distinction is made between evidence and interpretation of the evidence. Though Harlen (2006) admits that the dual use of evidence is “asymmetric”, she insists that assessment systems, with appropriate provision, can be planned and implemented to transform tensions into synergies. The synergy of the purposes, which comes from making use of the same evidence for both purposes and having the same person responsible for using the evidence, can be achieved in the processes of assessment (Harlen, 2005, 2006). As such, it seems that the integration is theoretically and practically possible. The potential relates to better performance in summative assessments (Greer, 2001; McDonald & Boud, 2003) and enhanced learners’ motivation and interest to learn (Liu & Carless, 2006). The key is to ensure that assessment is conducted and used in a way that fits the designed purposes (ARG, 2009).

Nevertheless, a number of challenges exist. First, the technical issues and theoretical doubts on validity and reliability of formative assessment need to be addressed before it is widely implemented in educational practices (James, 2006; Taras, 2005). Second, “context is a significant issue”, as Black and Wiliam (2005,
p. 256) concluded in a comparative review of the recent assessment policies and practices of compulsory education in England, France, Germany and the United States. Black and Wiliam (2005) found that the effective integration of formative and summative functions of assessment was likely to be extremely difficult technically. The assessment practices in these four countries took different forms as constrained and afforded by the policies, politics and cultural factors in their particular contexts. They concluded that there were no “royal roads” for individual education contexts to follow in reconciling the learning enhancement purpose of assessment with the accountability demands. The societal and cultural context, learning and assessment environment will be major influential factors to the assessment policies and practices (Black & William, 2005). The importance of context will be discussed further in Sections 2.4, 2.5 and 2.6.

Third, policy implementation is a complex issue (Spillane et al., 2002). Even with training and support provided, the principles of assessment for learning were found to be implemented to the “letter” but not to the “spirit” owing to unchanged teacher-student relationship and beliefs about learning (Marshall & Drummond, 2006). On other occasions, the principles were found distorted so that “real learning was sacrificed to performance on a test” (Klenowski, 2009, p. 1). Cases were also reported in which the meaning of formative assessment was narrowed and appropriated in contexts where the concept was imported (Hume & Coll, 2009).

Fourth, the subject matter needs to be taken into account. As Black (1999) put it:

The directions of desirable change are to be chosen in the light of the learning aims of each subject. Both pedagogy and assessment have to be fashioned in the light of assumptions both about how learning is best achieved and about the epistemology of the subject under consideration. (p. 121)

The subject matter of this study is ELL. In the following two sections, I will present a historic review of the evolution of ELL and assessment. The focus will be the particularities of the discipline in relation to formative assessment. It also arises from the need to situate this study in its disciplinary community of practice.
2.3 THE EVOLUTION OF ENGLISH LANGUAGE LEARNING

2.3.1 Second Language Learning (SLL) and Second Language Acquisition (SLA)

I start by clarifying some terms. When English language learning is mentioned in the theoretical sense, Second Language Acquisition (SLA) or Second Language Learning (SLL) is used (Swain & Deters, 2007). SLA, as a sub-discipline of Applied Linguistics (AL) is the term that is used predominantly in the research field at the theoretical level and associated with an established discipline and large body of work. SLL is a relatively new term. However, the use of the term SLL represents a deliberate departure from some of the key tenets embedded in traditional SLA notions of language and language acquisition (Mitchell & Myles, 2004). I explain these differences below. This study will use the concept SLL (at the theoretical level) while allowing the use of other terms as they appear in citations as a respect for the original literature.

2.3.2 Second Language Learning, a historical perspective

The theorising of SLL has been relatively short (Lantolf, 2000). It used to be a language pedagogy adjunct to linguistics before it developed into an independent field of research in the 1950s (Mitchell & Myles, 2004). Since then, SLL theories have witnessed noticeable advancement. The emphasis of theoretical development has shifted in response to the redefinition of the nature of language and the language learning process. Progressively over time, language and language learning have been viewed from the linguistic, functional, and social-cultural perspectives.

The linguistic view

The linguistic view of language remains by and large the study of language structure (Mitchell & Myles, 2004). Two significant views are that of behaviourism and structuralism. Behaviourism regards language as behaviour and habit and leads to a conviction that language learning is the formation of habits from a stimulus and response pattern (Bloomfield, 1933; Watson, 1924).

Chomskyan structuralism, with the introduction of Generative Grammar, holds that language is systems “governed” and “bound” by rules (Chomsky, 1959, 1986). The key to successful learning lies with learners’ creativity and exposure to naturally-occurring interactions (Chomsky, 1959, 1986). Language learning is, thus, viewed as a process of activating the language systems present in the biologically
built-in or "innate" faculty – the language acquisition device (Chomsky, 1986). That is, in response to language situations, generating language from the grammar that is "universal" to all natural human languages leads to a degree of linguistic competence (Chomsky, 1986). This revolutionary approach to the study of language gives a great stimulus to the field of linguistics and especially to the study of language learning (Mitchell & Myles, 2004). However, Chomskyan transformational grammar, revolutionary as it is in linguistics, remains the study of language structure as Bloomfield’s behaviourism does. It therefore has achieved limited effect on language teaching practices (Rajagopalan, 2004).

Within this linguistic view of language, tests are equated to language learning. An orientation towards "objective and scientific" measurement of discrete points of knowledge in the 1950s leads language assessment to include the multiple-choice questions in its traditional pencil and paper format (McNamara, 1996). That started a new era of standardised testing for language assessment, which has lasted for more than 40 years (Phelps, 2005).

**The functional view**

Systemic Functional Linguistics (SFL) of the 1970s claims that the meaning of language lies in its use and function in a social context (Halliday, 1977). Functional linguistics is firmly concerned with performance (language in use) rather than competence. This theory holds that language occurs in real life as "socially constructed and socially purposeful text" (Kress & Knapp, 1992, p. 5). The meaning as well as structure of a text is attributed to its use in the context of culture and situation and in its service of purposes (Kress & Knapp, 1992). Thus, language learning involves acquiring not only the syntax of language but how it is used in conversation. The form of the language is the vehicle for making social meaning, rather than an end in itself (Halliday & Hasan, 1985).

The systematic functional view of language and theory of communicative competence provides the rationale for performance assessment of language (McNamara, 1996). This approach to language assessment advocates measuring the "integrated and total communicative effect of an utterance" (Carroll, 1972, p. 318) and the communicative competence that reflect students’ ability in authentic communicative contexts (Savignon, 1983). The emphasis in performance assessment is on the directness and authenticity of assessment (Bailey, 1996). As such, positive
washback is explicitly encouraged. The models proposed by Canale and Swain (1980) and Bachmann and Palmer (1996), both modified later, have influenced communicative language testing since its inception (McNamara, 1996).

**The sociocultural view**

Starting from the late 1980s, the new understandings of mind and mental development have found their way into the SLL field. Given the socially and culturally embedded nature of language itself, the new theories of cognition and learning have been embraced with quite a speed and breadth (Lantolf, 2000). Consequently, SLL has incorporated the cognitive and sociocultural perspectives, captured so far mainly in the sociocultural theory of language learning (Lantolf, 2000, 2007). By viewing language learning in essentially social and cultural terms, the theory has led to the current interest in the approach known as SLL (Mitchell & Myles, 2004).

The sociocultural approach to language learning claims that target language learning cannot be viewed simply as a source of “input” for autonomous and internal learning mechanisms (Lantolf, 2000, 2007). According to this approach, English language learners construct their language proficiency by actively participating in “collaborative dialogue” (Swain, 2007) with peers or an expert other – teacher or other artefacts. During the process, the learners are expected to develop their autonomy so that they can outperform their competence and regulate their learning process when all the assistance is no longer available (Swain, 2007). This notion, regarded as a projection of Vygotskian Zone of Proximal Development (ZPD) (Lantolf, 2007), shifts the teacher’s role from an authoritative transmitter to a facilitator, guide, or coordinator (Nunan, 1999).

The sociocultural theory of language learning assumes that interaction itself constitutes the language learning process, which is essentially social in nature (Mitchell & Myles, 2004). Language learning is more likely to happen through social interaction in the form of collaboration and mediation (Lantolf, 2000), and socialisation (Watson-Gegeo, 2004), with consideration of social, cultural, and even political contexts.

This concept of language learning has problematised the previously-accepted understandings of language as based on an assumed set of common fixed norms, and
the traditional notions of English language development as simply the acquisition of English language proficiency (Davison, 2007; Firth & Wagner, 2007). Wells (1997) argues from his study of second language acquisition that:

The conception of (English language) teaching as transmission is a mistaken one, knowledge cannot be transmitted… it is constructed afresh by each individual knower on the basis of what is already known and by means of strategies developed over the whole of that individual’s life, both outside and inside the classroom. (p. 218)

These perspectives have contributed to a change in ELL practices. The conceptual shift in language learning is a manifestation of the conceptual shift in the understandings of learning itself from “symbolic cognition” to an increased awareness of the “situated and socially distributed” nature of ELL (Cross, 2010). On this shift, Block (2003) comments as the SLL field has taken a “social turn”. Through the social turn, both the understandings of language learning and the constructs of language have been redefined (Davison, 2007) and reconceptualised (Firth & Wagner, 2007). The implications of this conceptual shift relate directly to ELL practices and procedures in terms of what to learn and how to learn it. Specifically, the emphasis on situated learning enables the definitions of ELL-ness to cover much wider and complex domains, such as literacy, the sociocultural backgrounds and other constructs (Davison, 2007). Moreover, the foregrounding of learning autonomy in ELL includes a metacognitive component in ELL development and a need to update the teachers’ understanding in terms of their professional roles and relationships involved (Borg, 2003, 2006).

The implications of the conceptual shift in ELL provide impetus for corresponding change in English language assessment in regards to what to assess and how to assess it (Davison, 2007). This leads researchers to examine the prevailing language assessment mode in terms of content and format. The findings are unanimous in that the old assessment mode, which takes into account only linguistic factors or proficiency, fell short of capturing much of what is encapsulated by the renewed understandings of English language development (Lafford, 2007). The standardised large scale language testing format, which is based on behaviourism and psychometric theory, has become outmoded as a valid language assessment instrument (McNamara, 1997). Furthermore, the washback effects of the
standardised language testing are perceived to be more negative than positive, and constraining rather than affording ELL learning in many cases (Cheng, 2008b; Cizek, 2005; Coniam & Falvey, 2007; Gipps, 1994; Glaser & Silver, 1994). Hence, it is not surprising to find ELL assessment changing in ways similar to that in general education assessment. As such, language assessment has assumed a renewed role to reform as well as inform English language learning and teaching (Davison, 2007).

2.4 ASSESSMENT IN ENGLISH LANGUAGE LEARNING

2.4.1 English language assessment as a discipline

Assessment has been an important area of research, policy and practice in the field of Applied Linguistics (Brown & Hudson, 1998; Cheng, Rogers, & Wang, 2008; Kunnan, 2005; Leung, 2004). Language assessment is a field that is primarily concerned with the psychometric qualities of tests (Kunnan, 2005). Namely, it has been dominated by testing, that is large-scale, standardised, and divorced from the business of teaching and learning (Davison & Cummins, 2007; Davison & Leung, 2009). To be exact, language testing is the only field of study in which “the measurement of a particular thing has evolved into a separate academic discipline” and developed into a worldwide business (Davidson, 2004, p. 85). In the Inner-Circle countries where English is the dominant language and most people are native speakers (Kachru, 1996), the English proficiency test is something used as a gatekeeper for immigration and academic purposes (McNamara & Roever, 2006). Tests such as the Test of English as a Foreign Language (TOEFL) in North America and the International English Language Testing System (IELTS) in the UK and Australia are among the most influential. In some English as a Second Language (ESL) Outer Circle countries (e.g., Singapore and Hong Kong) and English as a Foreign Language (EFL) Expanding Circle countries, large scale tests are used as a tool of policy to promote English learning, such as the Hong Kong Certificate of Education (HKCE) in Hong Kong, the National Matriculation English Test (NMET) and College English Test (CET) in China. The results of these tests are used as gatekeeping devices for higher education entry, employment, promotion, replacement and even immigration. Thereby, language tests exert a powerful impact on the stakeholders (Shohamy, 2001; McNamara & Rover, 2006; Coniam & Falvey, 2007).
2.4.2 Transformations in English language assessment

Despite the established testing tradition in the area of English language assessment, during the last decade, the concepts, approaches and procedures of the emergent assessment in wider education (elaborated in Section 2.4) have projected into the English language assessment domain (Ross, 2005). Specifically, English language assessment shows a similar tendency of moving away from objective standardised testing and indicates an inclination towards classroom-based or teacher-based assessment (Davison & Leung, 2009). The transformations in language assessment, though lagging behind the general educational field by a decade, have engendered a major conceptual shift and impacted greatly on English language teaching and learning practices (Davison & Cummins, 2007; Davison & Leung, 2009).

The conceptual shift in language assessment relates to changes at epistemological and methodological levels (Lynch, 2001, 2003). At the epistemological level, the positivistic view that the psychometric language testing theories and testing practices are grounded in began to give away to interpretivism in a post-modern era (Lynch, 2003). Consequently, “solid belief in the validity and reliability of the objective quantified measurement of language is shaken” (Lynch, 2003, p. 2). The acknowledgement that standardised testing is not viable to measure all aspects of language learning in terms of knowledge, skills and ability has brought about changes to language assessment at the methodological level (Lynch, 2003). There is a tendency to include qualitative components in the construct and content validity of language assessment mechanisms, to shift the focus of language assessment to the classroom and the learner as an individual (Hamp-Lyons, 2007). There is a similar interest in introducing language assessment alternatives to the language testing domain (Bachman, 2002; Brown & Hudson, 1998).

Hamp-Lyons (2007) regards these shifts in language assessment as embedded in two different and often conflicting assessment cultures - the traditional examination culture and the emergent learning culture. Assessment practices in a learning culture are primarily concerned with learning and teaching, and are usually classroom-based and learning-oriented, whereas in an examination culture classroom assessment is seen as simply preparation for external set and assessed examinations. So learning and teaching are examination oriented. Hamp-Lyons (2007) illustrates
within the table below (Table 2.1) the differences between the two different assessment cultures, which lie in ideologies, principles, purposes, practices and results. These differences demonstrate the conceptual transformations of English language assessment. Similar to the assessment transformations in general education, the purposes of assessment for learning function that lies within formative assessment (Ross, 2005) is highlighted in the language assessment transformation.

Table 2.1
Two Assessment Cultures (Hamp-Lyons, 2007, p. 498)

<table>
<thead>
<tr>
<th>Classroom-based assessment culture</th>
<th>Classical large-scale testing culture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency-focused</td>
<td>Accuracy-focused</td>
</tr>
<tr>
<td>Individual-focused</td>
<td>Group- or “norm”-focused</td>
</tr>
<tr>
<td>Achievement-focused</td>
<td>Proficiency-focused</td>
</tr>
<tr>
<td>Learner-focused</td>
<td>Language-focused</td>
</tr>
<tr>
<td>Process-focused</td>
<td>Product-focused</td>
</tr>
<tr>
<td>Teachers’/students’ voices</td>
<td>Rule-makers’ voices</td>
</tr>
<tr>
<td>Leads to “assessment of and for learning”</td>
<td>Leads to “teaching to the test”</td>
</tr>
</tbody>
</table>

Hamp-Lyons (2007) regards the contrasts between formative and summative assessment and the two assessment cultures as “dynamic and highly contextualized”, namely, “multidimensional” rather than a “dichotomy” (p. 498). Researchers in Hong Kong (Carless, 2008; Kepple & Carless, 2006) identify these two assessment cultures with the terms examination-oriented assessment culture and learning-oriented assessment culture. This thesis will adopt the Hong Kong version, while keeping the term learning culture and assessment culture for later use in Sections 2.5 and 2.6 with a different connotation.

To shift the orientation of assessment culture from examination to learning so as to achieve better learning attainments is now a major driving force in the EFL educational reforms in the higher educational contexts of some Asian countries such as Hong Kong (Carless, 2008; Kepple & Carless, 2006), Japan (Ross, 2005) and China (CMoE, 2007). These initiatives try to incorporate formative assessment elements into a summative assessment framework.

In contrast to these initiatives is the fact that formative assessment in the area of ELL has “much variability, a lack of systematic principles and procedures and a
dearth of information as to the impact of teacher-based assessments on learning and teaching” (Davison & Leung, 2009, p. 389). Davison (2008), acknowledging the complexity of assessment practices within classrooms, propose that there could be a typology of possibilities (Figure 2.4).

A review of studies conducted in the Asian contexts (Carless, 2008; Kepple & Carless, 2006; Ross, 2005; Tang, 2006) sheds some light on the current study.

Kepple and Carless (2006) reported a case study that was carried out in a teacher education university in Hong Kong. This action research proposed a “learning-oriented assessment” framework (as reproduced in Figure 2.5 below), which was designed to support a learning function in addition to the more traditional measurement function.

The researchers assumed this assessment framework could be applied to either summative or formative purposes as long as a central focus was on enhancing student learning. Also, by putting learning at the centre of assessment and reconfiguring the assessment procedures, it was hoped to benefit both tutors and students and that “a gradual change in the assessment culture can be facilitated” (p. 181). However, this longitudinal study indicates a mixed result. Both positive elements in the learning process and the negative elements of heavy assessment workload are revealed.

In a later study, Carless (2008) illustrated a module designed to integrate formative assessment with summative assessment through feedback. The data
analysis revealed not only a rather positive attitude of the pre-service teacher participants towards formative assessment but also learning outcomes that were “superior to” the results from other similar cohorts in the final term assessment. There was evidence in the study that the students had changed their orientation toward assessment. Based on these findings, Carless (2008) suggested a guiding principle for the productive synergy between formative assessment and summative assessment: “summative assessment should be designed with learning foremost” (p. 10). Carless (2008) also emphasised the significance of a relationship of mutual trust and the necessity to communicate the rationale for an assessment process to students. He stated that students were generally willing to cooperate if they understood the rationale and could see the potential learning benefits.

More recently, Carless (2010), taking on a sociocultural lens, called for the need to address the challenge embedded in the sociocultural context. He cautioned that the incorporation of formative assessment, which is largely based on Anglophone and European literature and evidence, in the Confucian Heritage Culture (CHC) settings is “far from easy” (Carless, 2010, p. 1). Based on sociocultural standing, Carless (2010) proposed a “contextually grounded approach” (p. 2) and advocated adapting formative assessment to suit the contextual needs and cultural factors pertaining to the CHC settings.

Ross’s (2005) longitudinal study on a cohort of 2215 tertiary EFL learners in the Japanese context is informative about the effects of formative assessment and summative assessment integration on learning outcomes. The contrastive analysis of quantitative data showed that the involvement of formative components in assessment practices contributed to substantive improvement in academic listening. In contrast, little or no salient advantage was found in reading improvement. The positive listening result was attributed to the shift of the locus of control to the learners, which enhanced more learner engagement and provided a domain-specific stimulant to EFL learning, especially when “formative assessments are recognised by the learner as eventual inputs to summative criteria” (Ross, 2005, p. 337). However, Ross (2005) warns that formative assessment needs to be tempered with limiting conditions and caveats; fair and accurate formative assessment depends on responsible and informed practice on the part of teachers and on the assessment experience for learners.
Tang (2006) conducted a similar study in the Chinese context, which investigated the use of formative assessment for learning in a teacher university. It looked at the effects that formative assessment had exerted on College English learners’ learning outcomes in terms of language proficiency, learner autonomy, and self-efficacy. It adopted a mixed method approach. The analysis revealed that 12 weeks practice of formative assessment achieved a value-significant improvement in learner autonomy and self-efficacy, but the effect on language proficiency was not obvious. The constraints related to personal attitude, extra workload, and assessment culture.

These studies are encouraging in that they show that a synergistic effect of summative and formative assessments in the area of English language is possible in practice. The efforts can be understood as an attempt to bring the orientation of an assessment culture from examination to learning via fine-tuning along the assessment continuum (Keppell & Carless, 2006). This assessment design in these two studies is similar to the assessment model advocated in the CECR and therefore might present good lessons for the CERP to draw upon.

2.4.3 The significance of context

The significant role that context plays in the effective implementation of assessment principles (Black & Wiliam, 2005) is evident in the area of language assessment (Cheng et al., 2004, 2007, 2008; Davison, 2005; Hamp-Lyons et al., 2001). The context in the area of language assessment is wider (Hu, 2002; Miao et al., 2006), given that it not only involves the political and social contexts, but also encompasses the English learning contextual factors that Kachru (1996) defined in his Three Concentric Circle model (elaborated in Section 1.3.1).

The significance is evidenced in empirical studies in the field. Hamp-Lyons and her colleagues (2001) made an experimental effort to introduce process English writing in Hong Kong, a context where examinations are emphasised. The approach has been shown to contribute positively to English Language Learning in the American context and has helped to shift the assessment focus from summative to formative values. However, the same approach fell short of achieving comparable results in Hong Kong. The reason, according to the study, is that it is hard for the teaching community as a whole in an examination-oriented context to ignore the
high-stakes attached to external exams and in its stead, teach and assess adhering to formative values (Hamp-Lyons et al., 2001).

The comparative study that Davison (2005) conducted on ESL teacher assessment practices in Australia and Hong Kong (two different assessment contexts) also reveals the significance of the learning and assessment context in recognising salient differences in teachers’ attitudes to and beliefs in the formative purpose of assessment and their actual assessment practices.

The above findings find support in a series of comparative studies that were conducted across Canada, Hong Kong, and China (Cheng, Rogers, & Hu, 2004; Cheng, Rogers, & Wang, 2008; Rogers, Cheng, & Hu, 2007). It was revealed that the learning and assessment context exerted a significant influence on the assessment methods and procedures that teachers adopted in different contexts. These studies will be reviewed in Section 2.7 under the title of teacher and assessment practices.

Given the significance of context, I will situate this study in the Chinese culture of learning and the Chinese context of assessment. This action is also supported by the sociocultural framework this study adopts, which advocates exploring learning from historic, social, and cultural perspectives (see Section 2.1).

2.5 CHINESE CULTURE OF LEARNING

2.5.1 The role of culture in learning

The values shared by participants of a certain culture are critical determinants of what is perceived as appropriate to learning and how to go about learning in that society (Salili, 2001; Volet, 1999). These values form a particular framework of expectations, attitudes, beliefs, values and behaviours that are taken for granted (Cortazzi & Jin, 1996) and manifested in educational context (Salili, 2001). The whole set of values is termed a culture of learning (Cortazzi & Jin, 1996). It can be characteristic and particular to certain social contexts, and needs to be taken into account especially when trying to import educational theories or pedagogies based on different cultural norms (Biggs, 1996b; Ho et al., 2001b; Hu, 2002).

Sociocultural theory views the social contexts that share these values as a cultural community (Rogoff, 2003; Wenger, 1998, 2000). The values shared are “cultural tools for thinking” (Rogoff, 2003, p. 258) that mediate the meaning construction of learners within the shared community of practice. The section
focuses on the Chinese culture of learning and the key cultural tools that are embedded in this particular learning context.

2.5.2 Chinese culture of learning and English Language Learning

Cortazzi and Jin (1996) define the term *Chinese culture of learning* as “the whole set of frameworks characteristic of the Chinese way of teaching and learning” (p. 169). This culture of learning is deeply rooted in the Confucian educational and cultural traditions of China, and is affected by the socio-economic conditions in Chinese society (Cortazzi & Jin, 1996). It defines Chinese people’s understandings about the nature of learning and learning process, their attitude and expectation of a teacher/learner relationship, and exhorts their strategies and approaches to learning (Biggs, 1996b). The values are widely accepted at the different phases of education and used across the subjects. Hence, it is justified to say they have lasting impact on the Chinese way of learning and teaching (Cheng & Curtis, 2009b; Cortazzi & Jin, 1996; Paine, 1990).

The Chinese culture of learning persists even when it comes to learning a foreign language (Cortazzi & Jin, 1996; Hu, 2002; Zhu, 1992). China’s scholarly tradition in EFL teaching and learning has been a combination of the grammar-translation method and audio-lingualism (Hu, 2002). Despite the failure to develop an adequate level of communicative competence in Chinese learners of English, this traditional approach has been difficult to change despite the efforts and resources expended (Hu, 2002). The reason is, in a large part, due to the support it draws from the Chinese culture of learning (Cortizza & Jin, 1996; Zhu, 1992), which is arguably the largest constraint on the adoption of educational innovation of foreign origin in the Chinese context (Hu, 2002).

2.5.3 Two key tenets in the Chinese culture of learning

The Chinese culture of learning is known as a typical *Confucian Heritage Culture* (CHC) (Biggs, 1996b; Ho, 2001a). CHC, as the name suggests, has been much influenced by the Confucian philosophy on education (Biggs, 1996b). But it should be noted that the influence of the CHC is not limited to Confucius (551-479 BC) himself. His disciples (3000 in all and 72 of them well known) such as Mencius and scholars of later generations, contribute to the doctrines of Confucian philosophy as well. In addition, other scholarly schools, notably Taoist, and religions, notably
Buddhist, also have a share in shaping the Chinese learning culture (Kelen, 2002). Indeed, some of the ideas have become popular sayings and accepted by the general public. Its influences also extend to most of the Asian countries like Japan, Korea, Singapore, and Thailand, which gives CHC local diversity. The diversity exists in different regions of one country. In light of its development over the years and the local diversity, overgeneralisation and oversimplification have been more than once cautioned (Carless, 2010; Park, 2010). In this thesis I adopt a definition of CHC, which is generally accepted in the research community as: “a group of Asian nation-states with their motherland and overseas population who share Confucian values, which consistently reflect in social behaviour and practices, including academic outcomes and learning approaches” (Park, 2010 p.1). I also understand the development of CHC over the time and space, and therefore take a developmental view of CHC. That is, CHC and its values, as mentioned in this thesis, is to be guarded against generalisation. Similar to Carless (2010), I use CHC as “a useful umbrella term to refer to a relatively stable, recurrent set of practices” (p.65). Here only the two tenets most relevant to this thesis will be elaborated upon: the first being reverence for learning, and the second, teacher authority.

**Learning is revered and linked to utilitarian functions**

Learning in the CHC tradition is accorded great importance and high reverence (Biggs, 1996b; Ho et al., 2001a; Salili, 2001). Confucius explained the role of learning was to “cultivate oneself, then regulate the family, then govern the state, and finally lead the world into peace (修身，齐家，治国，平天下 xíushēn, jíjiā, zhìguó, píngtiānxià) (The Great Learning, IV).” That is, learning is viewed as important for the individual’s character building, family pride, as well as social development and world peace (Lee, 1996). The use of learning (now as institutionalised education) as a means to strengthen the country and enrich the people (强国富民 qiángguó fùmín) is evident in Chinese educational policy, including the current foreign language policy in China. These views in part account for the high regard for learning. The fact that learning opportunities in the past were available only to the privileged few who could afford them added to the reverence for learning (Zhu, 1992).

For the average person (老百姓 lǎobǎixìng), this high regard for learning was usually associated with the utilitarian functions that success in learning brought with it (Salili, 1996; Zhu, 1992). For example, it was believed that the achievement in
education can lead to upward social mobility and material rewards such as wealth, fame, a beautiful wife, and anything that was associated with high social status (Lee, 1996). The belief was that through learning, an obscure person can become superior (Hu, 2002). “A gold fish jumps over the dragon gate (鲤鱼跳龙门 lǐ yú tiào lóngmén)” illustrates the dynamic ascendency in social status available from successful learning in ancient China. The saying “in books one can find houses of gold as well as beautiful girls (书中自有黄金屋，书中自有颜如玉 shūzhōng zì yǒu huāngjīnwū, shūzhōng zì yǒu yánrúyù)” illustrates the added values of this ascendancy. Ho et al. (2001a) describe this phenomenon as an examination superstition, which he illustrates with an equation:

\[
\text{Examination results} = \text{academic performance} = \text{academic achievement} = \\
\text{future academic achievement} = \text{occupational success or failure} = \\
\text{socioeconomic status} = \text{personal and familial achievement} = \text{personal worth} \\
\text{and glory or disgrace to the family. (p. 37)}
\]

These benefits have served to motivate generations of Chinese to achieve highly in education in history as well as today (Biggs, 1996b; Ho et al., 2001a).

This view of education in its utilitarian terms has certain effects. Among others, it tends to encourage the view of learning as a means to an end (Ho et al., 2001a). Namely, it directs people’s attention to the outcome that learning brings about and how to achieve the desired outcome rather than the efforts that learners need to make and the process that learning involves (Hu, 2002). Namely, learning in the Chinese culture is largely oriented towards outcomes and examination. Education in this context is more about “gaining future socio-economic benefits” such as employment than personal development (Carless, 2010 p. 69). A pragmatic approach to learning is thus commonplace. This tradition holds fast in Chinese learners and prevails in today’s educational and social system (Biggs, 1996a, 1996b; Carless, 2010; Cheng & Curtis, 2009a, 2009b; Ho, 2001a; Salili, 1996).

This examination orientation in the Chinese culture of learning, manifested in examination obsession and certificate-addiction, has brought about serious consequences in Chinese education. According to Biggs (1996a), this orientation tends to entail surface rather deep learning strategies and learning outcomes, which are not good for sustainable learning for the future. More seriously, Ho et al. (2001a) assert that with the continuity of examination-orientation in learning, the education
might sacrifice knowledge generation and application for knowledge acquisition and retention only. Such being the case, the goals of education are subverted rather than advanced by examinations. The implication for this study is that this examination-orientation could be a situation that militates against the successful implementation of the current CERP in China, which hopes to achieve better quality by advocating the learning orientation. The consequences of examination will be elaborated in more detail in Section 2.6.

**Teacher authority**

Confucius advocates the restoration of the *Zhou Rites* (*周礼*, Zhōulǐ), that is, the ritual system in the Kingdom of Zhou for a better order in society (Zhu, 1992). *Xúnzǐ* (*荀子*), another great Confucian thinker describes the role of teacher in relation to rites and order in *tiántào* (*天道*, means the rules of heaven):

Rites have three bases. Heaven and earth are the basis of life, the ancestors are the basis of the human species, and the rulers and teachers are the basis of order... if there were no rulers and teachers, how could order be brought about? If even one of these were lacking there would be no safety for man. Therefore rites serve heaven above and earth below, honor ancestors and exalt rulers and teachers. (Adopted translation in Kelen, 2002, p. 232)

This notion has resulted in another eminent characteristic in Chinese culture of learning: respect for the teacher. This respect assumes the position of the teacher as an authoritative figure and further embodies a hierarchical teacher-student relationship (Biggs, 1996; Ho et al., 2001a; Hu, 2002). These two values are now discussed.

First, the teacher is an authority. Respect for the teacher can also be regarded as an extension of reverence for learning and by extension for wisdom in the Chinese tradition (Kelen, 2002). This is because the teacher, if not successful in obtaining high social status, is usually successful in learning. Teachers in the Chinese tradition are held to be in the same position as one’s father. As is said, “one day’s teacher is your father for life (一日为师, 终生为父 *Yīrìwéishī zhōngshēngwéifù*).” Respect for the teacher is also closely linked to the role definition of the teacher in Confucian tradition. This is elaborated by HánYù (*韩愈*), a well-known scholar in the Tang dynasty, in his classical essay *On Teacher* (*师说 shīshuō*): “The so-called teacher is
to transmit Tao (knowledge), teach a trade, and disabuse doubts (师者，传道授业解惑也 shīzhě, chuándàoshòuyějiěhuòyè)”. This well accepted definition indicates traditional Chinese understandings about what teaching is about and how teaching operates. The perception of teaching as knowledge-transmission further relates to the way teaching is carried out in that it is characterised by expository, analytic explanation (Hu, 2002). As well, this definition implies people’s expectations of teachers. The Chinese teachers’ authority in classroom is described by Ginsberg (1992) as:

Teacher decides which knowledge is to be taught and students accept and learn that knowledge. Teacher is the authority, the repository of knowledge, leading the student forward into this knowledge, a respected elder transmitting to a subordinate junior. (p. 6)

Students, on the other hand, are posited as a “listener” or “recipient”. They speak in class only when invited by the teacher and hold a deferent attitude towards him/her (Carless, 2010). The role definitions of the teacher and the student in Chinese culture of learning are the antithesis to the advocated roles of the teacher as a guide, facilitator and scaffolder and the student as a participative co-constructor in the constructivist paradigm. This poses a challenge to the imposed reform.

The teacher’s authority also demonstrates its power in the widely accepted role as the sole judge or assessor in the teaching and learning process (Hu, 2002). In line with this understanding, others such as peers, students themselves and even parents are not authoritative enough to judge or assess students’ work (Cortizza & Jin, 1996; Hu, 2002). This perception is evidenced in studies on Chinese students’ attitudes towards group work and peer feedback or evaluation (Hu, 2003; Meyer, 2003). Hu (2003) found that a great proportion of the Chinese tertiary students in his investigation are reluctant to give some value to peer feedback and are sceptical of peer-assessment. This finding is confirmed by Meyer (1999) and Zhang (1995) who showed that students had an overwhelming preference for teacher feedback over peer feedback. This, in a sense, reflected the Chinese teacher’s widely approved role as judge and assessor. This wide approval, however, could present a potential obstacle to the implementation of the CERP, which advocates student involvement by way of peer and self-assessment.
Second, the role definition of the teacher as authority informs a hierarchical teacher versus student relationship. Being an authoritative figure in knowledge, a knower, and usually an elder, teachers are to be respected, not to be challenged or offended (Hu, 2002). Therefore, the Chinese teacher/student relationship is hierarchical, not equal (Ginsberg, 1992; Ho et al., 2001a; Paine, 1990).

This hierarchical relationship is demonstrated by strict disciplines, and perceived student passivity, compliance and respectfulness in students’ behaviours (Dautermann, 2005; Ginsberg, 1991; Paine, 1990). This relationship implied teacher-centeredness and structured curricula (Burnaby, 1989) and led to serious consequences. Ho et al. (2001a), for example, noted that it is “potent determinants of learning outcomes” (p. 29). They maintain that this hierarchical relationship is closely correlated to Asian students’ (including Chinese students) poor performance in terms of approaches to learning and acquiring metacognitive abilities. It impedes the preparation of learners for their future development. In another study, Ho et al. (2001b) found Hong Kong Chinese students’ silence and passivity impeded not only effective learning but also their psychological well being. These attitudes and behaviours are regarded as inhibiting the development of individualism, autonomy, and creativity in learners (Biggs, 1996b), and free interaction between teacher and student (Ho et al., 2001a). They argue that to confront teacher authority in the world of learning is the key for transformation of education in the CHC contexts. This claim resonates with Hu et al. (2001a; Yu, 2001), who argue that the authoritative values in Chinese culture of learning have been a critical barrier that accounts for the partially failed attempt to implement CLT, an ELL approach highlighting student-centeredness and interaction. In Hui’s words (1997), after nearly one decade, CLT in China is still “New bottles, old wine” (新瓶装旧酒 xīnpíngzhuāngjiǔ) (p. 38). These characteristics of Chinese culture of learning appear at odds with the principles adopted in the CERP.

Studies by Western researchers similarly foreground these potential obstacles. For example, Carson and Nelson’s (1996) micro-ethnographic study revealed that Chinese students showed reluctance in initiating comments and showing disagreements. Their concern for social harmony within the group and vulnerability prevented them from critiquing peer work.
However, there are dissenting voices. Littlewood (2000), on comparing the learning attitude and behaviour of 2307 students from eight Asian countries and 349 students from three European countries, found that the differences between the two cohorts of learners were no more than those between the learners from the same country. Asian learners, including those from China, were not especially obedient and unquestioning, Littlewood concluded. Biggs (1996b) maintains that some stereotypes of Chinese learners are misconceptions. The Chinese learners, for example, are not necessarily surface learners. They adopt memorisation a lot, but with understanding. Some of them, despite perceived disadvantages from a Western perspective, outperform many of their Western classmates. This phenomenon, labelled as “the Chinese paradox”, has aroused great interest in Chinese learner research by Western researchers (Biggs & Watkins, 1996; Mok, 2006). Moreover, there are some studies which indicate that the supposedly undesirable features of Chinese culture of learning are changing (Chan & Rao, 2009; Cortazzi & Jin, 1996; Jin et al., 2006; Meyer, 2003; Shi, 2006).

2.5.4 Changes to stereotyped norms

Cortazzi and Jin (1996) point out that the Chinese culture of learning is changing. This is attributed to China’s Open door policy (Cortazzi & Jin, 1996; Jin, 1992) and the changes that the policy brings about in the Chinese social context (Shi, 2006). Jin (1992) points out that with the Chinese learners, especially the younger generation, developing an open mind towards cultural values different from the Chinese traditional ones, a synergy might be developed out of integration of the Western cultural norm and Chinese cultural values. In a recent study, Jin et al. (2006) further maintain that Chinese cultures of learning are changing in practice as well as in ideology. This claim is supported by studies conducted in the international context as well as the Chinese context.

Shi (2006) investigated 400 secondary students in Shanghai. The study found that contemporary Chinese students, different from the stereotyped characteristics mentioned above, showed a different inclination in their English language learning attitudes and behaviours, although some traditional values persisted. For example, they preferred to raise questions after class rather than in class when they disagreed with teachers. Overall, the younger generation favoured a friend-friend (equal and
interactive) relationship with their teachers rather than a parent-child (hierarchical) one. They actively participated in classroom interactions.

Meyer (2003) conducted a survey of 100 Chinese university students at a Singapore-based English Program. The focus was on Chinese students’ preferred learning styles, attitudes towards cooperation with peers, and attitudes towards peer feedback. It was revealed that 60% of the participants showed a preference for receiving feedback from the teacher compared with 3% from peers, while 37% were neutral. But the participants generally expressed positive feelings about feedback from peers (59% “no problem” versus 40% “uncomfortable” or “very uncomfortable”). These figures indicate that though teachers’ feedback is preferred as the Chinese tradition might predict, a considerable proportion of Chinese learners are open to input from peers and to new learning approaches.

This change is echoed in Miao et al’s (2006) study, which found that though teacher feedback was adopted relatively easily (90% as compared with 67% for peer feedback) and lead to more improvement, both teacher feedback and peer feedback were valued. Moreover, as the experience of peer feedback increased, students’ attitudes changed: 61% of the students from the peer feedback group, think feedback from peers useful or very useful as contrasted to 22% from the teacher feedback group. Thereby, Miao et al. (2006) state “even in cultures that are said to give great authority to the teacher, there is a role for peer feedback” (p. 179).

These studies were mainly conducted with students from either the top universities (Meyer, 2003; Miao et al., 2006) or the most developed regions (Shi, 2006) in China. This poses some limitations to an understanding of an overall picture of Chinese students’ learning attitudes and behaviours. Nonetheless, there is no denying that changes are happening to the Chinese education as multiple reforms were introduced in the Chinese context (Chan & Rao, 2009).

2.6 ASSESSMENT ENVIRONMENT IN CHINA

2.6.1 Examination tradition

Western literature (Dore, 1976; Spolsky, 1995) has indicated that the examination originated in ancient China in the Han Dynasty (206 BC-220 AD), when the government started to use a national examination system, called kéjū (科举), to select civil officials for administration purposes. Chinese sources have traced the
origin of the examination even earlier. Meng et al. (1961), for example, claimed that the Western Zhou Dynasty (1027-771 BC) held performance examinations regularly for the selection of officials. The imperial examination system continued, with periodical intermissions, until the end of the Qing (清) Dynasty (1644-1911) in 1911 when the feudal monarchy in China was overthrown during the Xinhai (辛亥) Revolution led by SunYàntsēn (孙中山) (Zhu, 1992).

Despite some minor variations, the examination content was invariably restricted to some classical sources, notably, Confucian classics (Jin, 1990; Zhu, 1992). Examinations were highly competitive and only a small proportion of the top candidates would be appointed to official posts. The selection followed strict procedures and fixed standards to ensure that the best talents were selected for government work while preventing corruption (Jin, 1990). The practice of this examination system for over 2000 years’ has greatly influenced the Chinese people’s ideology about education and schooling. Han and Yang (2001) summarise the influences as follows:

- a strengthening of the utilitarian values of education;
- a stress on the key role that examination plays in education;
- a stress on book knowledge at the expense of practical ability;
- an emphasis on the one-off result to the neglect of formative assessment.

(p. 5)

Extended practice of the examination in Chinese history has resulted in a deep-rooted examination tradition in China and the comprehensive influence in the Chinese culture of education and other aspects of social life (Tang, 2006; Tian, 2004).

2.6.2 A turbulent period

The years between the end of the Qing Dynasty and the foundation of the People's Republic of China in 1949 were turbulent in Chinese history. Chinese education and assessment during this period virtually “stagnated” (Han & Yang, 2001). The first few years saw the newly established government streaming the schools into a key and non-key scheme to cope with limited educational resources (Han & Yang, 2001). The unbalanced allocation of resources, facilities and staffing between the two kinds of schools encouraged a trend of examination-orientation in
education (Han & Yang, 2001). Mao Zedong, the Chairman himself held strong antipathy towards the rigid imperial examination system and its detrimental consequences (CMoE, n.d.). His criticism was regarded as partially responsible for the closure of schools and the complete abolition of examinations at all levels during the Cultural Revolution (1966-1976) (Han & Yang, 2001). Examinations came back with the restoration of the Entrance Examination of Higher Education (EEHE) in 1977 after Deng Xiaoping came to power. The period after 1977 is regarded as the modern period in Chinese education history (Han & Yang, 2001). During this period, education developed rapidly. Primary education was universalised in the 1980s. The 9-year compulsory education program was implemented during the 1990s. The years after 1999 have witnessed the higher education sector develop from an elite education system to a general one with the recruitment rate rising to 62% in 2009 (Xinhua News Agency, 2009).

2.6.3 The examination system in the modern period

The examination tradition continues and is highly competitive (Biggs, 1996a; Qi, 2007) in the modern Chinese educational system. This state is closely related to the way in which the education system operates, which is characterised by centralisation, stratification, and selection (Hu, 2003).

The Chinese Ministry of Education is the central policy-making and administrative body. It coordinates educational activities within the whole education system, which are mandated through national education policies, national curricula and teaching syllabus, appointed or recommended textbooks, and above all, the national EEHE (Wang, 1996). The provincial, municipal, and local levels of administrative organisations take care of implementation. Central education policies are applied throughout the country except in some centrally administered cities such as Beijing, Tianjin, Shanghai, and developed regions such as Guangdong and Zhejiang province. That is, the education policies in China usually follow a top-down approach.

The stratification of the education system originates from the key and non-key school scheme (Hu, 2003). With the growing regional disparity in economic development and accordingly in education investment, the gap between regional educational provisions has widened (Hu, 2003). It can be said that with the key and non-key stratification from primary through to higher education (Wang, 1996) and
the regional disparity, the educational development between developed and underdeveloped regions in China is lacking in balance (Hu, 2003).

Two factors illustrate the selective feature of the Chinese education system. Firstly, the Chinese education system is “a steep pyramid” (Wang, 1996, p. 76). At the end of junior secondary (that is, the end of compulsory education), 59.2% of the students continued on to senior secondary schools, 62% of whom went on to higher education in 2009. The relatively high enrolment rate is a result of nine years of compulsory education and higher education enrolment expansion since 1998. In 1992 for example, the corresponding figures were 26% and 3.9% respectively. The steep narrowing down of the pyramid itself means intense competition.

Second, although school-based assessment emphasises all-round development which involves moral, intellectual and physical ability, large scale, unified school leaving external examinations are used at every rung of the education ladder and serve as a gate keeping device for the next stage of education. Given the fact that the external examinations after junior secondary education are manipulated to narrow down the student population, the test results are mainly used for selective purposes (Qi, 2007). Hence, the external examination is norm-referenced rather than criteria or standard-referenced. The selection practices with one external examination among the largest student population in the world make the Chinese education system notably “harsh” and “highly competitive” (Biggs, 1996a).

It is justified to say that overall the assessment culture in China has generally remained the same – examination-oriented, highly competitive, attached to high stakes, and with an emphasis on product rather than process. The only difference lies in the fact that whereas this harsh assessment environment in the imperial period affected only the few who could afford to take the examinations, today it influences the majority of the student population. Whilst the competition in the past related to access to higher education, nowadays it is mainly up to the key universities.

The examination culture is further supported by the traditional value placed on education by parents and their desire to secure a better future for their only child (Hu, 2003). In light of these pressures, the government demands high levels of accountability from the assessment system to make it appear fair and objective. As
such, the weight that the examination carries in the Chinese education system cannot be overemphasised.

At the higher education level, usually internal examinations are administered at collegiate level, and the results are used primarily for certification purposes. The past few years, however, have seen an increase of its use for selection purposes. This has contributed to the enrolment expansion in the Chinese higher education sector. Graduates are pressed to climb higher up the education ladder or acquire more certificates in the hope of gaining an edge in the increasingly competitive employment market (Powell, 2008). The CET for non-English major students is one of the few externally mandated examinations at the higher education level. Even if it could be said that the overall assessment environment in the Chinese higher education context has changed, the assessment context for College English remains the same – high-stakes and competitive. Furthermore, with twelve years’ immersion in examination culture (from primary to senior high), students have formed particular learning styles and habits, which can help them to excel in examinations. And above all, students have formed their perceptions of learning and assessment. Such habits and ideologies are difficult to change and take time to do so. These, therefore, might be barriers to the ongoing initiative to innovate College English assessment.

2.6.4 The consequences of the Chinese examination culture

The consequences of the examination culture as seen through outsiders’ eyes have been mentioned in Section 2.5. Here I review key insights from the Chinese literature. ZhúXi (朱熹), a famous scholar in the Song Dynasty, who contributed greatly to the Chinese Confucian heritage, criticised the imperial examination system. He asserted the traditional Kéjǔ (科举) system, with content and format stipulated, restricted the test-takers’ learning and even thinking, and the longer the same format and content is used, the more serious the impact would be (Tian, 2004). Shen, another Chinese scholar in the Song Dynasty, links the negative impact of examinations to the social and political stimuli lying behind them, such as the high rewards associated with the success in examination (Tian, 2004). This matches contemporary notions about social consequence and the washback effect.

In the modern Chinese educational system, the consequences of the examination culture involve several factors. First, with the emphasis on students’
academic scores, students’ development tends to be one-sided rather than all-rounded (Wu, 1994). That is, affective attributes and practical ability are sometimes neglected. This leads to “high score but low ability” (literally translated from Chinese 高分低能 gāofèndìnégéng). That is, students are skilled at answering examination questions and able to achieve high scores in examinations, but have low ability in dealing with practical problems (CMoE, 1999; Wu, 1994). Second, to achieve high scores and high pass rates in examination, the teaching and learning tend to narrow the curriculum content down to only what is tested. Third, to achieve higher and better than their peers, students have to do a lot of extracurricular exercises, go to cramming schools or training centres after school. Such a burden leaves little time for pleasure, interests and physical exercises, which has proven to be harmful to their wellbeing (CMoE, 1999). Fourth, the emphasis on examinations impacts on the quality of education, and by extension, on the future of the country (CMoE, 1999).

Well aware of these consequences, the Chinese government and authorities launched a series of policies to improve the quality of education, to relieve the students of the extra burden, and to enhance the citizens’ learning for sustainable development (Han & Zhang, 2001). Examples are: Decisions on Deepening Reform in Education and Improving Quality-based Education in All round Way and Higher Education Expansion National Conference on Education in 1999, the Higher Education Quality and Reform Project in 2002 and Implement Quality Project File No 2, Uplift the Overall Quality of Higher Education in 2007. The shared focus of all these efforts is to reform the examination-oriented assessment system. It is still too early to say that the changes in the examination culture of the Chinese education system are under way. This study resolves to provide a snapshot of the potential dynamics that this policy initiative brings to College English classrooms. With a learning culture “paradox”, a problematic assessment system, a harsh assessment context, and a top-down approach to policy implementation, the effective operation of the College English innovation initiative faces multiple challenges.

The conflicts between advocated principles in a global forum, the reality of the local context, and the CE assessment policy initiative to reconcile these conflicts can be summarised in Table 2.2 below.

Table 2.2
The Contrast Between the Advocated Principles and Reality of the Local Context
<table>
<thead>
<tr>
<th>Advocated principles</th>
<th>Reality of local contexts</th>
<th>Policy initiatives in CERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student centeredness and engagement</td>
<td>Teacher authority Hierarchical T/S relationship</td>
<td>Student centred and Teacher directed</td>
</tr>
<tr>
<td>Formative assessment</td>
<td>Summative assessment</td>
<td>Summative assessment and Formative assessment</td>
</tr>
<tr>
<td>To improve learning</td>
<td>To measure learning</td>
<td>Improve and Measure</td>
</tr>
<tr>
<td>Learning-oriented assessment culture</td>
<td>Examination-oriented testing culture</td>
<td>From examination to learning orientation</td>
</tr>
</tbody>
</table>

The prospects, however, need not be necessarily pessimistic. Watkins (1996), on reviewing a series of studies on Chinese learners, concluded that the survivors of the harsh assessment environment of the CHC context are rather adaptive to assessment procedures of a different cultural norm. The studies on Chinese students studying in overseas contexts such as Australia (Ballard & Clanchy, 1991, 1997; Volet & Renshaw, 1996) and the UK (Jin, 1992; Zhou & Todman, 2008) all provide solid support for this claim. After a certain period, the Chinese learners usually catch up with and can achieve higher than their local classmates. This seems to indicate a synergy that is intended in the CERP initiative.

According to innovation theory (Fullan, 2004), the implementation of an education innovation policy is a complex process, and involves participants including teachers, learners, educational administrators, material writers and publishers. The top-down nature of this policy and restricted access to policy makers at the national level prevented their inclusion in this study. However, the administrators at the institutional level are included for an informed understanding of policy implementation at the local context. Further, as this study adopts a case study approach investigating the effects of assessment policy change on English Language Teaching (ELT) and Learning (ELL), the focus will be on teachers and students.

### 2.7 TEACHER AND ASSESSMENT PRACTICES IN THE ELL CONTEXT

#### 2.7.1 The role of teacher in assessment

The roles of assessment in ELT and ELL are central and important (Clapham, 2003; Davison & Cummins, 2007; Kunnan, 2005). However, it is acknowledged that the assessment practices in ELL classroom are complex and multi-faceted (Cheng et al., 2004; Rea-Dickins, 2007). Teachers play an important role in such complexity
(Cheng, et al., 2004; Scott & Erduran, 2004). They need to make “constant judgement about students’ learning according to their performance, and then make corresponding decisions over assessment practices” (Cheng et al., 2004, p. 360). Research has more than once demonstrated and confirmed that the tensions that emerge in the change process impact, first of all, on teachers (Arkoudis & O’Loughlin, 2004; McNamara, 2001; Rea-Dickins, 2004). Teachers “find themselves at the confluence of different assessment cultures and faced with significant dilemmas in their assessment practices: sometimes torn between their role as facilitator and monitor of language development and that of assessor and judge of language performance as achievement” (Rea-Dickins, 2004, p. 253). As the actual implementers of policy who stand at the interface of policy and practice and the interface of teaching, curriculum and learning, teachers are crucial to the delivery of the assessment innovation (Fullan, 2004).

### 2.7.2 Factors influencing teachers’ assessment practices

Teachers’ practices in the classroom are influenced by curriculum and policy at the social order level and beliefs and experience at the individual level (Murphy, 2008). However, in contrast to the vast body of literature on the impact of high-stakes testing on curricula content and classroom activities (Cheng, 1998; Qi, 2005, 2007; Wall, 1999), classroom assessment in the ELL field “has remained relatively under-researched until recently” (Rea-Dickins, 2004, p. 249).

Cheng and her collaborators (2004, 2007, 2008; Rogers et al., 2007) conducted a series of comparative studies on the English language assessment practices employed by the ELL teachers at the tertiary level. These comparative studies were set within three different ELL contexts – Canada (English-dominant ESL), Hong Kong (English and Cantonese bilingual ESL/EFL) and China (Mandarin-dominant EFL). These studies demonstrate the complex and multi-faceted roles that the ELL classroom assessment practices play at the higher education context of differing settings. I will review the four studies below. The findings have relevance to the Chinese EFL higher education context.

The first comparative survey (Cheng et al., 2004) involves 267 ELL teachers. It revealed that while assessment methods were closely correlated to the instructional goals (external testing) in both Canada and Beijing, Chinese university teachers’ assessment practices mainly focus on instruction-centred purposes rather than
student-centred purposes. The Chinese teachers showed less inclination to provide feedback via student conferencing (42.7% as compared to 80.0% in Canada and 72.7% in Hong Kong) and other labour intensive methods, but a much higher inclination to provide total scores as final reports (87.9% as compared with 25.0% in Hong Kong). In conclusion, the assessment methods they used and the procedures they followed revealed a marked influence of large class size and objectively scored external testing culture.

A similar instrumental linkage between classroom assessment purposes with the assessment methods and procedures applied by teachers was echoed in Cheng et al. (2008). It was found that the Chinese university EFL teachers based 80% (as compared to 25% in the other two settings) of their evaluation and assessment practices of students’ performance on standardised tests and examinations results. They adapted more examination-oriented assessment methods such as selection and translation rather than learning-oriented ones such as writing in their classrooms. The variations between Chinese and Hong Kong universities teachers, sharing similar cultural values, are attributed to the fact that Chinese teachers needed to prepare their students’ English language proficiency for the required CET4/6 whereas Hong Kong teachers’ assessment purpose was to enhance student learning. The variations between Chinese and Canadian university teachers are due to the different nature of learning contexts and the embedded cultural values.

A third study (Cheng et al., 2007) investigated the teachers’ assessment practices in terms of grading and reporting of student achievement. The results revealed Chinese teachers’ use of marking criteria was predominantly holistic, with 58.3% using holistic scoring only and 20.9% using a combination of holistic and analytical. Less than half of the Chinese teachers shared with students the criteria, in contrast, nearly all the teachers in the other two contexts did. Peer-assessment was used much less in the Chinese context than the other two (41.5% as compared to 75% and 73%), as was self-assessment (29.1% as compared to 50% and 37.5%). The differences were attributed to the impact of the high-stakes test with which these two contexts were faced.

Rogers et al. (2007) found that the teachers who held a positive attitude about assessment were more likely to use the new methods in their practices. The findings were achieved by surveying 95, 44, and 124 teachers respectively in Canada, Hong
Kong and Beijing via a questionnaire. The message this study delivered was that it was important to make sure the teachers’ attitudes agreed with rather than conflicted with the policy rhetoric.

Taken together, these studies contribute to an understanding about the ELL teachers’ assessment practices in the ELL domain, which are influenced by the assessment purposes, the cultural beliefs and attitudes. Meanwhile, a profile of the Chinese tertiary EFL context, as compared to the other two contexts, is depicted. It seems the learner position in Chinese EFL classrooms is marginalised and the examination orientation in Chinese assessment culture is predominant. The implications for this study follow. First, the close link between assessment purposes and teachers’ assessment practices indicates it is necessary to know teachers’ understanding of the purposes that the language assessment is stipulated to serve. Snow (1989) suggests satisfying the need of teacher understanding and use before a new assessment framework is implemented. Without it, he asserts, no progress can be made in educational practice. Second, since affective factors such as attitudes affect the teachers’ choice of assessment approaches and material as well as their decision on assessment practices, teachers’ attitudes need to be examined for an understanding of their assessment practices. Research also revealed that teachers’ attitudes can either support or be resistant to the assessment innovation and make a difference between success and failure out of an innovation attempt (Wang & Cheng, 2005).

A recent study conducted by Xu and Liu (2009), via a narrative inquiry of a Chinese College EFL teacher’s experience, found that prior assessment experience, power relationships in the workplace and professional knowledge are three major factors that influence teachers’ assessment practice.

These messages confirm the key role that teachers play in innovations and are resonant with what innovation theory advocates to support teachers with training to make sure they are willing and comfortable with the change, especially when it comes to a top-down policy (Fullan, 2004). These findings, however, are built upon teachers’ self reports (interview or survey), which seems to be a limitation of these otherwise carefully designed and meticulously carried out studies. This, in turn, constitutes a gap in the methodology for a close observation and in-depth exploration.
in related research. That, in part, explains why this study adopts a case study approach. The rationale will be elaborated in more detail in Chapter 3.

2.8 STUDENTS AND ASSESSMENT PRACTICES

Students, as “consumers” of education (Swagler, 1978), are consumers of the assessment system as well (Higgins et al., 2002). There is now consensus that learning and assessment are “inextricably linked” (Hargreaves, 1997, p. 401). On the one hand, assessment impacts greatly on students’ learning (Biggs, 1996; Black & Wiliam, 1998, 2006; Crooks, 1998; Gibbs, 2007); on the other hand, it is students who decide on how and to what degree assessment works (Alderson & Wall, 1996; Bailey, 1996; Boud, 2007; Tang & Biggs, 1996). Section 2.8.1 below elaborates on this mutual relationship; Section 2.8.3 focuses on the factors that influence students assessment practices as literature reveals.

2.8.1 Learning and assessment: mutually related

Assessment influences learning in many ways and is now acknowledged as “logically and empirically one of the defining features of students’ learning” (Struyven, Dochy, Janssens, Schelfhout & Gielen, 2006, p. 326). Crooks (1988), on reviewing extensive psychological research, suggests that assessment plays a key role in influencing students’ learning in terms of ability to retain and apply in varied contexts, continuing motivation and students’ self-efficacy as learners. Recent literature reveals that assessment has a considerable impact on how educational courses are perceived (Broadfoot, 2007), on conditioning the learning goals (Marton & Säljö, 1997), on shaping students’ approaches to their learning (Boud & Falchikov, 2007; Struyven et al., 2006), the quality of the learning outcomes (Ramsden, 2003) and even the development of their future learning (Biggs, 1996a, 1999; Sadler, 1983). Marton & Säljö (1997) caution that given the potential effects that assessment has in framing learning, decisions on assessment should be carefully made because assessment can induce surface approaches rather than deep learning.

How assessment impacts on student learning is recognised as a complex issue, in which many factors interact with each other (Messick, 1996; Tang & Biggs, 1996). Participants in the assessment process such as teachers and students play a major role and the key role is with students (Alderson & Wall, 1996; Bailey, 1996). As Tang and Biggs (1996) have identified from the washback studies: “It is not the kind of
test constructed but the student’s perception of the test and the demands it makes that generates the effects” (p. 163). Similarly, Boud (2007) observes that “what makes a difference in assessment is what students do and how they experience what they do” (p. xix). In light of the virtual role of students in assessment, Weaver (2006) asserts that without full understanding of students’ views and responses, education cannot possibly be “truly student-centred”. This is the rationale for the section below to focus on students’ views and responses as two major factors influencing students’ assessment practices.

2.8.2 Students’ views and responses to assessment

The early studies of students’ views and responses to assessment started with an original intention to explore students’ learning experiences. The researchers (Miller & Parlett, 1974; Snyder, 1971) found that students were quite strategic in their learning. They adopted various tactics to seek cues about what would be assessed and organised their time, efforts and approaches to learning according to the question types and assessment procedures. Assessment in this way becomes the “hidden curriculum” (Snyder, 1971), and students “play the examination game” (Miller & Parlett, 1974, p. 60).

Research following these early studies revealed mixed findings. The majority of higher education students investigated in Swann and Arthurs’ (1998) study seem to take an instrumental view of learning. They regard assessment tasks as obstacles to overcome in the pursuit of high grades. Formative feedback is used as a means to negotiate their needs.

The students in the study by Higgins and colleagues (Higgins, Hartley & Skelton, 2002), however, show a different orientation. It was found that they were not pragmatic and driven solely by extrinsic motivation of gaining a good grade as assumed. Rather, they were “conscientious consumers” of education who recognised the importance of grades while at the same time were “motivated intrinsically and [sought] feedback which [would] help them to engage with their subject in a deep way” (p. 54).

In an empirical study on students’ perceptions and responses to tutors’ written comments, Weaver (2006) also found that most students valued feedback. But they perceived the feedback could be more helpful if the comments were timely rather
than delayed, specific rather than vague, focused on the positive rather than the negative, or related to the assessment criteria.

An important theme that emerged from Struyven, Dochy & Janssens (2008) is that the students’ attitudes to and perceptions of assessment can be changed. The researchers found that after a period of hands-on experience, students began to support instead of resist the new assessment methods like portfolios and peer/self-assessment. This finding again highlights the significance of guidance and support to students while introducing new assessment methods or modes of assessment.

Gibbs and Simpson (2003, 2004; Gibbs, 2007) are convinced of the dynamics that assessment (through feedback and encouraging student involvement) has on contributing to effective learning activities and worthwhile learning outcomes. Based on a review of substantial literature, they categorised eleven conditions for assessment to support learning (Gibbs & Simpson, 2003) and devised an inventory – Assessment Experience Questionnaire (Gibbs & Simpson, 2003) to diagnose problems in an existing assessment system. I adopted part of this inventory that was relevant to the current research context and these selected interview questions were used in this study to elicit suggestions for improvement in the College English assessment procedures.

Research also reveals that students’ views of assessment and responses to it are closely interlinked and also linked to the quality of learning outcomes (Ramsden, 2003; Brown & Hirschfeld, 2008). How students view the nature, purposes and specific procedures of assessment directly relate to how they react subsequently in their approaches to learning. “What is assessed is likely to form the priority for learning” (Broadfoot, 2007, p. 10), whereas the nature of what is being assessed and how it is assessed informs implicitly the assumptions about what learning is (Hargreaves, 1997). For example, when the content assessed is focused on recollection of facts or knowledge, the underlying assumption about learning is reproduction of information; when the assessment adopts the format of multiple choice questions and true or false question, it is implied there are definite correct answers. But, when what is assessed concerns understandings and transformation, learning is assumed as seeking meaning and interpretation (Willis, 1993). The assumptions about the nature of learning and assessment relate to whether surface or
deep approaches are taken to their learning (Entwistle & Entwistle, 1991; Hargreaves, 1997; Ho et al., 2001a).

Students’ views of the purposes of assessment act as goals for students’ efforts in study by telling them what to aim for (Tang & Biggs, 1996). When students study for an exam, they will try to understand the learning materials in ways that they perceive will meet requirements and adopt the most convenient approaches to maximise their grades. But when they study for learning, they will lay their eyes on improvement rather than the grades and accordingly adopt high-level cognitive skills (Entwistle & Entwistle, 1991). Actually, Carless (2008) argues that whether or not students have been informed about the purposes and the potential results of an assessment reform program might make a difference between cooperation and resistance in their attitude. Harlen and Crick (2003) reiterate this sentiment, saying students’ understanding of the purposes and results of assessment as intended is a prerequisite for assessment to have a positive impact on students’ learning. Otherwise, it might well be “counter-productive”.

Students’ understandings of the specific assessment procedures are also regarded as closely linked to their approaches to learning. This is the conclusion that Tang and Biggs (1996) draw from their experiment on the use of two different assessment modes: tests and assignments. They found that how students cope with assessment depended in a large part on what they see as the assessment task requirements. If students are familiar with the assessment type and know how to do it, better learning can be achieved with the new assessment mode. Otherwise, they fall back on their established approaches. When new modes of assessment are introduced it is important to support students with appropriate guidance for completion.

This claim is confirmed by Smith and Gorard (2005), who reported an evaluation study of assessment practice change initiated to enhance the use of formative feedback. They concluded that formative feedback itself cannot bring about fundamental improvement to students’ learning. When technical support and attention are lacking, misunderstandings and misgivings of participants can lead to negative results. The guidance that students need is not only in technical terms but also affective terms. Orsmond et al. (2000) drew on their theoretical exploration into the introduction of formative assessment and found that when students are
encouraged to take responsibility of their own learning, students’ understanding of the assessment criteria is also important.

In summary, students’ responses to assessment are influenced by the way they view assessment in terms of nature, purposes, procedures and criteria. These four categories will be the major areas that this study will investigate. I acknowledge that the actions that students take are also affected by other factors, such as whether they are alert “cue-seekers”, whether they are equipped with appropriate strategies, and their personal agenda (Tang & Biggs, 1996). These factors are beyond the scope of this study.

2.8.3 Implications for this study

The role of assessment in influencing learners’ perceptions and shaping learners’ approaches to learning is evident in English education in the Chinese context. This can be seen in the washback studies conducted by Li (1990) and Qi (2005, 2007) at the secondary level and tertiary level (He & Dai, 2006; Yu, 2005). The literature on Chinese tertiary students’ views of and responses to the formative assessment and student engagement is limited since it is still quite new in the Chinese context, as evident from Section 2.5.

Hu (2003) illuminated a new perspective on Chinese university students’ attitudes and responses to group work and feedback in English as Foreign Language (EFL) Learning. The study spanned over six years, and involved six cohorts and 439 students from major universities in China. Data analysis revealed a clear disparity between students from economically more developed regions and less developed ones. In terms of classroom behaviour, the students from less developed regions showed notable reluctance to engage themselves in group work, a lack of learner autonomy, lower adaptability and considerable resistance to the communication-orientated, learner-centred learning context and pedagogy. As well, they valued teachers’ assessment more highly than self and peer-assessment. The disorientation they experienced, the struggle they made to adapt, and the frustration they encountered were troubling. Hu (2003) traced the contributing factors to their analytical view of the nature and process of language learning and learning habits that have taken shape in their learning experience in less developed contexts. Official statistics show that students from rural areas in 2006 account for 52.5% of registered undergraduates. The findings in Hu’s (2003) study pose the questions: How do these
rural students perceive and respond to the changes in the newly implemented College English assessment policy? Will they show a group orientation that is different to students from urban or developed regions? This perspective arouses my interest and hence is also included in my study.

2.9 SUMMARY

With the above review, this chapter (1) has established the theoretical perspective that this study adopts to explore the problem under research; (2) informed by this perspective, has set up a comprehensive background for the study by providing a critical review and synthesis of the literature on learning and assessment; (3) has situated the study in the Chinese cultural and the ELL disciplinary communities of practice; (4) has presented the argument for the changes in Chinese College English assessment policy and practice; (5) has demonstrated the significance of the teacher’s role in assessment practice and assessment innovation particularly; (6) has identified the research focus as it resides with students’ perceptions and responses for the purposes of revealing the impact of assessment policy change on English Language Learning; and (7) has synthesised theoretical explanations and empirical studies to provide a frame for the conduct of this study. With the existing literature canvassed and theoretical perspective established for the study, I now turn to the design and methodology that will be used to investigate the research questions.
Chapter 3: Research Design

This chapter starts with an elaboration of the sociocultural approach this study adopted as a research theory (Section 3.1). This is followed by the research design as informed by the approach, which includes the case study approach (Section 3.2), the selection of two local settings and participants (Section 3.3), the data collection methods used and the tools developed (Section 3.4). Then, the detailed description of the piloting (Section 3.5) and the actual data collecting process (Section 3.6) are presented. The measures taken to validate the study are also covered (Section 3.7). This is followed by the data analysis procedures (Section 3.8). Lastly, Section 3.9 discusses the ethical considerations of the research.

3.1 SOCIOCULTURAL APPROACH AS RESEARCH THEORY

Sociocultural theory is significant to this study in its cognitive sense as a learning theory, which reconfigures the understandings of learning, assessment and the interrelationship between the two (see Section 2.1). It is also of significant value to this study in its methodological sense as a research theory, which frames the research design as a whole.

3.1.1 Sociocultural approach as a research theory

Sociocultural theory is used to explicate the trajectory of human actions within its situated context (Wertsch et al., 1995). By expanding the cultural, contextual and political dimensions of our understanding, sociocultural theory has opened up an innovative perspective to educational policy studies in the past decade. A sociocultural approach to educational studies was demonstrated in an inaugural work by Levinson and Sutton (2001): Policy as/in practice – A sociocultural approach to the study of educational policy and a more recent article Education policy as a practice of power: Theoretical tools, ethnographic methods, democratic options (Levinson, Sutton, & Winstead, 2009). As a typical educational study of policy, this research bases its design on the sociocultural approach.

The sociocultural approach to educational policy, first, brings a renewed understanding to policy and related notions such as policy process and the role of actors. Educational policy for example is understood as “a complex social practice of
“power” which, through “governing statements about what can and should be done”, “defines reality, orders behaviour and sometimes allocates resources accordingly” (Levinson et al., 2009, p. 770). Meanwhile, policy is also regarded as being “instantly negotiated and reshaped in spaces of diverse levels” (Levinson & Sutton, 2001, p. 1). Policy process, which includes policy formation and policy implementation, is viewed as “a complex, ongoing social practice of nominative cultural production constituted by diverse actors across diverse social and institutional contexts” (Levinson et al., 2009, p. 770). Specially, policy process is viewed as dynamic and interactive with context rather than simply following a hierarchical and linear model. Actors involved in the policy process such as institutional administrators, teachers and students are not passive recipients of educational policy. Rather they are “fully cultural animals” who are empowered to enable the policy and construct meaning through negotiation with their cultural values and contextual constraints (Levinson & Sutton, 2001, p. 3).

In terms of methodology, the sociocultural approach to educational policy studies is “grounded” (Levinson & Sutton, 2001). It views the policy and its implementation from a “bottom-up” rather than “top-down” perspective (Fitz, Halpin, & Power, 1994). Specifically, the sociocultural approach does not concentrate solely on the policy formation at the top; rather, it is more concerned with the experience of those who are affected by it such as the localised actors or stakeholders, and how the policy impinges upon them (Levinson & Sutton, 2001). The role of the actors as meaning interpreters is highlighted against the conventionally assumed role of passive recipients who simply execute policy which has been handed down. The actors are considered to interpret the meaning/s of policy and engage themselves differently with the policy as it is implemented. A sociocultural approach argues that the meaning of policy is interpreted or appropriated, redefined and virtually recreated in the implementation. Policy formation and implementation are processes whereby actors at different levels appropriate and negotiate the meaning of policy in their community of practice. The appropriation is mediated by their cultural tools and contextual situations (Levinson & Sutton, 2001). During the process of appropriation and negotiation, the actors form or shift their identity in varied degrees. The variation in appropriation implies a necessity of multiple settings for informing purpose. Levinson and Sutton (2001) summarise the sociocultural approach to the study of
educational policy as aiming to “explicate policy as a practice of power and to interrogate the meaning of policy in practice” (p. 1).

Local actors’ localised appropriation needs rich documentation and ethnographical representation of a qualitative inquiry (Fitz, Halpin, & Power, 1994). So does their practice which “gets at the way individuals and groups engage in situated behaviors that are both constrained and enabled by existing structures, but which allow the person to exercise agency in the emerging situation” (Levinson & Sutton, 2001, p. 3). Accordingly, the sociocultural approach to educational studies tends to use research methods of a qualitative nature such as interviews and observations, which are suitable for detailed generation of data. What lies with the examination of assumptions in policy formation against the policy as interpreted and enacted in practice is “not only the concern about the relative effects or effectiveness of policies” but also a programmatic interest “to chart out new tools for research and change” (Levinson & Sutton, 2001, p. 4) on meaning negotiation with existing studies. The sociocultural approach informed the design of this study as described below.

3.1.2 Sociocultural informed research design

As detailed previously, this study is about examining assessment policy change in the CECR and its implementation. From a sociocultural perspective, the focus of this study is on the “agency of local actors in interpreting and adapting such policy to the situated logic in their contexts of everyday practice” (Levinson & Sutton, 2001, p. 17) and the aim is to inform research and policy change with bottom up information. Hence the research questions are as follows:

1. How is the College English assessment policy change interpreted at an institutional level?

2. To what extent does the assessment practice that is implemented reflect the policy intent?
   a. What are the teachers’ understandings of the assessment policy change?
   b. How do the teachers relate their understandings to assessment practice?
c. How do the students view the implemented approaches used to assess their College English learning?

d. How do the students respond to the implemented assessment approaches of College English learning?

The policy of interest, its implementation process, and actors such as administrators, teachers and students in this study are viewed from a sociocultural perspective. To start with, the CERP assessment policy (2004) was officially developed and promulgated by the Higher Education Commission of Chinese Ministry of Education, the policy-maker at the highest level in the Chinese context. It is viewed in this study as a kind of “legitimating charter” (Levinson & Sutton, 2001), which has complex and social power (Ball, 2006). It was worked into a viable institutional blueprint at the university level and became more of a local and non-authorised form of the CERP policy within the localised context. The policy formation at this level is an appropriation of the meaning of the CERP policy from the top level, and the appropriation is enabled and constrained by the contextual situations of these universities as well as the framework of the cultural meanings of the policy-makers at the university level. The policy appropriation at this level is of critical importance to the policy implementation at lower levels such as the classroom. Therefore, its validity is a major concern (Levinson & Sutton, 2001).

Second, the policy process is regarded in this study as dynamic and interactive with cultural values and contextual situations. It is a process of recreation of the institutional policy itself as applied and interpreted by a multiplicity of local actors on a local and individual level. The complexity of the process is located as the focus of this study.

Third, local actors involved in this study such as institutional administrators, teachers and students are viewed as active policy meaning negotiators and constructors. The exploration of their understandings and views aims at revealing how teachers and students appropriate the meaning of the CECR assessment policy in their local context, as mediated by “cultural tools” such as cultural values, traditions and contextual situations. The exploration of the teachers’ and students’ practice seeks to reveal how the meaning(s) of the CECR assessment policy are shaped and recreated in these actors’ locally situated community. The sociocultural
perspective also directs the attention of this study to the cultural tools that the local actors use to interpret their experience and generate actions in their shared community of practice. The cultural framework embedded in Chinese cultural community of practice has been elaborated upon in Sections 2.5 and 2.6 of Chapter 2.

Informed by the focus of the sociocultural approach on localisation and contextualisation in policy studies, this study locates its research settings in local cases; hence a case study approach is taken. In light of the perceived difference across local settings in their appropriation practice (Levinson & Sutton, 2001), this study chose to use two university cases rather than one; and hence a two case study. Senior administrators, teachers and students in the both universities were approached as major informants to reveal local actors’ appropriation and practice in relation to the assessment policy change. Qualitative methods such as interviews and classroom observations were used, and complemented by the collection of artefacts such as written documents and students’ assignment samples. The intent was to generate a rich and detailed account of actions at the lower level. The specific methods and their usage are detailed below in Section 3.3.

This study acknowledges that the appropriation of assessment policy meaning is interacted by factors of three different levels: at the macro level of ideology and social context, and in the epistemology of assessment, and policy-making; at the meso level of institutional conditions; and at the micro level of classroom (Ecclestone & Pryor, 2003). In this study close attention has been paid to the meaning regimes. Particular attention is focused on the negotiation of policy meaning by educators and students at the micro level to account for their behaviours in the classrooms. It should be noted, however, that identifying key decision points at different levels is not a major concern of this study. Rather, this study embraces the sociocultural approach to policy study to highlight the interactions between individuals, communities and agencies and the way they shape their actions in relation to the others. Attention is paid to the cultural tools people use to negotiate policy meaning and generate social behaviour.

By looking at the CECR assessment policy as/in practice in the local and daily context, the primary concern of this study is not only to illuminate the effect and effectiveness of the policy, but also to seek to provide new perspectives on College English assessment policy and practice and its functions. The significance of this
study also relates to the fact that there is little evidence so far of a sociocultural approach to the study of educational policy in Chinese contexts. The sections below detail the case study approach and the specific methods this study employed, and the process it followed to achieve these intended purposes.

3.2 CASE STUDY APPROACH

This study adopts a case study approach and involves two cases. According to Stake (1995), method is not a choice by researchers; rather, it is a choice of the research itself. Yin (2003) succinctly states the choice of research strategies depend on: (a) the research questions posed; (b) the kind of control an investigator has over actual situations; and (c) whether the research focuses on contemporary or historical phenomena. A prerequisite then is to know clearly the type of research undertaken and the research questions to be addressed before choosing the methodology to be used.

This study aims to explore and document the complex interactions between the College English assessment policy and practice within two nominated cases of the particular sociocultural context. It is exploratory, interpretative and descriptive in nature. To examine the dynamics between the policy and its uptake, an in-depth investigation at both institutional and classroom levels is needed.

The rationale to use a case study approach in this study is relevant to the sociocultural perspective taken. It is also relevant to the nature and functions of case study itself. The case study approach allows for probing deeply into a situation or problem and its causes, rather than settling for a general surface knowing of the facts (Yin, 2003; Stake, 1996). Hamel et al. (1993), in a historic review of the definition of case study, identify “depth” as an important feature of this approach. Creswell (2008) expands this point in the definition: “a case study is an in-depth exploration of a bounded system (such as an activity, event and system) based on extensive data collection” (p. 476) and thus identifying “boundary” is a second defining feature of case study. Yin (2003) views case study as a research strategy, and defines it in a more comprehensive way: “a case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between the phenomenon and context are not clearly evident” (p. 13). Yin (2003) posits a case study approach as particularly useful in the field of social
science to conduct in-depth investigation into a set of events or situations in which the researchers have little or no control, and ideal when contextual conditions are expected to be “highly pertinent” to the phenomenon under research.

These characteristics are relevant to this study. First, to explore the interactions of policy and practice, inquiry into what is experienced by the teachers and learners through an empirical study rather than a theoretical approach has been chosen. Second, the topic under research, the College English assessment policy was adopted in 2007, and so is current and timely. Third, the policy incorporates elements of foreign origin. Therefore, the adaptability and suitability of the policy in this particular social, cultural and historical context of China is a major concern. Regarding the contextual factors as highly pertinent to the phenomenon under research, this thesis has adopted a sociocultural perspective as the theoretical framework. A case study approach allows me to examine Confucian Heritage Culture and the Chinese assessment tradition, in order to analyse the affordances and barriers that the context provides for policy implementation (elaborated in Sections 2.1, 2.5 and 2.6). Fourth, the control that I have over the actual situation is limited because there are over 2000 higher educational institutions scattered over the landscape of China. The sheer number poses a challenge for any research that attempts to investigate higher education related phenomenon, and hence leaves little possibility of control on the part of researchers, let alone an individual PhD student.

Case study research has a unique strength of dealing with the variety of evidence such as documents, artefacts, interviews and observations (Yin, 2003) which an in-depth investigation needs. It can incorporate either qualitative or quantitative research or a combination of both methods (Yin, 2003). This flexibility allows investigations of various natures to retain the “holistic and meaningful” characteristics of real-life events (Yin, 2003). In light of the flexibility, Stake (1995) suggests defining a case study by the interested phenomenon or relationship rather than by the methods used. Lamnek (1995) regards case study as a research approach rather than a method or methodology. This study, being exploratory, interpretive and descriptive in nature, requires a mixed method design to investigate both within and between the two cases. My position is that case study is an approach which utilises a variety of methods.
Concerns relating to the use of the case study approach include criticisms such as a lack of rigor and weak generalisability (Yin, 2003). For these reasons, a case study approach is not often used in scientific studies (Hamel et al., 1993). However, it is used widely in the social science domains because it offers valuable insights for in-depth understanding of real-life contexts such as classrooms (Hamel et al., 1993; Flyvbjerg, 2006). Thoroughly executed case studies can provide a scientific discipline with exemplars and hence strengthen its effectiveness and identity (Flyvbjerg, 2006).

According to Yin (2003), a case study is essentially composed of five important components: (1) a study’s questions; (2) its propositions, if any; (3) its unit(s) of analysis; (4) the logic linking data to the propositions; and (5) the criteria for interpreting the findings (Yin, 2003, p. 21). This study aims to explore the impact of the revised College English assessment policy as experienced by teachers and students. The research questions were clearly stated in Section 1.7 and again at the beginning of this chapter. This study does not make propositions. Rather, a more open approach is adopted to explore the research focus as deeply and widely as possible. Therefore, the second and fourth components do not apply in this study. Informed by the sociocultural perspective (Vygotsky, 1978; Wertsch; 1995; Rogoff, 1990, 1995), this study used “activity” as the unit of analysis, which specifically refers to local actors’ negotiation of the meaning of the policy change. Now I will discuss the selections of cases and participants.

3.3 SELECTION OF CASES AND PARTICIPANTS

3.3.1 Case definition and selection

A university is a typical unit within the higher education sector, and hence a micro system in itself. I investigated the interactions between policy rhetoric and institutional reality in two cases. The use of two cases in this study rather than one is due to the sociocultural concern of the variations in local settings (Rogoff, 2003). As well, it is acknowledged that a two-case study design is stronger in external validity compared to a one-case study (Creswell, 2007). As Yin (2003) asserts “[T]he evidence from more than one-case study is often considered more compelling, and the overall study is therefore regarded as being more robust, even if it is only two cases” (p. 46). A third reason for choosing a two-case design relates to the type of
cases in this study. Universities in China fall into two categories: key and non-key (or key and ordinary) (Zhou & Campbell, 1995). The design for this study is to have one Key University (KU) and one Non-Key University (NKU), hence a two-case design. The two cases in this design are type representative of each category (Yin, 2003). The key and non-key categories of university in China occur at both state and provincial levels. I accessed one KU and one NKU at the state level for investigation.

The criteria for selecting the two universities include the formative plus summative assessment model of CECR as implemented in these universities and their accessibility. As CECR (2007) provides guidelines rather than specifications, most universities are still in an exploratory stage. Accordingly the universities suitable for this study were narrowed down to those involved in the two pilot studies of CECR in 2004 and 2006 (see Section 1.4.3). It is anticipated that after a certain period of implementation, the demonstrated effects of and attitudes towards the integrated assessment model will have stabilised. Given the acknowledged regional disparity in degrees of development and the influence on education (Peng, 2001), I chose one university from Beijing, the capital city of China, and the other from a province in the Midwest part of China, which is under-developed according to the 2008 China Development Index (China Investigation and Assessment Center, 2008). The profiles of the two cases chosen for this study, along with their respective characteristics as key or non-key universities, are described in more detail in Chapters 4 and 5.

3.3.2 Senior administrator selection

In each of the two universities, I approached one senior administrator who is responsible for the College English teaching of their university for an interview. The purpose was to get an overarching understanding of the curriculum arrangements and changes made over the years. The senior administrator at KU consented to an interview, but the senior administrator at NKU agreed to have an informal talk only.

3.3.3 Teacher selection

This study invited four teacher participants (two from each case). Teachers’ knowledge and understanding about assessment (Rogers et al., 2007) are key variables that influence their behaviours in assessment practices. However, Cheng et al.’s (2004) survey shows that only a minority of tertiary EFL teachers in China have
received any specialised training by taking a complete course in testing or assessment. Most of the remainder have received little (in the form of workshop) or no training in this area. Thus, I did not use teacher’s training as criteria to select teacher participants. Nonetheless, I decided on teachers who gave lessons to first year students, because I learned from communications that the two cases had different courses in the second year.

Personal communication with the deans of the two selected universities revealed that the majority of their College English teachers belong to the age group of 25-45. Informed by this information, this study selected teachers between 25-45, who are among the majority and have experienced both the old and the new assessment models. It was anticipated that they would be able to offer valuable insights into the CECR initiatives. With the two criteria: 1) teaching first year students and, 2) 25-45 years old, I approached the deans of the departments concerned. On their recommendations, I contacted two teachers in each case and communicated with them about this study. They all showed interest and a willingness to participate. Also on the senior administrators’ recommendations, I contacted other teachers in each case for a focus group interview. Four teachers at KU and five teachers at NKU agreed to participate.

3.3.4 Student selection

I followed the participating teachers in their classrooms and observed the classroom interactions. Their students were invited to volunteer to participate in the face-to-face interview and to be observed. Students’ engagement is emphasised for the formative purposes of assessment (Broadfoot, 2007; Black & Wiliam, 2006), and is thus included in the current CECR assessment policy. Hence, this study used it as a criterion to select student interviewees. The academic year in Chinese universities starts in early September. In March, when the data collection of this study was scheduled, the teachers had taught the class for a whole semester and it was assumed that they should know their students very well. I selected student interviewees according to their participation in assessment activity as observed by me and evaluated by the teacher. Five students in each case made up the total of ten students for the in-depth interviews, which were conducted face-to-face and individually.
3.4 DATA COLLECTION METHODS AND TOOLS CONSTRUCTION

With a case study approach, this study aims to make an in-depth investigation of the interactions between policy and practice in the area of College English learning and the new types of assessment introduced by a change in policy. An intensive study needs rich sources of information (Merriam, 1998). The methodology and the case study approach of this study involve four types of data from three sources. The first data type is the relevant College English policy documents from the national and institutional levels. The second type is the audio data from interviewing the senior administrators, the selected teachers and students in the two cases. Data gathered from video-recording classroom observations, and artefacts such as the written assignment samples and the end-of-semester assessment tasks, provide further information. The written samples may provide evidence of teachers’ feedback and students’ responses to the feedback. These data are triangulated for validation purposes. Rich data from triangulation by source aims to substantiate the findings of a study (Yin, 2003). As such, the study aims to achieve more accurate and valid qualitative results (Oliver-Hoyo, 2006).

Table 3.1 details the links between the research questions, the target data, and the specific methods for data collection in this study.

Table 3.1
Link between Research Questions, Data Sources and Data Collecting Methods

<table>
<thead>
<tr>
<th>Research questions</th>
<th>Purposes</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>How is the College English assessment policy change interpreted at an institutional level?</td>
<td>Institutional responses</td>
<td>Documents Interviews</td>
</tr>
<tr>
<td>What are the teachers’ understandings of the assessment policy change?</td>
<td>Understandings Views</td>
<td>Interviews</td>
</tr>
<tr>
<td>How do the teachers relate their understandings to assessment practice?</td>
<td>Assessment practices</td>
<td>Interviews Observations</td>
</tr>
<tr>
<td>How do the students view the implemented approaches used to assess their College English learning?</td>
<td>Views</td>
<td>Interviews</td>
</tr>
<tr>
<td>How do the students respond to the implemented assessment approaches of College English learning?</td>
<td>Responses</td>
<td>Artefacts Classroom observation</td>
</tr>
</tbody>
</table>

As shown in Table 3.1, the main methods of data collection used in this research included interview and classroom observation, and were supplemented with
field notes and artefacts. Face-to-face interviews were conducted with the senior administrator, the participating teachers and the selected students in each university. Classroom observations were also made. In all, four tools were to be used. The development and application of the four tools are explained below.

### 3.4.1 Interview schedule for the senior administrator

Interviews are frequently used in qualitative research to elicit detailed information from participants (Creswell, 2008). They can be open-ended, semi-structured or structured, can be conducted with individuals or in groups, and can be conducted face-to-face or via telephone and computer. A semi-structured interview was conducted with a senior administrator in each university. The rationale for semi-structured interviews relates to the flexibility afforded to the subjects’ responses (Creswell, 2008). A semi-structured interview enables the researcher to capture the informants’ own accounts of their experiences and understandings of assessment while at the same time keeping the communication focused on the topic of interest (Higgins et al., 2003).

The interview schedule for the senior administrator focuses on major questions concerning the overarching curriculum arrangement, assessment policy in particular, of College English within the university. Changes made in response to the CECR (CMoE, 2004, 2007) and support and training provided to the teachers are also included (see Appendix A).

### 3.4.2 Interview schedule for the teacher

Semi-structured interviews were conducted with teachers in each university, firstly individually with the two individual teachers, and then with a focus group. The interview schedule for the teachers (Appendix B) was developed mainly through incorporating literature and interview protocols from Cheng et al. (2004, 2007). The protocols were adapted to apply to the key areas of the research questions. The interview schedule consists of four sections: demographic information, teachers’ understanding of assessment policy initiatives, teachers’ assessment practices as related to the new assessment policy, and teachers’ perceptions of assessment in the area of EFL.

The demographic section seeks information about gender, age, duration of teaching experience, assessment-related training received, and academic status.
Section Two focuses on teachers’ understanding of the current College English assessment policy. Specifically, this section investigated (1) how teachers understood the new College English assessment policy and its initiatives, (2) how they had been informed about College English assessment procedures, and (3) whether and how they were supported with training or guidance. Section Three relates to teachers’ assessment practices. Specifically, this section inquired about (1) the procedures followed, (2) the types of assessment task used to assess students’ learning, (3) the extent to which the assessment demands and criteria were shared with students, (4) the degree to which students were involved in assessment practices, and (5) the nature of the feedback provided to students. The last section of the interview dealt with teachers’ views of (1) the purposes of assessment, (2) the effects of formative assessment on learning, (3) the effects of the summative assessment on learning, (4) effective assessment, and (5) making compulsory external examination optional and its effects on students’ learning. With the above categories covered, the individual teacher interview gathered a detailed account of each teacher’s views on the College English assessment policy initiatives in relation to their own assessment practices. The interview took about 40 to 45 minutes.

The teacher interview schedule was piloted with two eligible Chinese College English teachers visiting QUT during the process of its construction. The questions were modified using simpler language and some explanations to eliminate possible misunderstandings and confusion based on the teachers’ feedback. For example, I added a question about “class size” to the demographic information section, as this was emphasised as an important inclusion by both teachers. As well, the piloting interviews took longer than expected, I accordingly adjusted the order of the questions to make them more logically linked and eliminated others that did not elicit the needed information.

In addition to individual interviews, a focus group interview was adopted in this study. Focus group interviews, which often lead to lively interactions between group members, are apt to generate richer data and are therefore used a lot in qualitative studies to provide parallel data (Barbour, 2007). It is used in this study to triangulate the data from other sources. The same teacher interview schedule was used in the focus group interview.
3.4.3 Interview schedule for the students

The purpose of the student interviews was to generate more detailed information for an understanding of the students’ views and responses to the CERP assessment policy change. It was also designed to explore the underlying reasons for the students’ participation in the assessment procedures in the classroom environment. The student interview schedule (Appendix C) includes 17 semi-structured interview questions. It was constructed via incorporating the key questions of Assessment Experience Questionnaire (Gibbs & Simpson, 2003) as relevant to this study. It was used in a face-to-face fashion in individual interviews with ten students selected from the two research sites.

3.4.4 Classroom observation

Observation is the process of gathering open-ended, first hand qualitative information by observing at a research site (Creswell, 2008). It is distinguished from interviews and questionnaires in that rather than relying on participants’ self report, observation provides direct evidence drawn from a setting as it occurs. As a result it provides an opportunity to study actual behaviours (Denscombe, 2007). For this study, the emphasis is on behaviours as practices, that is, the ways that teachers undertake assessment in their classrooms and the ways that students involve themselves in the assessment practice. I took the role of a non-participant observer and video-recorded three continuous sessions of the participating College English classes to observe the detail of the assessment practices in the classroom context.

An observation schedule (Appendix D) was developed to gather most direct firsthand data about the teachers’ and students’ assessment practices in the classrooms. The schedule was designed in a way as to “facilitate note-taking and retrieval of information” when needed (Angrosino, 2007 p. 58). It mainly focuses on teachers’ assessment practices: the assessment tasks and types used, instructions and feedback given, student involvement, and students’ assessment practices such as their responses to the instruction and feedback given, their participation in assessment activities.

The classroom observations were used to provide useful validation of the data from other sources. While the overall engagement of the whole class was observed, the focus was on the selected students and the teacher. Multiple observations were
conducted to overcome the initial awkwardness with an outsider in the room (Hammersley & Atkinson, 1995) and provided a more comprehensive picture of classroom practice. In this way, there is more possibility for obtaining the best understanding of the classroom practice (Creswell, 2008).

The data collection went through two phases. The first was a pilot study to refine and develop the data collection processes. This was followed by the main study that incorporated the data collection procedures. In the following two sections I outline the procedures and the timeline that I followed to collect the needed data for this study.

3.5 PILOTING

The purpose of piloting was to trial the interview and observation schedules to be used and then to refine them according to the feedback from eligible participants who completed and evaluated them (Creswell, 2008). Three tools for the collection of data were piloted. These included the teacher interview schedule, the student interview schedule and the observation schedule. The senior administrator schedule was not piloted because no other participants with similar qualifications were available. The main study included two sites. Only one of these case sites, the NKU, was involved in the piloting phase, as this study used the same criteria to select participants in both cases.

3.5.1 Teacher interview schedule

The semi-structured teacher interview had been piloted during the process of construction with two eligible Chinese College English teachers visiting QUT, as mentioned in Section 3.4.1. Moreover, the teacher interview schedule was also piloted at NKU before it was actually used. The rationale for this was the sociocultural perspective’s emphasis on context (Lave & Wenger, 1998; Wertsch, 1995). The two visiting Chinese College English teachers were from higher institutions of different types and both had the experience of studying abroad. I was unsure as to whether the teachers in the selected sites might provide different responses, given their different contexts and backgrounds. Accordingly, it was decided that it was necessary to pilot the teacher interview schedule with a sample of teachers who were more representative of the selected teacher participants in this study.
I piloted the teacher interview schedule with two College English teachers at NKU. They both conducted College English lessons to first year undergraduate students and both were in the age group of 25-45 years. They, therefore, met the participant selection criteria as specified in the research design. I intentionally chose two who were more senior and experienced in the target group, with up to 15 years’ College English teaching experience, anticipating that they could provide insightful comments on the soundness of the schedule. They both agreed to participate and chose to use English as the language for the interview. The pilot proved to be fruitful in three aspects. First, one teacher advised a reorganisation of some of the questions for better coherence; the other teacher advised the use of Chinese, the native language, rather than English for better elicitation of effective information. The advice from both teachers was adopted. Second, their responses to the questions gave me a better understanding of what the institutional College English assessment policy was like at this site and how the teachers were going about their assessment practices. Third, the interview itself gave me an opportunity to practice my interview skills and to familiarise myself with the interview schedule, which enabled me to do the study interviews more effectively.

For the purpose of piloting the other two tools, that is the student interview and observation schedules, I approached one teacher and one of her classes. The teacher met the participant selection criteria as specified in the research design. She taught College English to two classes of year-one undergraduate students and her age group was within the age group of 25-45 years. For the sake of convenience, I coded this teacher as NKU-T1. I used NKU-T1’s class A for the purpose of piloting the observation schedule.

3.5.2 Student interview schedule

I piloted the student interview schedule with four students from the same class. I selected two students who appeared passive as they had kept silent throughout the session I observed and another two who seemed relatively responsive. These more active students positioned themselves at the front (near the teacher’s desk) and they responded quickly and explicitly to the teacher’s questions. Their responses in the interview did not give me many grounds to alter the constructed student interview schedule. However, the communication did give me some understanding about how the students perceived the present College English assessment policy and practice,
and why they performed the way as they did in College English classrooms. Certainly, the interview itself gave me the opportunity to practice my interview skills and familiarize myself with the interview schedule, which enabled me to do the actual interviews more effectively.

3.5.3 Classroom observation

Originally, I had planned to focus both on the teachers (how they arrange assignment tasks and how they give feedback) and the students (how they respond). The actual videoing demonstrated that this plan was not realistic, especially when I needed to take notes as well. Consequently, I used the initial observation to select students to focus on in the subsequent observation and for the interviews.

The piloting was finished by the end of March 2009 with 16 visits to the university and multiple phone calls to arrange the times and places. The timeline for the data collection is detailed below in Table 3.2. Pseudonyms were used for the participants involved in the pilot. The limitation in this phase is that I piloted the data collection methods in one site. The financial and time constraints meant it was not possible to include the other site.

Table 3.2
The Piloting

<table>
<thead>
<tr>
<th>Stage</th>
<th>Method</th>
<th>Date</th>
<th>Participant</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Interview</td>
<td>January 01, 2009</td>
<td>Loy</td>
<td>Teacher interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>January 09, 2009</td>
<td>Fiona</td>
<td>Teacher interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 12, 2009</td>
<td>Jade</td>
<td>Teacher interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 14, 2009</td>
<td>Felix</td>
<td>Teacher interview schedule</td>
<td></td>
</tr>
<tr>
<td>Student interview</td>
<td>March 15, 2009</td>
<td>Lieu</td>
<td>Student interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 15, 2009</td>
<td>Young</td>
<td>Student interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 15, 2009</td>
<td>Yuna</td>
<td>Student interview schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 15, 2009</td>
<td>Wendy</td>
<td>Student interview schedule</td>
<td></td>
</tr>
<tr>
<td>Observation</td>
<td>March 09, 2009</td>
<td>NKU-T1’s class A</td>
<td>Observation schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 11, 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>March 16, 2009</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.6 MAIN STUDY

As mentioned previously, the research involves data collection in two university sites which are located in a Midwest province and Beijing respectively. I
outline below the data collection process at NKU first since it was done earlier. This is followed by the process used to collect data at KU.

### 3.6.1 NKU

The data collection in this site started in late March 2009 right after the pilot and continued through to the end of April 2009. It took four weeks and 13 visits. First, I started data collection with classroom observation. On my request, the teachers invited volunteer students for a face-to-face interview. I waited for a week, yet no student responded to my invitation. Therefore in the next week I approached five students whose performance in the classroom had attracted my attention as they appeared to be either passive or active. The students agreed to be interviewed, which from my point of view was related to cultural understandings and the students’ response to a teacher’s or authority figure’s request. I explain the related ethical issues in the ethics section (Section 3.9). Altogether, I interviewed five individual students at NKU. The students’ demographic information is shown in Table 3.3. Similar to the teacher participants, I coded them as NKU-S1, NKU-S2, NKU-S3, NKU-S4, and NKU-S5. The interviews were conducted in Chinese, the native language of these interviewees, and were translated and transcribed into English. Fifty pages of script resulted. I also asked the interviewees to give me some samples of their assignments. Four students gave me their assignment samples. I labelled and filed these artefacts.

*Table 3.3*

Demographic Information of the Students Interviewed at NKU

<table>
<thead>
<tr>
<th>Student</th>
<th>Gender</th>
<th>Age</th>
<th>Major</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKU-S1</td>
<td>Female</td>
<td>20</td>
<td>Arts</td>
</tr>
<tr>
<td>NKU-S2</td>
<td>Male</td>
<td>20</td>
<td>Arts</td>
</tr>
<tr>
<td>NKU-S3</td>
<td>Male</td>
<td>19</td>
<td>Science</td>
</tr>
<tr>
<td>NKU-S4</td>
<td>Female</td>
<td>20</td>
<td>Science</td>
</tr>
<tr>
<td>NKU-S5</td>
<td>Female</td>
<td>19</td>
<td>Arts</td>
</tr>
</tbody>
</table>

I observed and video recorded three class sessions of NKU-T1 and three sessions of NKU-T2 within four weeks of my visit in April 2009. I observed class proceedings and took notes. After the last sessions of observation of the two teachers’ classes, I interviewed the two teachers individually. I also conducted a focus group interview of another five College English teachers to generate more data.
about teachers’ understandings of the College English assessment policy in the university and their assessment practice in classrooms. The demographic information of this focus group is shown in Table 3.4. I coded them as NKU-T3 to NKU-T7. The interviews were conducted mainly in Chinese, with occasional use of English. The interviews were recorded, translated and transcribed into English as is the assessment related section of the video recording. This resulted in 68 pages of transcription in all.

Table 3.4
Demographic Information of the Focus Group Interviewees at NKU

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Gender</th>
<th>Age group</th>
<th>Prof. title</th>
<th>Years of College English Teaching</th>
<th>No. of classes and students</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKU-T3</td>
<td>Male</td>
<td>30-35</td>
<td>Lecturer</td>
<td>6</td>
<td>4 classes, 190 students</td>
</tr>
<tr>
<td>NKU-T4</td>
<td>Male</td>
<td>30-35</td>
<td>Lecturer</td>
<td>9</td>
<td>2 classes, 128 students</td>
</tr>
<tr>
<td>NKU-T5</td>
<td>Female</td>
<td>30-35</td>
<td>Lecturer</td>
<td>11</td>
<td>3 classes, 165 students</td>
</tr>
<tr>
<td>NKU-T6</td>
<td>Female</td>
<td>30-35</td>
<td>Lecturer</td>
<td>5</td>
<td>2 classes, 106 students</td>
</tr>
<tr>
<td>NKU-T7</td>
<td>Female</td>
<td>40-45</td>
<td>Professor</td>
<td>23</td>
<td>2 classes, 90 students</td>
</tr>
</tbody>
</table>

I also approached a senior administrator who was responsible for the College English Department for an interview. She agreed to have an informal talk rather than a formal interview. She also asked not to be recorded for fear of possible inconsistency between her and the top administrators of the school. Notes taken in the talk were translated and became part of the data set collected in this case. All the data collected in this university were organised into a database and stored in a space QUT has allocated specially for data storage.

3.6.2 KU

The data collection at KU started in early May 2009 and continued through to early June (the timeline is presented in Table 3.7). It followed the same procedures as in the NKU case.

As I did at NKU, I first invited volunteers for student oral interviews. Similarly, no students responded to my invitation. I therefore approached five students for the interview, whose performance in classroom attracted my attention because they appeared either passive or active. They all agreed to be interviewed. All but one of the interviews were conducted in Chinese and translated and transcribed into English. A girl chose to use English but switched to Chinese midway through the interview when she could not express herself freely. A 50-page transcription comprised this
data set. Apart from the interview, I asked the interviewees to give me some samples of their assignments. Two of them later gave me their assignment samples. I labelled and filed these artefacts.

At NKU, the same teacher gave both *Speaking and Listening lesson* (SL) and *Reading, Writing and Translating lesson* (RWT) to the same class of students, whereas at KU, College English lessons were presented by two different teachers. One gave the SL lesson and the other gave RWT lesson. Accordingly I also included the RWT teacher who gave College English lessons to the same class of students with the selected teacher. I observed and interviewed these two teachers and coded the selected teacher (who gave the SL lesson) as KU-T1, and the RWT teacher as KU-T2. Originally, I planned to observe four sessions of each teacher’s class. However, KU-T1 went to a conference for one session and another teacher’s sessions happened to be a traditional Chinese holiday – the Dragon Boat Festival. As a result, I ended up with three sessions for each teacher over the four weeks’ visit at this university.

The interview of the two individual teachers was conducted after the three sessions of observation. Also, an interview was conducted with a focus group of four teachers; their demographic information is shown in Table 3.5.

*Table 3.5*

Demographic Information of the Focus Group Interviewees at KU

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Gender</th>
<th>Age group</th>
<th>Prof. title</th>
<th>Years of College English Teaching</th>
<th>No. of classes and students</th>
</tr>
</thead>
<tbody>
<tr>
<td>KU-T3 (RWT)</td>
<td>Male</td>
<td>45-50</td>
<td>Assoc. Prof.</td>
<td>20</td>
<td>1 class, 40 students</td>
</tr>
<tr>
<td>KU-T4 (RWT)</td>
<td>Male</td>
<td>35-40</td>
<td>Assoc. Prof.</td>
<td>14</td>
<td>1 class, 40 students</td>
</tr>
<tr>
<td>KU-T5 (RWT)</td>
<td>Female</td>
<td>30-35</td>
<td>Lecturer</td>
<td>11</td>
<td>3 classes, 120 students</td>
</tr>
<tr>
<td>KU-T6 (SL)</td>
<td>Female</td>
<td>45-50</td>
<td>Assoc. Prof.</td>
<td>25</td>
<td>2 classes, 80 students</td>
</tr>
</tbody>
</table>

I interviewed five individual students at KU. The students’ demographic profiles are detailed in

*Table 3.* The five interviewed KU students were freshmen at KU, now in the third and last semester. They were all approximately 20 years, the average age for entry to the Chinese universities. Three were female and two were male. For ethical reasons I have coded them as KU-S1 to S5 to protect their identity.
To understand the assessment policy of the two universities better, I also approached a senior administrator in each university for an informal interview. The interview was conducted mainly in Chinese, with intermittent use of English. The KU senior administrator interview was recorded, translated and transcribed. During the NKU senior administrator interview, I took notes, which were then translated. The two senior administrator interviews resulted in 40 pages of English transcription.

The collected data at KU were organised in a database. The hardcopy were locked in a cabinet allocated to the researcher and the electronic version was stored in a specially allocated space on the QUT website. Table 3.7 presents all the data collected and the specific time when the data were collected.

Table 3.7
Timeline of Data Collection Procedures

<table>
<thead>
<tr>
<th>Method</th>
<th>Date (2009)</th>
<th>Participant</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>KU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom observation</td>
<td>March 26</td>
<td>NKU-T1’s class 50 students</td>
<td>Observation schedule</td>
</tr>
<tr>
<td></td>
<td>April 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>April 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>April 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>March 18</td>
<td></td>
<td>NKU-T2’s class B 48 students</td>
<td>Observation schedule</td>
</tr>
<tr>
<td>March 25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>April 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NKU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student interview</td>
<td>April 7</td>
<td>NKU-S1</td>
<td>Student interview schedule</td>
</tr>
<tr>
<td></td>
<td>April 8</td>
<td>NKU-S2</td>
<td>Student interview schedule</td>
</tr>
<tr>
<td></td>
<td>April 9</td>
<td>NKU-S3</td>
<td>Student interview schedule</td>
</tr>
<tr>
<td></td>
<td>April 10</td>
<td>NKU-S4</td>
<td>Student interview schedule</td>
</tr>
<tr>
<td></td>
<td>April 14</td>
<td>NKU-S5</td>
<td>Student interview schedule</td>
</tr>
<tr>
<td>Teacher</td>
<td>April 8</td>
<td>NKU-T1</td>
<td>Teacher interview schedule</td>
</tr>
</tbody>
</table>
### 3.7 VALIDITY AND RELIABILITY ISSUES

Validity and reliability are the criteria used to judge the qualities of academic research including case studies (Yin, 2003). Yin (2003) lists four criteria for case study research: construct validity, internal validity, external validity and reliability. Internal validity is not the concern of an exploratory study (Stake, 1996; Yin, 2003) as in this study. Accordingly only the other three will be discussed here.

Construct validity tests whether the operational measures for the concepts being studied are established correctly (Yin, 2003). This is a major concern for case study research because it is often a major flaw for which case study research is criticised (Yin, 2003). In this study, construct validity is ensured by following the two procedures suggested by Yin (2003). The first is to use the evidence from multiple sources: interview, observation, documents and artefacts, as detailed in...
Section 3.2 and 3.5. The triangulation of the multiple sources of evidence enhances the construct validity of this study. The second procedure is to maintain the chain of evidence in data collection and the research report so that the operational measures taken and the procedures followed can be traced.

External validity deals with the generalisability of a study, that is, the degree to which the findings of a study can be generalised (Yin, 2003). This is another major reason that case study research is undervalued, as case studies “offer poor basis for generalizing” when contrasted with large scale survey research (Yin, 2003, p. 37; Tashakkori & Teddlie, 1998). An argument against this criticism is that case studies rely on analytical generalisation rather than statistical generalisation as used by large scale survey research (Yin, 2003). The results from case studies cannot be generalised to a large population, but they are generalisable in a theoretical sense (Yin, 2003). In this study, two cases are conducted to generate understanding of the interactions between assessment policy and practices in two sites – a KU and a NKU. The multiple-case design is more robust in terms of external validity (Stake, 1998; Yin, 2003). Nevertheless, I understand the findings in this study are not to be generalised to the general practices of College English education throughout the country, though the lessons drawn from this study will be valuable for other universities.

Reliability is another criterion that relates to the quality of a research design. It refers to the efforts made to maximise the possibility that the same results can be obtained if the same procedures are followed. Specifically, the goal of reliability is “to minimize the errors and biases in a study” (Yin, 2003, p. 37). The study addresses this issue by documenting the procedures and describing in detail the specific measures. Further, the researcher documented the procedures and constructed a database to organise and manage every piece of evidence collected. This database is available for examination of the reliability of the study.

As the participants in this study are Chinese, English is a foreign language to them. Even if they are teaching or learning English, most of them chose to use their native language in the interview. Therefore, this study involves a validity issue in translation. This study utilised a “negotiation process” (Eco, 2004) to address this issue.
Eco (2004) defined translation as a process of negotiation “between all the parties involved, between the structure of two languages and the encyclopaedias of two cultures” (p.34). That is exactly how this study went about the translation. The researcher did all the translation on the grounds that she is proficient in both languages and she knows both languages and cultures well. Then, the translated transcripts were sent to the interviewees for validation or change as they saw appropriate. The translation then addressed the suggested changes. As the interviewed teachers are all English teachers in higher education institutions and students have gained reasonable proficiency after several years ELL, the participant check in this study served two purposes. The first was to validate the data transcription and the second, to validate the data translation. My three supervisors, who do not speak Chinese, however via feedback and constant questioning for clarification, acted as the readers and critics of the translation. The three parties (interviewees, researcher and translator, supervisors as readers and critics) were actively involved in the negotiation process.

The negotiation also related to that between two languages and two cultures. I acknowledge that despite the vast amount of time and meticulous efforts spent on translation, there may be instances when the meanings of the participants have not been accurately translated. This can occur when concepts and experiences are specific to one culture and language, and not another. To illustrate, the literal translation of the Chinese term 哑巴英语 (yābā yìngyǔ) into English is “deaf or dumb English”, which does not communicate the Chinese meaning of limited proficiency that many English learners in China have in oral communication. The meaning of some Chinese idioms and slang were also difficult to convey fully in English. A typical example is “假、大、空 (jiǎ dà kōng)”, which a teacher used to describe the training provided by the publishing house. Translation of it is “false, big and empty”, which again does not convey what is implied in the three-word colloquialism. Therefore, notes were added to explain the implied meaning so that all the parties involved were satisfied.

All interviews, which include two senior administrator interviews, four individual teacher interviews, two teacher focus group interviews, and ten individual student interviews, were fully translated and transcribed. The observation data were
viewed several times and only those relevant to teachers’ and students’ assessment practices were translated and transcribed.

3.8 ANALYSIS

The design as laid out previously incorporated elements of both interpretive and explorative studies (Yin, 1989). Accordingly, the data analysis was logically interpretative in nature. For an interpretive study, it is important to identify the perspective taken and the researcher’s self (Simons, 2009), which are foregrounded in Sections 3.8.1 and 3.8.2. The transparency of the data analysis process is regarded as critical to enhance the quality of qualitative research (Flick, 2007). Therefore, the following section (3.8.3) for transparency purposes is devoted to detailing the coding process, which was computer-assisted.

3.8.1 Sociocultural perspective & community of practice

This study has adopted a sociocultural perspective to underpin its framework and methodology (elaborated in Section 2.1 and Section 3.1). The sociocultural perspective emphasises the role of “the subjective knower in interpreting the known” (Guba, 1990, p. 17). This is the theoretical lens that I adopt to view the issue under research and also the perspective I adopted to interpret the collected data within its situated social, cultural, historical context. The sociocultural approach to educational policy highlights the notion that “community of practice often comprise the conditions under which policy is both formed and appropriated” (Levinson et al., 2009, p. 773). Informed by this notion, this study adopted an analysis approach of community of practice (Lave & Wenger, 1991; Wenger, 1998, 2000) and tried to explicate the complexity of the local actors’ practice through the lens of community.

3.8.2 Researcher’s self

According to Denscombe (2007), “the researchers’ identity, values and beliefs cannot be entirely eliminated from the process of analysing qualitative data” (p. 300). I admit the influence of my personal experience and social background in data interpretation. To start with, I was an English learner who had gone through the competitive assessment procedures with the Chinese traditional approaches to teach and learn English as a foreign language. As well, I was a College English teaching practitioner who has witnessed the struggles of my students to pass the exams and their limited English language proficiency even after they passed the examination as
required by the authority. And I did this study as a researcher who was sponsored by China Scholarship Council (CSC) to examine the College English policy change and the implementation and the local actors’ responses to it. More importantly, I was receiving professional research training in a higher degree program in a different cultural mode, which enabled me to view the issue under research as an insider and from outside (Evered & Louis, 1981). All these factors added to the sociocultural lens, with which I viewed and interpreted the data.

3.8.3 Coding & categorising using NVivo 8

The extended process of translation and transcription helped me to become familiar with the data, which was the first step in coding. The transcribed data, which include student interviews, teacher individual and focus group interviews, classroom observations, and artefacts, were then imported into NVivo 8, a software package developed by QSR International. I used it to facilitate the coding of the qualitative data. I was aware of the fact that this software is only a tool that assists researchers to work with data more efficiently and effectively; it cannot do the analysis itself (Bazeley, 2007). Bearing this in mind, I used this software mainly to manage and query the data. Moreover, rather than run the software to code automatically, I did the coding myself.

It is well acknowledged in the research community that qualitative data analysis is a complex process (Merriam, 1998; Dye, Schatz, Rosenberg, & Coleman, 2000; Simons, 2009). With two cases involved, this study calls for special attention to the differences and similarities between the cases. Constant comparison (Merriam, 1998), originated from the grounded theory (Glaser & Strass, 1967), was therefore used to capture the wide variations and complexities of the data. The process, which is facilitated with NVivo 8 software, is detailed below.

The coding process was, first of all, a descriptive process (Richards, 2005). The imported data were firstly coded as nodes, an NVivo expression of the basic codes. The coded nodes were, then, clustered and categorised in accordance with the emergent themes. The nodes, categories and themes were developed, which generated key themes as related to the research questions at the three levels: the institutional, the teacher and the student. For instance, the coding at the institutional level, as informed by the senior administrator interviews and collected documents, generated themes such as the institutional responses to the external test initiative and
the institutional responses to the formative assessment initiative. The coding at the teachers level, as informed by the teacher interviews (individual and focus group), classroom observation and artefacts, generated themes as related to the teachers’ views and responses to the two policy initiatives. The coding at the students level, as informed by the student interview, classroom observations and artefacts, generated themes as related to the students’ views and coping practices with the assessment policy changes. An example to illustrate this follows.

An excerpt from KU-T1 interview: I do not actually say what is bad about their (the students’) presentations ... for the sake of encouragement. This quote, which is about how KU-T1 gave feedback, is coded as KU-T1’s feedback practices. This code also incorporates excerpts of KU-T1’s oral feedback on students’ presentation as observed (reproduced in Snapshot One in Chapter 4) and his written feedback on students’ assignments (as shown in collected assignment samples Appendix G). Another excerpt, which is from the KU focus group interview, is: I usually give some nice words to make them (the students) comfortable (KU-T6). These excerpts were then, along with other KU teachers’ feedback practices as reported in the interviews, categorised as KU teachers’ feedback practices. KU teachers’ feedback practices and NKU teachers’ feedback practices were further clustered as a broader theme - teachers’ feedback practices.

The broad themes covered subthemes which related to the local actors’ specific views and practices. Teachers’ assessment practices, for instance, encompassed all the relevant data that related to the teachers’ specific practices in terms of use of criteria, student involvement, feedback, which are from interviews, classroom observations and artefacts. The themes and their subthemes formed a node tree of hierarchy (see Figure 3.1 for an illustration).
The coding process is also an analytical process (Richards, 2005). On coding according to the levels, themes from different levels as informed by varied data sources, were connected into categories. Feedback, for example, emerged in both teacher and student interviews, artefacts and observation data. These data were clustered under feedback category. The coding process however was by no means linear. Some themes, when found not closely relevant or exceeding the original category, were reconstructed. For instance, the policy category, which originally included both national policy documents, institutional policy documents, was later replaced. In its stead, national policy became one category, and the institutional policies, along with relevant data from the senior administrator interviews, were clustered under the institutional level category. The themes and categories were in this way identified and connected.

It needs noting that this was a constant process of going back and forward, which was described by Denscombe (2007) as “an iterative process” (p. 289) and Creswell (1998) as “a data analysis spiral” (p. 142). The use of NVivo 8 made this process more efficient. For example, when I realised that some teachers’ comments reflected implementation issues; I could simply move the relevant data which were coded as teachers’ comments to the implementation issue category. Figure 3.2 below is a screenshot of NVivo, which illustrates the interface between the programme and the user. The screenshot also clearly shows the frequency for each node and the source counts.
NVivo 8 also facilitated data query. This study involves two cases: accordingly, the data not only needed to be analysed within the cases, but also across the cases to identify the similarity and differences, and to elicit more significant findings. There was therefore a need for the data to be retrieved from both within and across the cases. The query function of NVivo 8 made it easy and efficient. Through the “query” feature, I could search for patterns and explore the data as I saw necessary. This feature is illustrated in Figure 3.3. Otherwise, the data retrieval and connection process would have been difficult and time-consuming. In this sense, the time spent on learning how to use the software proved worthwhile.

The use of software in analysing qualitative data, which is also referred to by the acronym CAQDAS (Computer Assisted Qualitative Data Analysis System), has the potential to increase the reliability of analysis since it allows the researcher to ascertain whether “the patterns reported actually existed throughout the data rather than in favourable examples” (Silverman, 2001, p. 228). Therefore, the use of NVivo 8 in the present study not only makes the data management easier but also has the potential to improve the quality of the analysis. The coding and analytic process is illustrated below in Figure 3.3.
Through such a process, the data from all sources were integrated to inform credible findings and further implications and conclusions.

To summarise, this study adopted a “congruent methodological approach” (Duggleby, 2005) to analyse the data gathered. That means it used the same approach to deal with the data from various sources such as interview data, observation data and artefacts, rather than adopted different approaches for different data sources. Congruency is regarded as a vital issue in qualitative studies (Simons, 2009). However, maintaining congruency between data sources does not blur the focus of different data types when care is taken. Rather, it helps to achieve coherence of research purpose, analytic method and reporting style (Duggleby, 2005). Moreover, a congruent methodological approach allows flexibility with regard to reporting of different data types. The interactions of focus group interviews and the directness of classroom observations in this study, for instance, were both maintained. The analytic process involves understanding complex relationships and connections between data sources. Dye et al. (2000) describe the constant comparison process as “a kaleidoscope of data”, which involves categorising data bits, comparing, refining and further refining categories. That is how this study went through in the process of data analysis.
3.8.4 Ethics

The necessity of human participation and information for the purpose of this study involves ethical issues. According to the *National Statement on Ethical Conduct in Human Research* (2007), the ethical requirements for this research belonged to the low risk category, which I applied for and which was approved in December 2008 with the code number 0000971. With this proof I applied for and was given permission to access the two participating sites on December 2008 and January 2009 respectively. I contacted two potential teacher participants as recommended by the dean for participation and they granted permission and acceptance to be involved. Before actual data collection, a teacher participant consent form, which states clearly the research problem and the objectives, the types of information to be collected and the confidentiality terms, was presented to the teachers and they signed it. Similarly, a student participant consent form, which clarifies similar information, was handed out to the selected students for their consent and signature. In the actual data collecting process, few students volunteered to participate. I approached in person the potential participants for agreement and showed them the ethics document, which is in English. With a concern to the possible limitations in their English proficiency, I explained the ethics documents item by item to them to make sure they understood the ethical issues before the interviews were conducted. Moreover, it was indicated to the participants that the information collected was to be used for this study only. Data would be anonymous and confidential and be stored securely. Participants would be free to withdraw at any time and contact details of the researcher and her principal supervisor were left with them.
Chapter 4: Case Study One: Key University (KU)

This case study of the Key University (KU) presents the implementation of the assessment policy changes initiated in the College English Reform Program (CERP) (CMoE, 2004). The case study addresses the research questions in this local context, focusing on responses to these changes at the levels of the institution, the teachers and the students.

Section 4.1 adopts a sociocultural perspective to document KU’s interpretation of the policy change and its appropriation of these changes. The following two sections detail the teachers’ (Section 4.2) and students’ (Section 4.3) lived experience of the assessment policy changes and their responses in relation to their prior experiences and perceptions. Section 4.4 summarises the findings of this case and brings the interpretation and discussion to a conclusion.

The Key University is referred to as KU for ethical reasons. Likewise, detailed references are not provided when citing from the university website or from articles published by the staff of the university. This order and style will be replicated in the Non-Key University case in Chapter 5.

4.1 INSTITUTIONAL LEVEL RESPONSES

Sociocultural theory emphasises the role of context in affording and constraining the local actors’ behaviours (Lave & Wenger, 1998; Rogoff, 1995; Wertsch, 1995). Understandings about the institutional level responses to the assessment policy changes in the CERP (CMoE, 2005, 2007) are derived from an analysis of data sources, which include an informal interview of a senior administrator responsible for College English teaching, documents collected from the university website, and personal communications with the teachers. This section starts with a profile of the local context (Section 4.1.1), which focuses on what makes KU a key university in the Chinese higher educational system. It is followed by the changes that KU made in response to the assessment policy changes (Section 4.1.2). The research question about the interpretation of the College English assessment policy change at an institutional level is addressed in the discussion (Section 4.1.3).
4.1.1 Profile

KU is a comprehensive university located in Beijing and has an enrolment of over 30,000 students. It is one of the top universities in China and has a prestigious reputation at home and abroad. Being a key university and located in the capital city puts this university in an advantageous position. Specifically, it is well supported in terms of finance, resources and policy, and attracts highly qualified teachers and high achieving students from around the country.

KU was included in both the “985 Program” and the “211 Project”, two major programs that the Chinese government launched in the late 1990s with the aim of driving 100 Chinese higher educational institutions to become “first class universities at a world level” (CMoE, 2004). Large amounts of national funding and policy support have been provided to these selected universities (CMoE, 1999). Locally speaking, Beijing was classified as a “level one” region according to the 2008 China Development Index (China Investigation and Assessment Center, 2008), which means it is among the most socially, economically and educationally developed regions of the country. Its investment in higher education accounted for 5.75% of the regional Gross Domestic Production (GDP) in 2005; in contrast, the educational input in all other provinces was lower than 3% (Treasury Department CMoE, 2008).

Many renowned scholars and top academics congregate at KU, adding to its prestige. Only the College English teachers will be introduced here as they are the key focus of this study. There are 50 teachers in the College English Teaching and Research Department. The teaching staff includes six professors, 20 associate professors, and seven lecturers with doctorates obtained at domestic or international universities. Nearly all of the teaching staff have studied or visited English-speaking countries for an extended period of time (Personal communication, May 8th, 2009). Their profiles on the university website detail extensive research experience in relevant fields. Together, they provide elementary (College English Band one to four) courses and over 20 advanced English courses at the university. In addition, the English Department recruits around ten foreign teachers, that is, English as native language teachers. Also owing to staff shortages, some higher degree students with English majors are employed when necessary on a casual basis to give elementary lessons.
The university enrols 3300 undergraduates annually who are primarily the top-scoring students on the national Entrance Examination of Higher Education (EEHE) from the 32 administrative regions. In 2008, the university claimed to have attracted 70% of the top 10 students from various provinces and regions around the country (source: university website). Around 3000 of these enrolled students have a specified area other than English as their major, and are hence regarded as non-English major students. They are the target group for College English at this university.

4.1.2 Changes to the institutional assessment policy

The current College English assessment framework at KU consists of an independent English proficiency testing system for gate-keeping purposes and an internal assessment framework for individual College English coursework. The changes in response to the policy initiatives in the CERP are now the focus.

The English proficiency testing system

The external test for College English (CET4/6) became optional and was disconnected from qualification requirements at KU in the late 1990s. This was several years before the China Ministry of Education (CMoE) (2005) advocated the change at a press release. In its place, KU developed an independent English proficiency test. The development of the proficiency test was conducted as directed by a policy document that the Higher Education Commission (HEC) of the Chinese Ministry of Education (CMoE) issued to eight top universities in China. The rationale for this directive related to the students’ above average results in the CET4 at these institutions. Originally CET4 was designed for all undergraduate students in China but it did not challenge those from top universities like KU. Pass rates of over 97% with distinction or high distinction were frequently reported. Despite the high pass rate on CET4, a majority of the students were found to be weak in speaking and listening. That is, “high score, low capability” (Wu, 1994) and “low efficiency and effectiveness” (Dai, 2001; Hu, 2002; Jing, 1999, 2000) was an issue for these students. Given this situation, these universities were authorised by the HEC to be excluded from the external test and to develop a test more suitable for their students.

The KU English proficiency test consists of an oral test to measure speaking and a written test to measure listening, reading and writing. Students need to pass both the oral and written tests. According to the senior administrator, the test uses
minimal multiple choice questions (MCQ) owing to perceived doubts regarding the validity of this question type. Currently, more than 70% of the test consists of short answer questions, translation tasks, summary writing, cloze, error-identification and correction. The employment of non-MCQs is also intended to avoid results reflecting test-wiseness. The test format, requirements and marking rubrics are publicised on the university website. The proficiency level in English that the test sets out to measure and its structure are regarded by both the senior administrator and the teachers in the interviews as much higher and sounder than the CET4. Its use was officially recognised by the HEC. Students can sit the test multiple times and gain four credit points when successful. For ten years it has served as a baseline qualification requirement for the graduates of KU. The testing system has undergone some major changes with the most recent change, having accommodated the changes in CECR (CMoE, 2007) and paying more attention to communicative ability.

**The internal assessment framework**

The internal assessment for individual College English courses comprises assessment conducted during the process of a semester, which was referred to as *process assessment* (guòchéngpíngjià 过程评价), and an achievement test at the end of a semester, which was referred to as *final term exam* (qīzhōngkǎoshì). The former is a literal translation of the term used by both the senior administrator and the teachers interchangeably for *formative assessment*. Grades are awarded to both the final exam and process assessment, which is combined for reporting purposes. The recent assessment changes are reflected in this framework (Figure 4.1) with the weightings for process assessment increasing in 2002 from 40% of the overall grade to 50% in 2004 and to 60% in 2008.

![Figure 4.1. Changes to grade distribution.](image)

The final exam is a textbook-based achievement test with over 50% of the testing material derived from the textbooks. The students need to study the textbooks meticulously and attend classes regularly to gain good scores. Process assessment is
composed of 20% for classroom performance and attendance, 20% for quizzes or assignments and 20% for online learning (Figure 4.2).

**Figure 4.2. College English internal assessment framework.**

Online learning, which was introduced at KU in 2008, is still in an experimental phase. It is not yet familiar to teachers or to students, who are accustomed to the traditional face-to-face classroom mode of teaching and learning. Assessment for online learning is being developed. It began as an assessment of the time that students spent online. But recently, spot checking or quizzes in class were introduced to check on students’ understanding, as it was agreed that time, as outlined above, was not a good measure of learning. Given these limitations this case study focuses on the other areas of process assessment which include attendance, classroom participation, and assignments and/or quizzes.

There are strict requirements regarding attendance. Three absences and the students automatically fail according to the institutional regulation. The senior administrator related the strict regulation regarding attendance to discipline and to the weekly College English class, which was designed to teach *how to learn* more than *what to learn*.

Syllabus and assessment arrangements are shared within each teaching and research group and communicated face-to-face at a meeting organised at the beginning of a semester. At the end of a semester, another meeting is organised to address problems and make changes if needed. Email contact is maintained between group members throughout the semester. All this is done to ensure that common criteria are maintained and principles are adhered to. Teachers have flexibility,
4.1.3 Discussion

From a sociocultural perspective (Levinson & Sutton, 2001; 2009), the College English assessment policy in practice at KU has involved an appropriation of the CERP assessment policy change (CMoE, 2007) in accordance with its particular contextual needs. This is demonstrated by a comparison of the above described College English assessment arrangement at KU and the prescribed policy change of CERP (see Table 4.1).

Table 4.1
A Comparison of Assessment Policy at National and Institutional Levels

<table>
<thead>
<tr>
<th>Assessment policy at the national level</th>
<th>Assessment policy at the institutional level</th>
</tr>
</thead>
<tbody>
<tr>
<td>CET4, the external test, made optional</td>
<td>CET4 no longer obligatory and disconnected from qualification requirements a decade ago</td>
</tr>
<tr>
<td>Formative assessment incorporated</td>
<td>Greater weightings allocated to process assessment</td>
</tr>
</tbody>
</table>

The two innovative initiatives prescribed in CERP (CMoE, 2004) have been implemented in ways that suit the local context. No external test is required, which aligns with the CERP initiative. The independent English proficiency test provides a baseline for the undergraduate students, which demonstrates the traditional and cultural value of the key role that examinations fulfil in education (Han & Yang, 2001). An emphasis on proficiency and an effort to avoid the negative washback of standardised tests (Han, 2002; Han et al., 2004; Niu, 2001; D. Wang, 2007) seem apparent. However, the aim to eliminate the social consequences (Gu, 2005; J. Wang, 2007) of external tests, which was intended in the policy initiative (see Chapter 1), is not clearly demonstrated in the institutional policy.

The CERP policy initiative of incorporating formative assessment was realised by the allocation of greater grade weightings to the process of learning. Formative assessment, interchangeably used with process assessment, was represented by records made about students' attendance, classroom participation and assignment tasks during the learning process. The results, referred to as “process grades (píngshíchéngjì 平时成绩)”, were used in the reporting at the end of semester. A shift of assessment focus from the product of learning to the process seems to be
indicated, but this does not necessarily mean that a shift from summative to formative purposes has occurred.

Feedback and student involvement have been argued as essential for formative assessment, and it is the particular uses and purposes that matter in assessment (Harlen, 2005, 2006; Newton, 2007). KU’s interpretation of formative assessment as process assessment reveals a possible misinterpretation of formative assessment to fulfil a reporting purpose rather than a learning purpose. Feedback and students involvement, both clearly articulated in the CECR, are not emphasised in the institutional assessment policy.

In terms of involving students in assessment, the senior administrator did mention a period when some teachers tried practising formative assessment by handing more control in class to the students whilst the teachers focused on the role of grading. This led to confusion and reconsideration about the feasibility of student involvement in assessment. It was concluded that the teachers’ directive role was indispensable for learning as well as for assessment. Nonetheless, without the inclusion of feedback and student involvement, the assessment practices appear to fulfil more of a reporting purpose than the enhancement of learning, that is, KU’s interpretation of the assessment initiative are more directed towards summative purposes. The influence of Chinese assessment traditions that emphasise the product or outcome, the summative function (Han & Yang, 2001), and the influence of teacher authority in the Chinese culture of learning (Biggs, 1996; Ho et al., 2001a) might be the attributing causes for the results.

The appropriation of the authorised national policy at the institutional level is crucial in that it constitutes the policy to be enacted within this particular local context, and prescribes the procedures for the translation of policy meaning into the classroom practice (Levinson, Sutton, & Winstead, 2009, p. 768). The findings related to how the policy has been interpreted at the institutional level hold major implications for teachers’ actual implementation of the policy in their classrooms. The intention to support and promote students’ English Language Learning (ELL) is iterated and reiterated in the teacher interviews and documents that were analysed. The teachers’ response to the policy changes now follows and adds to the complexity and richness of this case study.
4.2 TEACHER LEVEL RESPONSES

The investigation of the KU teachers’ uptake of the institutional assessment policy in classrooms is derived from an analysis of data sources which includes two individual teacher interviews and six observation sessions of their classes, a focus group interview and collected artefacts. It starts with the teachers’ profiles (4.2.1), which are followed by the major themes on the teachers’ perceptions (4.2.2), their understandings of the CERP policy change (4.2.3) and their accounts of any changes they have made in their assessment practices in response to the policy change (4.2.4). A discussion of the main findings (4.2.5) addresses the research question about the teachers’ enactment of the policy change.

4.2.1 Teacher profiles

Teachers’ backgrounds are important in framing the teachers’ perceptions about what English teaching and assessment should be like. They are also likely to influence their interpretation of the change and their teaching and assessment practice (Murphy, 2008; Rogers et al., 2007). The profiles of the teacher participants are presented from the data as either an individual teacher or as a member of a focus group.

KU-Teacher 1 (T1)

KU-T1, a male in his early 40s, started College English teaching in 1989 when he studied at KU for his Master’s degree, and then began his career as a College English teacher in 1991 upon graduation. He had experience of studying and teaching in England and the United States. On obtaining a PhD degree in Applied Linguistics in the USA, he returned to KU and resumed his College English teaching. His professional title is now Associate Professor and he gives Speaking and Listening (SL) lessons to a class of arts students. His background has positioned him as a confident and independent education practitioner.

KU-Teacher 2 (T2)

KU-T2 is a Master’s degree student majoring in Cognitive Linguistics. He was employed by the College English Unit on a casual basis to give Reading, Writing and Translating (RWT) lessons. He has no experience of studying abroad. He was included in this study because he gave lessons to the same group of students as KU-
T1. KU-T2 teaches and assesses under the supervision of a senior teacher. KU-T2 self-identified as a novice, and conceded “I am teaching mostly as I was taught”.

**Focus group**

The demographic information of the teachers involved in the focus group is presented in *Table 3*. As the table shows, all four teachers in the focus group have extensive experience of College English teaching, ranging from 11 to 25 years. Their professional titles vary from Lecturer to Associate Professor. They all have some experience of visiting or studying abroad. Three of them give RWT and one gives SL lessons to first year undergraduate students.

4.2.2 **Teacher perceptions**

Perceptions, established in prior experiences, are another identified means that mediates the teacher’s assessment practice (Rogers et al., 2007). It is important in this practice-related research to identify teachers’ perceptions to understand if a shift has occurred in their understanding of English Language Learning (ELL) and assessment (Broadfoot, 2007; Shepard, 2000). With the emphasis on formative assessment, there is also a need for a conceptual shift in terms of the relationship between learning and assessment which requires a change in role for both teacher and student (Dochy & McDowell, 1997; Shepard, 2000). The findings about the teachers’ perceptions are presented in order from their general views on ELL to the more specific themes related to assessment.

**English Language Learning**

Analysis of the interview data shows that the teachers saw the four skills of reading, writing, listening, and speaking as basic to English proficiency. However, they held different opinions regarding the importance of, and need for, emphasis on these skills. KU-T1, for example, held that the students had established a sound language base in their pre-university years. Therefore, their focus of ELL should be more about disciplinary content than basic language skills; they should be oriented more toward one’s future career than learning English for the sake of learning only. The basic skills and knowledge, nevertheless, should still be emphasised according to KU-T2, although he did not think “there is a general rule to follow” for how students could achieve these skills successfully.
The teachers’ views also differed as to how to facilitate English proficiency. KU-T1 viewed practice as valuable in ELL, which applied to both receptive skills like reading and listening and productive skills such as speaking and writing. “Doing is improving” and “doing is learning” were his typical expressions of ELL strategies. He also used the saying “practice makes perfect” to support this view. On the other hand, culture was valued by KU-T2 and he strongly advocated “learning English in its culture rather than taking it as a tool only”. A “sound cultural background” according to him could help students learn the language in such a way that it would encourage a cognitive process. It seems ELL in KU-T2’s opinion was not only about gaining a certain degree of proficiency; it was also about culture acquisition and cognitive development. There appears to be an influence from his major, Cognitive Linguistics.

The roles of teacher and learner in ELL

The teachers expressed a consensus regarding the role they played in students’ ELL. They viewed it as different from that of high school teachers, for whom the focus was mainly imparting knowledge, developing skills and preparing students for the EEHE. College English teachers, as they saw it, were more like guides, who facilitated and motivated students’ learning (KU-T1, KU-T2), and provided resources (KU-T1, KU-T4), learning opportunities and contexts for learning (KU-T4, KU-T5). KU-T1’s words below were representative:

… Let the students know that they can do this and that now, and provide some relevant resources … the teacher guides, assesses and checks to see whether they have done satisfactorily. If not, facilitates … and motivates them when they are not highly motivated.

By giving teachers a facilitating role and highlighting the how aspect of instruction, the students were positioned to take an active and central role in ELL. The notion that “English is not taught but learned” was mentioned both in the KU-T1 individual interview and the focus group interview. It was agreed that “it is up to the students themselves to make the differences” (KU-T1, KU-T4, and KU-T6). Another saying “one can bring a horse to water but they can’t make it drink” was quoted by KU-T4 in the focus group to support this point. The teachers’ understanding of the role definitions of teacher and learner in ELL appears to primarily align with the
constructivist paradigm (Vygotsky, 1978) that learning is a co-construction between the teacher and the student rather than a transmission from one to the other.

**The roles of teacher and learner in assessment**

The investigation of the teachers’ understanding of teacher’s and learner’s role in assessment reveals two views. One view, with KU-T1 as a representative, sees the teacher taking a leading role in assessment, designing the assessment tasks and the evaluation. KU-T1 also attached a cultural perspective to this view by saying “it is especially so in the Chinese classroom”. The other view, held by KU-T2 and KU-T4, values the role of teacher as an advisor. According to this view, teachers, based on their understandings of students and their expertise that allowed “them to see something that is invisible to the students at their stage” (KU-T2), can provide constructive feedback to “let the students know how to improve” (KU-T2, KU-T4). This view, highlighting teachers’ scaffolding function with feedback and its purpose to improve learning, tends to reside within the constructivist paradigm (Shepard, 2000).

The teachers’ views on the role of students in assessment are also divided. One view (KU-T3) holds strongly that to involve students in assessment was “not quite feasible”, and “both assessment and learning need to be done under the teacher’s guidance”; otherwise, “everything would mess up” because the students “all want high scores and they have got the skills and strategies about how to get it”. This view maintains that it might be “problematic if you let the students assess for each other”.

KU-T6, on the other hand, was positive about involving students in assessment. She believed that the students could be trusted and they had the capability to assess fairly:

… The students can do it very well. They give a very good score when the presentation is really good. They probably are reluctant to give very low scores even if the presentation is not quite good though. But overall they have very sharp eyes and open their eyes wide while assessing…

KU-T2 presented a different perspective. He saw democracy in involving students in assessment, and thought it could “motivate students more and make students aware of where they are”, but was conservative about it in practice because the “effects were uncertain” and the procedures were not yet established. He
regarded the teacher as in a better position to observe and assess “fairly and objectively”. The mixed nature of these views indicates that some teachers may have adopted a role shift in their understandings of assessment but some have retained the teachers’ authoritative role in assessment in their understanding (Hu, 2002).

**The purposes of assessment**

The teachers’ understandings about the purposes of assessment are again divided. KU-T1 saw measurement as a major, if not the only, purpose of assessment, designed to measure “how well the students have learned” and “how well you teachers have taught”. This view resonated with the focus group, which regarded assessment as “a necessary part of coursework” that ensured the course requirements were met. Furthermore, the focus group held that objective measurement of students’ learning was “indispensable” for large student populations in courses such as College English. They related the reason to the strict accountability demands from the administrative sectors.

KU-T5: The objective part of assessment is definitely needed.

KU-T6: The school will check on the details of assessment. Not a bit should be missing. Without tests? That would be a total mess-up.

KU-T2 similarly highlighted the necessity for “objectiveness” but also acknowledged both measurement and learning improvement functions in assessment. On the one hand, he regarded assessment as aiming to “give an objective description of students’ learning with an established and accepted system”. On the other hand, he also noted that assessment should be able to inform learners about “what their learning is like and where they are” and he valued the “constructive advice” that identified and addressed the students’ learning problems. The overall assessment, as he saw it, was largely product-orientated and designed to direct students towards the ends rather than the means. To combat this orientation, he suggested teachers balance students’ learning and testing needs so that “the students could focus on learning, without worrying too much about the test”.

It seems that objective measurement still dominates the teachers’ understanding about the purposes of assessment which aligns it primarily with the administrative system. The learning enhancement function is acknowledged by some teachers, but appears related only with teaching and learning practices rather than assessment.
This finding seems to indicate that the educational system is a factor influencing the teachers’ understanding about the purpose of assessment and possibly their assessment practice as well.

**Formative assessment**

The interview data reveals that the teachers’ perceptions of formative assessment were closely related to their training in the relevant field. All the teachers except KU-T1, who had taken a course on testing for his Master’s degree, reported to have no training in assessment. They acknowledged that their understandings about assessment were primarily related to examinations and tests. There was no formal training about formative assessment before the implementation of the policy change. When asked about formative assessment, both KU-T1 and KU-T2 asked the researcher to clarify the concept, and the focus group teachers conceded they did not have a theoretical understanding about it. Rather, they knew it by “tacit knowledge”. This was represented below.

Researcher: How do you perceive the concept of formative assessment?

KU-T3: Simple introduction.

KU-T5: No perception of theoretical level. Assessment is really not our major. We come in touch with it mostly when it comes to examinations…

Researcher: I mean how you perceive formative assessment.

KU-T5: Not quite clear.

KU-T6: No theoretical understanding.

Researcher: Could you just tell me how you understand it?

KU-T3: We know it by tacit knowledge.

The teachers tended to relate formative assessment to when it was conducted, that is during the learning process as opposed to a final summation of student performance. The various process tasks were seen as formative. To illustrate: “I used quizzes, which is kind of formative assessment” (KU-T1). The teachers largely took the institutional interpretation of formative assessment as process assessment.

KU-T1: It [formative assessment] was quite little in the past.

Researcher: Oh?
KU-T1: [It was] 10% for the process, the rest all went to the final term exam.
Now it is around half to half.

These perceptions will influence the teachers’ interpretation of the policy change and their enactment of it in practice, which follow in the section below.

4.2.3 Teacher understandings of the policy change

This section examines how the teachers understood the assessment policy change of making the external test optional and incorporating formative assessment. The understanding is that the ways the teachers interpreted the meaning of the policy closely relates to their uptake of it in practice (Murphy, 2008).

**External test becoming optional**

An analysis of the teachers’ responses confirms the institutional assessment policy change that has been outlined previously, such as the optional nature of CET4 since the late 1990s and the use of an independent English proficiency test as a qualification requirement. The teachers attributed this policy change to students’ varied English proficiency levels in different institutions. For example,

CET4 was originally designed for the undergraduate students nationwide. The students in some top universities stand out considerably from the average level … That means, CET4 is no longer a proper measurement for these students; it kind of undervalues them … I think making the CET4 optional is to take into consideration the individual differences of various universities (KU-T1).

The differences between the English proficiency test in use and the CET4 were also well understood by the teachers, who agreed that their self-developed test was valid given that it included very few items “that can be guessed” (KU-T1, KU-T6). However, the teachers acknowledged that this policy change had minimal impact on their practice. KU-T1, for instance, had experienced a shift away from testing-oriented practices that marked his early years of College English teaching, but he denied any relationship between his teaching practice change and CET4’s optional status. Rather, he attributed this shift to his realisation that “English is not to be taught to or learned for testing”. KU-T2 claimed he never taught to the tests, regarding preparing students for testing “a dumb thing”, which could not possibly “bring about noticeable results” within the limited time, even with the test-wise skills. Improving students’ English proficiency was thus the focus of his teaching and he
attributed this to his perception of ELL and the use of proficiency testing rather than the practice of optional CET4 at KU. Overall, learning or proficiency appeared to take a focal position in the teachers’ teaching practice. This is also visible in the teachers’ feedback focus, which will be discussed later under the theme of Feedback.

Similarly, the impact of the policy change on the students was scarcely perceived by the teachers. First, according to the teachers, most of the students would still take CET4 despite the optional practice. The teachers attributed the situation to “social” reasons, which is explicated below.

… Our students seem to show great interest in taking the CET4. The reasons are social. When it comes for them to seek employment in the domestic market, it is the test results in CET4 rather than the test results in the institutional proficiency test that speak louder (KU-T1).

Second, the teachers believed students’ learning behaviours were closely related to their goals. They saw three goals in the KU students’ learning of English. The first was the Graduate Record Examination (GRE) or Test of English as a Foreign Language (TOEFL). The students who had this goal aimed to study abroad. The second was practical use of English for their future careers, which was highly probable for graduates from key universities like KU, especially in this globalised era. The third was to be a qualified graduate of KU, which featured strengths in Information Technology (IT) and English. Given that most of the students were highly motivated to succeed in areas other than just passing CET4, making the external test optional did not seem to affect their approaches to learning English.

**Incorporating formative assessment**

The teachers confirmed the incorporation of process assessment within the institutional assessment framework, but understood this change differently. Some teachers such as KU-T1, KU-T3 and KU-T5 reported to have witnessed the variation of the ratio distributions between process assessment and the final exam over the years. They also indicated that by introducing process assessment, more professional control had been handed to teachers, which in turn made it possible for teachers to conduct assessment rather than solely rely on the administrative body. In contrast, the change was not thought as substantive to some teachers. KU-T6, for example, said: “[The institutional policy stays] roughly the same. I do not see much change over these years” and claimed to have been doing assessment as she thought it was
best. KU-T5, who had seen the ratio variations, iterated: “we have been doing [assessment] like this all along. Only it is now put forward and regulated in documents”. KU-T4 cited some visiting guests from other universities to claim: “the assessment framework regulated in the new requirements (CECR) is quite similar to that of our school”. They unanimously denied that the incorporation of process assessment had exerted an obvious impact on their assessment practices. It seems that designating larger weightings of assessment to the learning process had not led to changes that were recognised by all the teachers.

Despite their different understandings of the change, the teachers agreed upon the rationale for incorporating process assessment as intended to “do assessment across the learning process” rather than conducting it only at the end of the learning period (KU-T1, KU-T2, KU-T3). This change, according to the teachers, had the potential to enhance learning. To start with, it could be an effective means to reduce absences from class since the students needed to attend each session to gain their process grade (KU-T1, KU-T3). In addition, it could help reduce pressure on students (KU-T3, KU-T5), who were usually keen to achieve high scores. “Knowing they have done well during the process, they do not need to worry too much about the final term exam” as KU-T5 said. Moreover, it could make assessment more “objective” given that the evidence collected from multiple sources at different intervals helped to build a comprehensive list of assessment results that reflected the students’ actual achievement (KU-T2, KU-T4). Further potential from the incorporation of formative assessment included helping to prevent the use of test-wise skills (KU-T1), sustaining the degree of students’ effort throughout the learning process (KU-T1, KU-T2), and keeping the students motivated (KU-T1, KU-T2, KU-T3). All these were considered to be positive for learning.

The above analysis suggests that the teachers understood the formative assessment initiative more in the light of its administrative functions. The improvement of learning function was recognised, but primarily from the view of increasing students’ motivation and encouraging students’ efforts in the learning process. The important aspects of formative assessment - feedback from teachers and the involvement of students - appear to be missing. A close relation seems to exist between these understandings and the teachers’ perceptions, particularly their understandings of the purposes of formative assessment.
### 4.2.4 Assessment practices and procedures as enacted by the teachers

From the teachers’ assessment practice-related data themes surfaced about what and how assessment was practised, which includes measures, criteria, student involvement, feedback and grading practice. The focus of the analysis in this section is to detect how, if any, formative assessment was enacted in the classrooms.

**Alignment with the institutional policy**

The procedures that the teachers reported to have taken for assessment during the process are attached in Appendix E. To start with, classroom participation was reported by three teachers; attendance was adopted by all the interviewed teachers. The teachers also followed the institutional regulation about assignments, with all of them assigning tasks and/or conducting quizzes. Online learning, which was stipulated in the institutional assessment policy, was not reported by the teachers. This is perhaps owing to its experimental status and hence it was not considered in this study. Overall, the assessment procedures that the teachers adopted in the classroom aligned with what was framed in the institutional assessment policy. This was acknowledged by the teachers, who stated they practised assessment mainly in accordance with the prescribed assessment arrangements, especially in terms of the assessment procedures and the distribution of grades. KU-T1’s assessment arrangements illustrated in Figure 4.3 is a further support to this claim.

![KU-T1’s grade distribution](image)

**Figure 4.3.** KU-T1’s assessment arrangements.

Analysis of how these assessment procedures were understood and used by the teachers is expected to shed light on the assessment as practised. Attendance, for example, which was adopted by all, was understood differently by the teachers. KU-T2 and KU-T6 viewed it as a matter of attitude, which was important for learning in general. KU-T1 thought it showed students’ respect for teachers. Further KU-T3
perceived it as a matter of discipline, and especially important for those students who cannot discipline themselves. The teachers also reported to check attendance in different ways. KU-T2 and KU-T5 called the roll while KU-T3 had students sign their names. They tended to do it especially when “not many show up.” KU-T1 used the quiz papers and quiz results, and KU-T4 used the writing assignment, both weekly, as a form of checking. Absences meant there were no grades for this particular assignment, which would eventually affect students’ overall grade at the end of the course. KU-T6 knew her students very well and checked their attendance by calling each student to answer questions. Regardless of the differences in checking practice, “absences of twice and more would lead to downgrade of final assessment results or even failure”. This was clearly stated in KU-T1’s syllabus. That means attendance as an aspect of assessment was actually used as a means for control of student behaviour and for punitive purposes.

Classroom participation was reported by three teachers, who mainly looked at how the students responded to their questions (KU-T3, KU-T6) or involved themselves in learning activities (KU-T1) in class. Nonetheless, they all admitted the limitations of this measure in terms of objectiveness and exactness. KU-T4 for example stated: “Only those who are particularly good or particularly bad are observable. The middle ones I can’t tell”. Hence the teachers tended to use it as a minor rather than major means for assessment. Flexibility and manoeuvring were also evident in the teachers’ reports. KU-T3, for example, commented: “We say some percentage, but actually it is not quite exact. I use it as kind of reference. If the students’ final score is near to fail. It can be very useful”.

In terms of assignment tasks, the teachers conceded that given that the institutional assessment policy only provided a guideline rather than specifications, they had used the flexibility allowed and adjusted it accordingly to suit their students. The two SL teachers, for example, adopted totally different assessment tasks regardless of the shared syllabus. KU-T1 relied on quizzes to check if the students had attended to the listening materials in the textbooks. In contrast, KU-T6 tended to favour recitation to check students’ mastery of knowledge in the textbooks. KU-T1 used presentations to encourage students to present in a professional way in class, whereas KU-T6 chose to have students take turns to recite or answer questions as a kind of speaking drill. Even the same assessment task type was used in different
ways by different teachers. Presentation, for example, was used by five out of the six
teachers across RWT and SL courses. Yet, KU-T1 and KU-T3 gave students the
freedom to choose the topic of their interest, whereas KU-T2, KU-T4 and KU-T5
asked students to present on a given topic, which was relevant to the theme of the
units or culture in general. Most teachers let students present individually, whereas
KU-T2 used group work. The time allocated for presentations also differed and
ranged from three to six minutes in different teachers’ classes. The type of task to use
and how to use these appear to be primarily individual teacher’s decisions.

The teachers’ interviews also confirmed that prior notice was issued in the case
of assessment policy change; however, the ratio distribution between classroom
assessment and the final term exams differed between groups and teachers. The four
teachers in the focus group, who belonged to two different teaching groups, gave
different ratios. The RWT teaching group used 40% to 60%, whereas the SL teacher
KU-T6 reported a 50% to 50% distribution, which, according to her, was unilateral
within the group. It is worth noting however that KU-T1, also a SL teacher, did not
adopt the group distribution (see Figure 4.3). Hence, it is justified to say that
although the teachers generally followed the assessment procedures suggested in the
institutional assessment framework, they also interpreted how to implement them
differently. The following three sections look closely at how the teachers’ assessment
practices in the areas of criteria use, student involvement and feedback.

The use of criteria

Most teachers stated that they used similar criteria when assessing tasks. However, some teachers made them explicit while others preferred to keep them implicit. The two individual teachers for instance both claimed to have used criteria, which KU-T1 referred to as “preparation, effects and language” and KU-T2 identified them as “preparation, presenting and the audience’s response”. Their further explanations indicated that the three criteria were virtually the same irrespective of the different wording. KU-T6 in the focus group reported similar criteria, as can be seen from the excerpt below:

KU-T6: You know it is good when it is good.

Researcher: Can I say there are no specific criteria?
KU-T6: We do have. The presenter needs to present very good content with fluent English. Good content without good speaking ability won’t win any applause.

However, these criteria appeared to be implied rather than articulated clearly. The implied nature of criteria was echoed by KU-T4, who stated that he judged assignments mainly according to his “tacit knowledge”. KU-T3 too used content and language as criteria. He differed from the others; he did not think students’ responses to the presentations a reliable criterion since his students “would applaud for politeness or encouragement’s sake.” The similarity in criteria was mostly related to the teachers’ expertise and tacit knowledge about what a good presentation is.

The data analysis also reveals different means of sharing criteria. KU-T1 reported to have done a presentation for students about “how to do a good presentation” and explained the requirements for a good presentation beforehand. He saw this as training and a sharing of the criteria. Different from KU-T1, KU-T2 reported to have communicated these criteria orally only to students in class before the task type was introduced. The teachers who chose to keep the criteria implicit assumed the criteria were understood by the students.

**Student involvement**

Student involvement in assessment appears in three places in the data. The first was peer response used as a criterion for the presentation task. The other two both related to KU-T1, who reported to have negotiated the assessment arrangements with the students and introduced a writing blogging program in the previous semester.

As elaborated in the section above, three teachers claimed to have used audience, that is, other students’ responses as one of the three criteria for judging the presentation task. However, the use was primarily limited to the “effect” of the presentation, which was indicated by the amount of applause rather than students’ comments in relation to explicit standards. This limits its validity given that politeness may intervene (KU-T3). This doubt was proved reasonable by the observation data (see Snapshot Two), in which applause occurred both before and after the presentation, and was given to each presenter. It seems the use of students’ responses as a criterion did not mean that students had been actually involved in assessment.
KU-T1’s first activity that involved students in assessment is illustrated as:

… I make sure to share the syllabus and all the assessment arrangements with them and ask for their opinions about it beforehand. I would make adjustments according to their opinions. But it will be carried out strictly as class law once agreed upon.

The excerpt seems to indicate that students have been involved in negotiating the assessment arrangements.

KU-T1’s writing blogging program was another activity that related to student involvement in assessment. He encouraged peer review so that the students wrote and read from a writer’s as well as a reader’s perspective. This suggests the students were quite involved in assessment in this program. However, primary data about the program were not available because it had ended when the data were collected. However, it will be discussed further in the student section.

Student involvement in assessment was not reported or observed elsewhere. KU-T1 related his practice to his prior experience:

I learn from my teaching experience that students are scarcely motivated during the process if only the final term exam is used. The other influence may come from my experience in USA. They do it this way.

This seems to indicate that the teachers’ prior experience influenced their decisions about whether to involve students in assessment.

**Feedback**

The data revealed that some teachers provided comments to students’ assignment tasks as feedback. Quite a few insights emerged about the feedback. First, the comments were brief in most teachers’ cases but not for KU-T4, who reported giving “very detailed” comments for students’ weekly writing assignments. Second, the comments were general for some teachers but there were teachers who reported to have attended to individual student’s needs and given constructive feedback. KU-T4, for example, reported that his detailed comments told the students “what was good about the good ones and what to improve for the bad ones.” Third, most teachers gave comments in written form and mostly only to those who performed well, but there were exceptions. KU-T3, for instance, reported that he preferred to point out problems face-to-face, especially for those students who “did very poorly”.

Chapter 4: Case study one
KU-T6 too stated that if she found the students had problems, she would “definitely let them do it again.” She attributed the rationale to her view that attitude was the most important. Fourth, the comments were more positive than negative, and the purpose for encouragement was highlighted. This was evidenced by KU-T1, who articulated: “… I do not actually say what is bad about their presentation … for the sake of encouragement”. The same sentiment resonated with KU-T6, who noted: “I usually give some nice words to make them comfortable” and shared by others such as KU-T2 and KU-T5. KU-T1 attributed his value for encouragement to the purposes he designed for assessment tasks. Presentation for example was intended “more as an opportunity for [the students] to summon up their courage to talk in public rather than as an activity to assess if they could do it well or badly”. Last, the comments also appeared to relate to the teachers’ perceptions. KU-T2, for example, reported that his feedback on writing assignments focused mainly on grammatical errors. He regarded the key to feedback as residing in raising student awareness “about their grammatical problems.” KU-T2’s personal valuing of grammar in ELL was now a central focus for his feedback practice. Other factors that were reported to constrain the teachers’ feedback practice included time and class size, which related to a heavier workload. Overall, the comments provided by the teachers were diverse and reflected individual teachers’ decisions. Feedback for formative purposes was visible, for example, in KU-T4’s detailed and constructive comments and in KU-T1’s oral feedback to the presentations (sees Snapshot Two). In a sense, it can be said that the teachers were incorporating formative assessment in their practices, even though they might not be aware of it.

Grades

All the teachers reported grading and making records of students’ performances on assigned tasks (a record sheet by KU-T1 is attached in Appendix H). The grades were usually categorised into five bands from A to E, which appeared to be determined tacitly rather than according to specified standards.

The teachers’ approaches to grades were diversified. The grades appeared high and undifferentiated in some teachers’ cases. KU-T1, for example, stated that he gave “almost all of [the students] full scores except when they presented it extremely badly. But even so I give a B.” He stated that his rationale was to be encouraging. KU-T2 too said he would not “baffle them”. Other teachers took different approaches.
KU-T4, for example, reported giving “roughly the same” grades for most of his students but not as high as near full scores. Thereby, he intended to direct the students’ attention to the comments rather than the grades. For KU-T6, the grades were usually high but differentiated. According to her, it was easy for students to get very high scores with about half going to the process grade. However, she thought “it is not alright to give all the students 90 points or above”. She insisted on giving some 78 or 80 points and admitted that she “deliberately [brought] down the scores a little bit, for their good”.

Regardless of the diversified grading practices, the resultant grades were used invariably as part of the process grades and integrated into a final grade. The final result aligned with the system of allocating grades up to one hundred points and was used for reporting purposes. This practice was shared by all the teachers and largely accorded with the institutional assessment policy.

**Snapshots of assessment practice in classroom**

Two snapshots of KU-T1’s assessment practice in classroom are reproduced to illustrate how assessment was practised. The first shows how he undertook a quiz, and the second demonstrates his use of criteria, student involvement and feedback to students in a presentation task.

### Snapshot One

**Time: 10:00 – 11:30am (Beijing time), 12th May, 2009**

KU-T1 asked students to produce a piece of blank paper and do a quiz. Before he started, KU-T1 made brief comments on the quiz they did last session:

> I am very glad to tell you that we got 36 “A”s last time. That is really, really good.

KU-T1 displayed the exercise on the monitor screen (attached in Appendix E). He took two minutes to explain to students the instructions of the exercise and new words, *symptom* for example. Then KU-T1 asked students to close their textbooks and put on their headphones before he played the audio recording on the computer. Students listened and completed the questions and handed them in. The process took six minutes altogether.
This is one of the three quizzes observed in the three observation sessions conducted at KU-T1’s class. The same procedures were followed each time. A quiz paper sample with teacher’s comments was collected and is attached in Appendix G. KU-T1 made it clear in the interview that the quiz was conducted to spot check if and how the students had finished the assigned listening tasks. Comments, if given, were for encouragement mainly. No constructive advice for improvement was provided. Analysis of the quiz practice with the interview, observation data and the collected artefacts show that this assessment practice, though done progressively throughout the learning process, was not formative in function. Rather it appeared to take the form of a series of mini-tests that served summative purposes.
KU-T1 started with two questions to elicit other students’ responses about the presentation content. That was followed by a word of encouragement to the presenter. Then KU-T1 put forward three points for the presenter’s as well as other students’ attention, which involved time limits, pronunciation and presenting skills. He also mentioned the possible influences on what he called the effect of the presentation. These areas corresponded generally with those reported in the interview to assess by: preparation, language and effects. These comments were targeted at issues that emerged in the presentation: Xing, the presenter, had exceeded the time limit and

---

**Snapshot Two**

Time: 10:00 – 11:30am (Beijing time), 12\(^{th}\) May, 2009

KU-T1 started the lesson with a brief introduction of the session content and then introduced the first of four scheduled presentations.

KU-T1: Now, it is Xing’s turn for his presentation. Let’s welcome.

Teacher and students applauded. Xing went up to the teaching podium and presented the PowerPoint presentation on the musician Mozart. The presentation lasted for 10 minutes. Students applauded again on Xing’s finishing. Then KU-T1 started to give comments orally.

KU-T1: Thanks, Xing. (Then turn to the class)
KU-T1: His topic is …?
Students: Mozart.
KU-T1: Right. Now see if you got this information. At what age did Mozart begin to perform? Three, eight?
KU-T1: Five! Yes, that’s right. (Turn to Xing) Ok, thanks Xing, a very good presentation.

KU-T1 commented further on while facing the whole class:

We learned a lot about a great musician, Mozart from this presentation. However, a few things need your attention. Firstly, watch your time. Secondly, pay attention to your pronunciation. Check the pronunciation that you are not sure about. And please rephrase the materials in your own words rather than quote directly and read from the screen. Otherwise, people will just follow the screen rather than listen to you, which will influence the effects.

KU-T1 started with two questions to elicit other students’ responses about the presentation content. That was followed by a word of encouragement to the presenter. Then KU-T1 put forward three points for the presenter’s as well as other students’ attention, which involved time limits, pronunciation and presenting skills. He also mentioned the possible influences on what he called the effect of the presentation. These areas corresponded generally with those reported in the interview to assess by: preparation, language and effects. These comments were targeted at issues that emerged in the presentation: Xing, the presenter, had exceeded the time limit and
mispronounced some words. The comments were provided to foster improvement. In this sense KU-T1 was practising formative assessment. However, given that there was no time for the presentation to be done again, the advice was intended for the presenters for future presentations or for those who were to present in later sessions. It can also be seen that although the teacher had tried to engage students with questions, the comments were largely dominated by the teacher and more of a monologue. Overall, no effective interactions with students and involvement of students in the assessment process occurred in this episode.

4.2.5 Discussion

At the outset of the investigation of the teachers’ assessment practice it is acknowledged that my presence influenced their practices. This was confirmed by KU-T1, who said: “I should say I did not do much of this [referring to his feedback practices] before. I did it because you are there”. With this understanding the following insights are proposed.

A sociocultural approach (Levinson & Sutton, 2001, 2009) views teachers as important local actors, who engage with the policy and its implementation. Through them, the policy meanings are interpreted and delivered into classroom practice. As the above analysis demonstrates, the teachers interpreted the rationale of making CET4 optional as related to the disparity of proficiency levels between the top key universities and other institutions. They were able to see the institutional efforts to locate English proficiency as the focus of instruction, which they mostly did. However, they tended to attribute the change of direction to their perceptions and the independent proficiency test that the institution adopted as a baseline requirement. The impact of the shift in policy on teaching and learning was scarcely perceived or in fact often taken for granted. After all, the external test has been optional for over a decade at this university.

The teachers understood the rationale of incorporating formative assessment as mainly intended to direct students’ attention from the product or test to the process of learning and using the language. The weightings to process assessment were not regarded as a substantive change by some teachers, who claimed to have been doing this all along. They were aware of the potential benefits of formative assessment for better learning, but mostly their tacit conceptualisation of formative assessment holds
implications for students only rather than for teachers themselves and feedback. These understandings were related to the teachers’ assessment practice.

Analysis of the teachers’ assessment practice revealed that most of the assessment procedures, the weightings and the use of assessment results, which were framed in the institutional assessment policy were taken up. However, the ways they practised assessment were different from each other. They claimed to use criteria, which were specified by some and shared with students but implied by others. There is evidence of student involvement in assessment but it was primarily dependent on the individual teacher and largely related to his/her prior experience. There is evidence of constructive feedback for improvement. However, despite the formative function that constructive feedback can play, conceptual awareness of its relation with formative assessment was lacking. These findings, along with the invariable use of assessment results for reporting purposes, positioned process assessment, currently represented at KU as formative assessment, as primarily assessment of learning rather than assessment for learning. The relationship between teachers’ assessment practice with their interpretation of the meaning of policy at the social order level, their perceptions and experience at the individual level (Murphy, 2008) could be seen clearly.

These findings on one hand indicate that the teachers’ assessment practice reflects the policy intent only to a limited degree in this context; on the other hand, they point to a strong need for professional development in supporting teachers as they come to grips with the intent of the changes in assessment policy (Fullan, 2004; Snow, 1989). The students’ experience of the policy change and how they take up the apparent formative feedback is examined now to build a more detailed picture of the implementation of the policy changes.

4.3 STUDENT LEVEL RESPONSES

Informed by the sociocultural perspective, the investigation of the students’ experience of the College English assessment starts with the informants’ profiles (Section 4.3.1) and their perception regarding ELL and assessment (Section 4.3.2). This is followed by an examination of their experience of assessment change (Section 4.3.3) and their responsive actions to cope with it (Section 4.3.4). The investigation was based on a series of individual interviews that were conducted with
The role of the students, which is posited as active in formative assessment principles (ARG, 2002; Black et al., 2003), is the focus of the investigation. The research question that relates to the students’ responses to the policy change will be addressed in this section.

### 4.3.1 Profiles

The sociocultural perspective’s emphasis on prior experience (Vygotsky & Cole, 1978) holds true with the students as it does with the teachers. The students’ profiles are accordingly presented below.

**KU-S1**

KU-S1, a female student, came from a “very poor” county of Sichuan province, a level three development region (China Investigation & Assessment Center, 2008), which means under-developed in China. She started her English learning in junior high school with an interest in English, but this declined in the face of test-driven teaching practice that was teacher-dominated and involved “strict control”. KU-S1 described her English language learning experience as “quite passive” and “frustrating”, especially in her senior high, when she just followed the “programmed procedures” and was “pushed to study for the only sake of passing the EEHE”. She evaluated her English as “very poor” and especially weak in terms of expressing. KU-S1 related her weakness in expressive ability to the previous contextual situations, when she seldom had access to authentic materials or opportunities to practice authentic use of English and the exclusion of listening and speaking in the EEHE.

KU-S1 perceived her College English learning at KU as “much freer” and “much better” in terms of interest but still a struggle due to her poor English foundation. She identified herself as a passive learner and conceded she had been and was still studying for the test. The only non-testing experience KU-S1 mentioned was KU-T1’s writing blog program.

**KU-S2**

KU-S2, a female student, came from Dalian, a developed seaside city in Liaoning Province, which is a level two region (China Investigation & Assessment Center, 2008). That means it is developed in various aspects. The high school she had attended was a well known arts school, where English was emphasised because
of its status as one of the compulsory subjects for the EEHE. KU-S2 described her English learning at high school as “mainly EEHE oriented and teacher-centred” but teachers did more than just prepare the students for the EEHE. Speaking was included in the curriculum and was “full of interest” because topics about their arts life were used. In learning English, KU-S2 gained not only “a lot of knowledge from the textbooks”, but also views about the world. More importantly, KU-S2 attributed her choice of major in landscape engineering to an English article she read in high school. She said that “sometimes I think I find my future in it (English)...” KU-S2 had previous experience of involvement in peer- or self-assessment, but she did not seem to value it much. Rather, she preferred to use self-reflection strategically to achieve her learning aims.

**KU-S3**

KU-S3 was a female student from a medium sized city in Zhejiang Province, a level two region according to the 2008 China Development Index (China Investigation & Assessment Center, 2008). She attended a key high school at the provincial level, which means it was one of the best high schools within the province. Her English learning started from primary four when dialogue was the focus and she learned the alphabet and simple expressions for daily use. English contests in high school impressed KU-S3 deeply because preparing for and attending these contests helped improve her English level quickly and noticeably. The positive attitude she held towards contests extended further to testing. She did not think too much of the test results. Rather she viewed testing as a “most effective approach” of motivating her to memorise and master the language. The only non-testing assessment she was involved in was the writing blog program by KU-T1.

**KU-S4**

KU-S4 was a male student from a county of Gansu Province, a level four region according to the 2008 China Development Index (China Investigation & Assessment Center, 2008), which he described as “a mountainous region, under-developed, a poorest county in the country”. He had been to a key high school in the county, which according to him was not as good as the worst municipal school in big cities. He described the English education at the high school as “rotten”. Although English classes had been provided from primary three on, his English learning actually did not start until he was in senior high. The reason was that English as a
subject in his primary years had been experimental and was of a lesser status. Consequently, English was not taken seriously by either the school or students. No qualified English teacher was available so a salesgirl was found to fill in. English teaching during the junior high school years was not much better. He went over a book entitled *Through the Senior High Entrance Exam* for the senior high school entrance test and passed. English lessons in senior high were EEHE focused and didactic, which was given in Chinese and dominated by the teacher’s explanations of the test items and there were few interactions in class. But he did not think the experience was “painful” except for the memorisation of vocabulary, which he saw as a basic and decisive index of one’s English level and his headache. He rated his English as “average” at the high school but “very poor” here at KU, especially in terms of listening and speaking. Similar to KU-S1, he attributed the weakness to his poor English learning background and the EEHE’s non inclusion of listening and speaking. KU-S4 categorised himself as a passive learner and did not participate in the classroom activities unless he “has to”, which as he put was a “long established habit” for him.

**KU-S5**

KU-S5, male, came from Shanghai, which is a metropolitan city and one of the two top level regions listed in the 2008 China Development Index (China Investigation & Assessment Center, 2008). He attended one of the best key high schools there. His English learning started from primary three and was quite enjoyable until junior three, when he was streamed to a different teaching mode to prepare for the senior high school entrance exam and later for the EEHE. KU-S5 specially mentioned his English teacher in junior high school, who had studied abroad and appeared to favour communicative pedagogy. The senior high school KU-S5 attended had foreign teachers who organised various activities such as summer camps. However, the test-driven practices from junior three through to senior high seemed to exert an obvious negative impact on KU-S5. As he said: “My original interest was worn away by these practices. Actually I developed an antipathy towards it (English).” He felt “quite oppressed because of these tests.” But overall, KU-S5 regarded ELL as a “smooth” and “proud” experience for him, especially in terms of test results.
As the above profile descriptions made clear, the five KU student interviewees’ English learning and assessment experiences are similar in some ways. For instance, they had all attended formal English lessons at school and they all had been through a testing oriented system for the purpose of achieving good results in the EEHE to gain entry to KU. Nevertheless, the English education they had each experienced and the way they each approached English learning was quite different. Of the two students from the under-developed regions, KU-S1 had started her ELL late; KU-S4 started early but with unqualified teachers. They both had little chance to use English in authentic contexts. Their ELL was restricted and totally oriented towards the EEHE. Listening and speaking, for example, were rarely attended to at the high school for the reason that they were not included in the EEHE. They both believed that this had resulted in a weakness in these two basic skills for them and put them in an inferior position compared with their peers at KU. In contrast, the three other students, who are from big cities or the developed regions, had all started their ELL early, usually from mid primary years. They had qualified English teachers and were provided with opportunities to use English while learning. The washback effect of the EEHE on their ELL was obvious too, but their classroom activities were by no means monotonous and totally test-driven. Listening and speaking were covered in the syllabus. The experiences as such had equipped them with sound English proficiency and confidence compared with their peers at KU. The students’ experience had a marked influence on their perceptions, which are presented below.

4.3.2 Student perceptions

Similar to the teachers’ perceptions (Section 4.2.2), themes emerged about the students’ views of ELL and assessment and their understandings of the role of teacher and learner.

**English Language Learning**

The students’ views of ELL are closely related to identity which they established in their prior learning experience. KU-S1 and KU-S4, who identified their English as “poor” or “rotten”, saw ELL as difficult and a struggle. ELL hence seemed to be an important but overwhelming task. Vocabulary was a big concern for KU-S4. He claimed to have spent nearly half of his ELL time on memorising vocabulary. Despite the time he spent on vocabulary, he still lost a lot of marks in EEHE because of it. His aim for ELL was to communicate with others and to
understand English movies and songs. The situation is different for the other three students, KU-S2, KU-S3 and KU-S5, who were from well-resourced ELL backgrounds. A detached attitude towards ELL was shared and they unanimously referred to it as “tool”. KU-S2, for example, questioned the need to take English too seriously if one did not actually need to use it in future. Also shared was a valuing of learning opportunities that involved using English in practical contexts. KU-S3, for example, attended well-known English training programs run by an English education company in vacations because of the opportunities for communicating with native English speakers and she claimed noticeable improvements in her communicative ability as a result. KU-S2 once volunteered to be a guide for an American student tour group, which she valued and regarded as “useful and meaningful” experience. They appeared to have attained good English proficiency and hence confidence in using it. The current College English learning was described as “easy and comfortable” (KU-S3). According to KU-S3, it was still important but no longer a major worry so that they could spare their energy and time for specialty study or social activities.

The roles of teacher and learner in ELL

The students’ views split over the role that the teacher played in their ELL. It was an important guiding role for KU-S1 since the teacher told her about learning strategies and provided her with learning resources. She seated herself in the desk nearest to the teacher’s table and tended to follow the teacher’s instructions and did mostly “as I[she] was directed”. The guiding role of the teacher was viewed as minimal by KU-S4, although he admitted he was sometimes motivated by the teachers. He said:

Some teachers are very kind to us. That makes me feel I should learn English since they are so kind to us. As to guiding role, I really don’t think they guide me much and don’t quite know they guide me in what sense. As to what we need, and how we can achieve it, there is even less guidance.

In contrast to KU-S1’s reliance on and KU-S4’s non-acknowledgement of the teacher in their ELL, KU-S2, KU-S3 and KU-S5, the three students from developed regions, shared a more ambivalent view of the teachers’ role. They acknowledged the teacher’s guiding and supporting role, but they seemed to hold a view that what teachers did fell short of their expectations. KU-S2 for example expected the teacher
to “tell me what I don’t know and what ability I have not got yet for my English study.” Failing to have her expectation met in classroom, KU-S2 seated herself in the back row and tended not to participate in classroom activities “passionately”. Instead, she took the control of her ELL and established herself as an independent learner. KU-S3 too stated her partial reliance on teachers or classroom learning, which she estimated as “no more than 30%”. She recommended that learners take the initiative in learning and teachers introduce more interactive activities and less lecturing. KU-S5 specified the reliance on the teacher as at the initial stage of ELL and then learners took up an increasingly active role. However, given that the English classes were usually teacher-centred and lacked real interactions, KU-S5 advised learners to:

- take [their] role on the basis of teachers’ guidance, fetching up [their] disadvantages and playing [their] advantages according to [their] particular situations, since [they] know [themselves] the best. And also do something of [their] interest.

A general dissatisfaction with the traditional role that teachers took and a consciousness for learners to take the initiative in their English learning seems apparent.

**The roles of teacher and learner in assessment**

The students’ perceptions regarding the roles of teacher and learner in assessment were revealed through an analysis of their views of self, peer-assessment and teacher assessment. The analysis first of all revealed different understandings in the concepts. One understanding about self and peer-assessment was the assessment that was formally administered and related to grading practice. However, this was understood by students who had the experience only such as KU-S2. Those who did not have this kind of experience tended to understand self-assessment as self-reflection, which is an individual strategy of self-monitoring and evaluation in relation to personal goals in ELL. It is referred to as self-reflection in this study. When peer-assessment was mentioned, it was sometimes regarded as a synonym for advice or feedback from friends or classmates, to which no stakes were attached. I term it peer feedback. To these different understandings the students expressed different attitudes, which are elaborated upon below.

The formal use of self and peer-assessment was reported by KU-S2 in previous experience and by KU-S4 and KU-S5 in a different subject. They shared a sceptical
attitude towards it. According to KU-S2, that was an administrative thing and closely related to one’s grades. “Nobody actually takes it seriously.” KU-S4 talked about his experience in a different subject, where students ended up with “giving full scores to each other … it turned out to be a comparison in personal relationships…” KU-S4 accordingly argued: “Student assessment, when not included in scores, is tokenism; when included in scores, is more of tokenism. Everybody wants high scores, you know”. This view was echoed by KU-S5, who did not think self or peer-assessment was reliable for use in high stakes circumstances and considered objective tests as “more accountable” for this purpose.

Both KU-S2 and KU-S5, however, spoke favourably of peer-assessment in high school, when the teacher asked them to assess each other’s writing (KU-S2) or presentations (KU-S5). They claimed to have been actively involved in the activities, and consequently learned a lot from the experience, which KU-S5 referred to as a “different and valuable” kind of learning. This seems to shed a different perspective to their understanding and attitude towards peer-assessment. But they stated having not had similar experiences at KU.

In contrast to their sceptical attitude towards self and peer-assessment, the students embraced self-reflection enthusiastically. They all value it highly on the ground that they best knew themselves and had all used it constantly to regulate their learning towards self-initiated goals. This was exemplified by KU-S2: “Only I know what I want to do and what I intend to achieve, while others judge from their perspectives and according to their criteria”.

Mixed views were revealed regarding the students’ attitude towards peer feedback in relation to feedback from the teacher. KU-S3 regarded feedback from both teachers and peers as important and necessary, since they might come up with something that she failed to see for herself as a learner. In contrast, KU-S2 maintained that the feedback from both teachers and peers was “of minor importance” to her. She tended to use it only as a reference. KU-S1 and KU-S4 too doubted the usefulness of the feedback from peers. Comparatively, they preferred feedback from the teacher. It seems mostly that the teacher still takes a predominant role in assessment in the students’ views, whereas students’ involvement in assessment in the form of self and peer-assessment was doubted, although the use of reflection to regulate their ELL was valued.
Assessment

The five students unanimously identified testing as the major if not the only means that had been used to assess their ELL and tended to view assessment as testing. However, their views about the impact of testing on their ELL differ. KU-S1 and KU-S5, for instance, regarded that the testing experience had impacted negatively on them as learners, and especially so in terms of interest in ELL. For KU-S2 and KU-S3, tests had never been “a heavy burden”. Rather, they regarded it as an efficient and effective way to memorise vocabulary, grammar and the like. KU-S4 perceived testing as the focus of his ELL. He did not seem to hold negative feelings towards tests, which actually served as a motivation and “in a way forced” him to learn. These perceptions might potentially mediate the students’ responses to the assessment policy change, which is presented below.

4.3.3 Assessment practices and procedures as experienced by the students

Similar to the teachers’ experience, the students’ accounts of their experience of institutional assessment policy can be categorised into five major themes, which are assessment procedures, criteria, student involvement, feedback and grades. The key findings on each of these are now presented below from the student data.

Assessment procedures

Quite a few insights were revealed regarding the assessment procedures that the students experienced. To start with, all the five students affirmed the incorporation of process assessment within the institutional assessment framework. They were aware of an intention to shift the assessment focus to the process of ELL. This was evidenced with KU-S3: “What was emphasised is that the process is the most important”.

Second, they saw greater variety in the specific procedures used to assess their College English as compared with previous assessment modes. The students were able to specify the procedures, which included attendance, assessment tasks such as presentation, quiz, and writing tasks. These findings were consistent with what was revealed about the teachers’ assessment practices in the previous section.

Other insights included first of all a concern about the effectiveness of the final term exam. The exam’s focus on the textbook and some teachers’ purposefully asking students to recite certain paragraphs tended to make it test memory more than
proficiency (KU-S2, KU-S5). This practice in KU-S2’s opinion tended to demean the learning purpose since “practice and practice only will suffice”.

Also revealed was an impression that “different teachers do differently”. This was stated by both KU-S4 and KU-S5 and related to the use of assessment tasks and criteria. They coincidentally mentioned an anecdote in the previous semester when most students in a band one class got very high scores, but there were only a few in a band four class that did. This occurred to them as being “not fair”. It posed a reliability issue related to the practice of process assessment, which involved teachers’ subjectivity in maintaining standards and scoring according to the bands that the students were in.

**The use of criteria**

A consensus reached by the students is that they were not clear about what criteria, if any, had been used and how they were used by the teachers to assess their College English learning. KU-S2’s description below is quite representative.

The criteria are ambiguous to me. We don’t know what is expected…
Different students choose to present on different topics, which can be of different degree in terms of difficulty. Then how does the teacher give scores
I am really not clear…

This claim resonated with KU-S3, who saw the presentation that KU-T1 intended as criteria sharing as “showing a goal to work towards”. She asserted “either the teachers did not have clear criteria for assessment or they deliberately blurred them”. The ambiguity in criteria, however, was thought as not necessarily a bad thing in the sense that this uncertainty actually led the students to make concerted efforts in every task. Moreover, it appeared that the students had developed an ability to derive the criteria implicitly by using their common sense. This is evidenced by the excerpt below:

… When the teacher said the grammar was poor in this presentation, I know grammar is a criterion; when he said the grammar was all right but the expression was so so, I know he wants beautiful expression; when he said everything was beautiful but the logic. I know he wants logic…
These findings revealed a lack of transparency in the teachers’ use of assessment criteria. It seems the criteria, even if used or communicated, were not explicit enough or not clearly understood by the students.

**Student involvement**

In agreement with the report of KU-T1, the students admitted having been communicated to the assessment arrangements and being asked their opinions. However, they still demonstrated confusion about the percentage allocated to various assessment procedures and did not seem to care much about the exact weightings. The reason, as KU-S3 stated, was that she did not think the students would alter their level of engagement in learning in relation to the assessment allocations. Rather, they tended to remain consistently committed to all tasks. This involvement in assessment does not actually affect them much.

The peer review in KU-T1’s writing blog program was reported by KU-S1, KU-S2 and KU-S3. These three students expressed three different views. According to KU-S3, her classmates had examined her writing “very carefully” and pointed out the strengths as well as weaknesses for her. The feedback was much more detailed and specific than that from the teacher and came up with something that she failed to see for herself, so she valued it highly. KU-S2, however, maintained that the feedback from both teachers and peers was “of minor importance” to her. She tended to use it as a reference but neither would affect her study. KU-S1 reported to have received some feedback from the teacher and other students, but the feedback did not seem to give her constructive advice for improvement. She also showed a concern about the possibility of exposing more weaknesses to peers. It seems peer review in KU-T1’s writing blog was experienced differently. The different experiences appear to relate with how the students perceived the value of feedback.

**Feedback**

Analysis of the students’ response to the questions about feedback generated several insights about the feedback they received from the teachers. First, it was general rather than individualised. KU-S2 commented: “Feedback? … The teacher will tell us the merits or demerits of our presentation in general. Overall speaking, I don’t think the teacher gives much individualised feedback”. Second, the feedback was undifferentiated. This was stated by KU-S3, KU-S4 and exemplified by KU-S1, who observed: “… Actually I think he gave roughly the same feedback to us…”
Also, the feedback was more complimentary than critical and constructive. KU-S3 complained that her teacher tended to say “good, very good” to everyone and then “pointed out a small problem.” This was echoed by KU-S4, who said: “If only [the teacher] saw your efforts, he will find heaps of praise for each presentation”.

The encouraging value that the teachers intended in their positive feedback seems to have been taken up by some but not by others. Some students like KU-S5 and KU-S4 reported “feel[ing] good” on teachers’ positive comments, which in turn became motivation for them to carry on. In contrast, other students such as KU-S1, KU-S3 and KU-S5 tended to ignore the comments, which, being undifferentiated, appeared to be of no substantial value to them. They accordingly expressed a wish for more critical, individualised, and constructive advice, as was illustrated by KU-S3 below.

We hope he [KU-T1] can point out more [issues]. We Chinese students had experienced too much frustration; we don’t actually need so much praise.
We hope for some beneficial advice for improvement.

Differing from other students who took no follow-up action, KU-S2 reported acting upon the feedback KU-T1 gave and sent a revised version to the teacher. She frankly admitted a concern with the grades, as she said: “If I don’t, my score will be affected. I am direct in this point but that is true.” It seems the formative value attached to feedback for improvement, when taken up, is for the purpose of better grades. Summative ends are still the students’ major concern and decide to a great extent whether the feedback is acted upon or not. This further indicates an obvious product orientation in students’ responsive behaviour and in a way affirmed the summative nature of the assessment practice at KU.

Grades

As observed by the teachers and self-reported by the students, grades had been the students’ major concern. However, they did not seem to feel great pressure because of it. This was typically stated by KU-S1: “Basically, you can get the 10% by being there every time; so is the 10% for presentation. You get it only if you do it…” KU-S4 concurred and commented “… as far as I know all the points will be given to us…” The finding confirmed the undifferentiating nature of the grades that the teachers reported in the previous section.
A general dissatisfaction with the grading system was revealed in students’ responses. KU-S4, for example, reported that because of little pressure, he had spent less time on ELL and hence felt a decline in his English level. KU-S2 and KU-S5 too articulated dislike of the grading system, which as they understood failed to indicate the students’ proficiency. They hence held a doubt about the fairness of the teachers’ professional judgement.

4.3.4 Student responsive behaviours

Students’ responsive behaviours to cope with the external CET4 and process assessment are analysed in this section. The focus is to find out if the implemented assessment policy had actually functioned to improve learning and to what extent, if any, the policy change had helped to shift the students’ orientation towards learning.

**External test as optional**

All the students acknowledged an understanding of the optional nature of the external CET4, but they had either taken or registered for it. This finding concurs with what the teachers observed previously. The social reasons identified by the teachers were confirmed but responses were not limited to these. Actually only KU-S4 mentioned the better approbatory status of CET4 in the social context and the possible use of the certificate for employment purposes. Other students held different opinions. KU-S3, for instance, regarded the CET4 or 6 certificates as “just [an] accessory, definitely no guarantee for a job”. They all seemed to have plans to obtain better proof such as the GRE or TOEFL to use for the same purpose.

Several other reasons were identified. The first relates to the students’ habitual reliance on tests for measurement purposes, which the educational system had established in them. KU-S1 for example felt “study and tak[ing] tests a kind of habit” and she registered for CET4 to measure her English proficiency level. There were also students who reported to use the test to ascertain some time for ELL. KU-S2 and KU-S5 were cases in point. To both of them, CET4 was no challenge at all. They took it only so that they could perform well in tests required for studying abroad. The reason that KU-S3 expressed was not as pragmatic. She, with a positive attitude towards testing, saw multiple benefits in taking the test, which included “[motivating] one to learn”, “[cultivat[ing] one’s integrated ability in a real sense”, practising one’s ability in time management, self regulation and even psychological endurance.
Learning context was also identified as a major influence on students’ taking the now optional CET4. This is evidenced by KU-S4 below:

It does not sound good if I say I just follow suit. But that is the fact. Plus, there is CET6 after CET4 and TOEFL after CET6 etc. so I start with the first step… There are not many reasons for me to refuse to take the test, are there?

There seemed to be signs of taking the test for learning purposes (see KU-S3’s reasons). But the emphasis on the learning product and the use of the result for utilitarian purposes still appeared strong. Despite the high stakes was removed from the test, the students still seemed to be under pressure from a culture of expectations, which were now not from the institutions or teachers but from employers and parents in the social background, or students themselves.

**Process assessment**

With a highlighted concern about scores, the students reported to have adopted a pragmatic approach to cope with the incorporated process assessment. The approach, described by KU-S2 as “do as I am supposed to”, was described by KU-S4 using a Chinese saying – “confront soldiers with generals and stem water with earth” (兵来将挡，水来土掩 bīnglái jiàngdǎng, shuǐ lái tǔ yǎn), which means “taking measures as the situation calls for them”.

Insights were revealed as to the impact that process assessment had on students’ ELL time, approaches and outcomes. First, the effect that process assessment had on the students’ time of ELL and effort distribution is affirmed, which is illustrated by KU-S5 as: “Without it, I would probably come to class only, spending little or no extra time on ELL and then cram when the test is approaching”. However, the impact of process assessment on ELL approaches was seen as minimal. KU-S2 asserted bluntly: “No influence for me… I can’t possibly change my ways of ELL because of it. That is impossible”. By this, KU-S2 seemed to imply another approach for learning purposes, which differed from the pragmatic approaches that the students adopted for the sake of high scores. That approach was aware of by other students too, though they might not probably have obtained it yet. KU-S4 said: “… If I want to learn English, I will definitely learn it well. Not that kind of learning to cope with tests. I mean learning it well. If only I make it, tests will be no problems”. Third, the students’ views were divided regarding the effect that process assessment had on their learning. Some saw it as positive: KU-S3 and KU-S5, for
instance, maintained that the meticulous preparation the students made for the assessment tasks like presentation, which included searching for proper materials, reading to choose the best ones, extracting and synthesizing to make it a well organised PowerPoint, were “very good learning and improvement”. Others disagreed. KU-S2 for example refuted the possibility that her English was improved as a result of preparing for a presentation. KU-S1 too regarded that with tests “taking the lion’s share,” the positive impact of process assessment on her ELL was minimal.

As the above analysis shows, the students’ responses to the enacted assessment policy were still dominated by a concern about the results and hence primarily oriented towards product rather than process. There was evidence that the enacted process assessment had functioned to support learning (see KU-S3’s and KU-S5’s responses). But given that it had impacted on only the time students allocated to ELL but not their learning approach and orientation, it is justified to say that this formative function had been achieved only in a limited degree for the students of this context.

4.4 SUMMARY

Seen from a sociocultural perspective (Levinson & Sutton, 2001, 2009), the assessment principles advocated in the CERP have so far undergone a reshaping process in the local context of the KU. The reshaping practices were closely related to the local actors’ appropriation of the meaning of the policy within the situated cultural, institutional, and historical context (Wertsch et al., 1995).

The appropriation at the institutional level resulted in adoption of the two policy initiatives in a modified manner: the external test became optional but an internal proficiency test was still used to maintain standards; formative assessment was implemented as process assessment. Students’ involvement in assessment and feedback, which were encouraged in the national policy, were not emphasised in the institutional assessment policy. With these modifications made, the interpretation of formative assessment at the institutional level seems to take on a different perspective from the national and international levels. Nonetheless, the KU seems to be a leading university interested in innovation. The introduction of process assessment provides evidence of a shift in focus to assessment for learning and
capacity for change. This can also be seen from the larger weightings allocated to process assessment.

The reshaped institutional assessment policy was communicated to the teachers with the expectation that it would be translated into classroom practices. Like the appropriation that occurred at the institutional level, this translation underwent a process of re-creation. Lack of training was a major factor influencing this re-creating process. With no systematic training provided on the assessment principles, the teachers relied on their tacit knowledge rather than theory for an understanding about these policy initiatives. It was found that the teachers largely took the institutional interpretation of formative assessment as process assessment and enacted most of the procedures prescribed in the institutional assessment framework. However, further analysis of themes that emerged from the data showed that the specific assessment practices were mostly teachers’ individual decisions and related to the teachers’ perceptions and beliefs. There was evidence of use of criteria, although these were not clear to students. There was evidence of student involvement in assessment, which however did not seem to be taken seriously by most of the students given that they did not see it as contributing to their ELL. There was evidence of constructive feedback for learning purposes, which was not substantial and there appeared to be disappointment and suspicion over its effectiveness in improving English proficiency. Hence, the students tended to act upon only when scores might be affected. The undifferentiating grades too indicated a weakly fulfilled function of assessment in measurement. A need for professional training and support thus surfaced.

The students’ uptake of these policy initiatives was complex. The findings indicate firstly that all the interviewed students took or would take the CET4 despite the knowledge that it was optional. Actually, the CET4/6 was only the first and lowest of the many external English tests that most of the KU students were to take during their university years. Their taking of the test was related to a culture of expectations in the social context. The influence of the examination tradition seems strong. Second, it was agreed among the students that the incorporation of process assessment had impacted on their ELL time and effort, but not their approaches to and outcomes from ELL. Some of the students seemed to retain a double orientation to testing and learning, they also seemed to adopt different approaches to achieve
The students’ responses appeared to relate to their educational background and prior experience in ELL and assessment. However, given the limited sample size in this case study, this claim is tentative and needs to be confirmed by a large-scale quantitative study.

Overall, these findings seem to point to a conclusion that the impact of these initiated policy changes on the teachers and students was minimal. The policy intent, to measure as well as to improve learning via formative assessment, was achieved only to a limited degree in this context. The examination of the College English assessment policy at KU indicated that policy implementation is by no means a linear process from the text to practice; rather, it is an appropriating and recreating process in “situated locales and community of practice” with unexpected outcomes (Levinson et al., 2009, p. 768).
Chapter 5: Case Study Two: Non-Key University (NKU)

This chapter presents the findings related to the Non-Key University (NKU) case. By documenting the localized responses to the assessment policy change initiated in the College English Reform Program (CERP) (CMoE, 2004) at the levels of the institution, the teachers and the students, this chapter addresses the research questions in this particular context. It duplicates the format of Chapter 4 and adopts the same sociocultural perspective.

Following the introduction, NKU’s appropriation of the two policy changes, namely, to make the external test optional and to incorporate formative assessment, within the local context are described, interpreted and discussed (Section 5.1). The next two sections detail the teachers’ (Section 5.2) and students’ (Section 5.3) lived experience of the assessment policy changes and their responses in relation to their prior experiences and perceptions. The final section (Section 5.4) summarizes the findings and brings the interpretation and discussion to a conclusion for this case. This university is referred to as NKU. Considerations are given to the ethical obligations of keeping the identity of the university and all the informants anonymous.

5.1 INSTITUTIONAL LEVEL RESPONSES

The understandings of the institutional level responses to the initiated assessment policy changes in the CERP (CMoE, 2005, 2007) are derived from an analysis of data sources which include an informal interview with a senior adminyster responsible for College English teaching at the university, documents collected from the university website and personal communications with the teachers. It starts with a profile of the local context (Section 5.1.1), which focuses on what makes NKU a non-key rather than a key university in the Chinese higher educational context. This is followed by the changes that NKU made in its College English assessment policy in response to the advocacy of CERP (Section 5.1.2). The discussion (Section 5.1.3) related to the first research question focuses on the institution’s interpretation of the national assessment policy.
5.1.1 Profile

NKU is one of the 2000 Chinese higher educational institutions. It is located in a Midwest province of China and has an enrolment of around 15000 students from around the country. It is a comprehensive university and enjoys a relatively long history within the province and country. It is a major university in the local region, but is not classified as key at the national level. Being non-key and located in a provincial city implies some disadvantages for this university in terms of financial support, staff quality and student quality.

In terms of financial support, this university was not included in the Chinese government’s “985” or “211” Projects, two programs that the government launched at the higher educational level with the necessary funding to accelerate some major universities to become world-class. Regionally, the province was categorized as third out of four levels of development by the 2008 China Development Index (Center for China Investigation and Evaluation, 2008). This means that the region is relatively under-developed socially, economically and educationally. The higher educational input of the province was 1% of its regional Gross Domestic Production (GDP) in 2005, which was much lower than the national average of 3% and 5.75% in Beijing (Treasury Department CMoE, 2008). Given that government support is the major financial source of educational resources in China, it can be said that the financial support to this university is by no means adequate.

There are 50 teachers in the College English Department of NKU, which include two professors, 13 associate professors, 22 lecturers and 14 assistant lecturers. None of the staff currently have a doctorate, but four are presently engaged in doctoral studies. All of the College English teachers are local Chinese, 20% of whom have studied overseas as visiting scholars (Personal communication, April 4th, 2009). They are now teaching College English at this university.

This university enrolls 2500 undergraduates annually. The majority of these students’ test results in the Entrance Examination of Higher Education (EEHE) are slightly above the national higher education admission line. Half of them are from local schools within the province. Around 2300 of these students take an area other than English as their speciality and are hence called non-English major students. They are the target group of College English education at NKU.
Sociocultural theory highlights the role of context, which includes its history and the expectations that have built among the community, in affording and constraining the agents’ behaviors (Lave & Wenger, 1998). The institutional responses to the nationally initiated changes to the College English assessment policy are presented now from this underpinning theoretical perspective.

5.1.2 Changes to the institutional assessment policy

The College English assessment system at NKU has witnessed major changes since 2005, when the university was involved in the second round of piloting of College English Curriculum Requirements (CECR) (CMoE, 2004). Previously, standardised achievement tests were organised at the end of each semester and served a summative purpose. The band four achievement test was required for a certificate that indicates the successful completion of four years of study, whereas a pass of CET4, the external test, was a qualification requirement for the Bachelor degree. Yet, the requirement for CET4 was not strictly followed. Specifically, although 60 points constituted a pass in CET4, the institution had allowed students to graduate with a lower grade such as 55 points in the event that too many students failed to achieve the required 60 points to pass the CET4.

The changes made in 2005 include, first of all, that CET4, the external test, became optional and separate from the qualification requirements for the Bachelor degree. Nonetheless, the test was still organised at the campus and the use of the students’ results in the test for purposes such as the award of scholarships was retained. The second change relates to the internal assessment of College English coursework and will be elaborated upon below.

The current College English internal assessment framework consists of two parts and differs from the past assessment which relied solely on the achievement test. One part now relates to assessment of the learning process which is referred to as process assessment (过程评价 guòchéngpíngjià). Similar to KU, process assessment at NKU is used interchangeably with the term formative assessment by the senior administrator and the teachers. Ten percent of the overall College English learning assessment is allocated to process assessment, which is further divided as follows: 4% for classroom participation, 3% for assignments and 3% for attendance. Students need to perform well in class to achieve the 4% for classroom participation and
submit a minimum of three assignments to achieve the 3%. One recorded absence means one point is subtracted from the 3% of the marks allocated for attendance. The teachers are required to record students’ performance on these three aspects. The combined result is recorded as a process grade (平时成绩 pingshichèngji) and represents 10% of the assessment for reporting purposes.

The final term exam administered at the end of semester consists of the other part of the College English assessment. The exam consists of an oral and a written test, which are organised separately. The oral test is conducted face-to-face and involves individual oral work. It contributes 10% towards the College English final report. The remaining 80% is allocated to the written test that is textbook-based and achievement-focused. Online learning is currently being trialled and has recently been allocated 10% of the 80%. The College English assessment framework at NKU is inclusive of both the process and the product of College English learning as illustrated below in Figure 5.1.

This framework applies to College English band one through to band four assessment and is regulated at the institutional level. Band four is now used as the English gate-keeping qualification requirement at NKU. Students who successfully pass the test will have met the English language qualifications for the award of the bachelor degree.

5.1.3 Discussion

The above description of the current assessment arrangements for College English at NKU was iterated by a senior administrator, with whom I communicated
personally, and was confirmed in the interviews of the NKU teachers. From a sociocultural perspective, I now consider the institutional College English assessment policy in practice that has resulted from an appropriation of the assessment policy of CECR (CMoE, 2007) within the situated context of NKU. The way in which the policy has been appropriated can be seen from a comparison of the described College English assessment framework with the two policy changes advocated in the CERP at the national level.

Table 5.1
A Comparison of Assessment Policy at National and Institutional Levels

<table>
<thead>
<tr>
<th>Assessment policy initiatives in CERP</th>
<th>Assessment policy at KU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making external test optional</td>
<td>CET4 no longer obligatory and disconnected from qualification requirements; but still encouraged</td>
</tr>
<tr>
<td>Formative assessment incorporated</td>
<td>Weightings allocated to process assessment</td>
</tr>
</tbody>
</table>

NKU responded to the first initiative by making the CET4 optional and separate from the qualification requirements, which aligns with the prescribed policy at the national level. However, the high stakes attached to the external test have only been partially removed as taking the test is encouraged and linked to the award of scholarships. These associations and functions of the CET4 potentially will compromise the substantial nature of the change.

NKU responded to the formative assessment initiative by introducing a process assessment, which combined grades for students’ attendance, classroom participation and assignment tasks. A shift of the assessment focus from the product of learning to the process seems to be the intention. However such a move is compromised by two facts. The first is the weightings allocated to process assessment is minor in NKU’s case compared with the 60% at KU. The second is the use of these assessment results for reporting rather than enhancing learning purposes. This approach suggests a possible misinterpretation of the purpose of formative assessment. Moreover, feedback and student involvement in assessment, which are key to formative assessment (Black & Wiliam, 1998), were not emphasized. The focus of the internal College English assessment thus appears to be used more for summative rather than formative purposes.
How the policy was appropriated by the institution can also be seen in the senior administrator’s explanation on the weightings allocated for process assessment and examination. The senior administrator is responsible for the overall administration of College English teaching. She referred to herself as the implementer of policy, stating that she could initiate a change, but it was those in the more senior positions like the Teaching Administrative Office (TAO), who decided whether it was policy or not. Notably, as an intermediate administrator and a teacher herself, she stood between the policy maker and the practitioners and sometimes felt a sense of powerlessness in balancing some teachers’ behaviours and the demand for accountability from above.

According to the senior administrator, the allocation of certain weightings to process assessment was based on the belief that “to combine formative assessment and summative assessment is a more appropriate way to assess the students”. The resultant minor weightings allocated to process assessment at NKU were due to concerns regarding reliability. She singled out an intervention from the TAO, the teaching administration body and the “virtual assessment policy maker” at the institutional level of the instant where a teacher gave almost all students full scores for process assessment. When he was asked to explain, the teacher insisted that he was doing justice to his students, who all performed very well according to his standards. The TAO could not do anything about it but refused to give teachers greater control for assessment when this idea was proposed. The development of the NKU’s College English assessment framework illustrates the power and control operating at the institutional level in an effort to regulate assessment practice. This was detrimental to the aim to empower practitioners, as required by formative assessment. This control exerted major influence on the teachers’ attitude and actual enactment of the policy, as is elaborated on in Section 5.2.3 and 5.2.4.

How the authorised national policy was appropriated at the institutional level is important. On the one hand, it prescribes the specific procedures for the practitioners’ enactment of policy in their classroom practice and constitutes the implementation guidelines within this particular local context (Levinson et al., 2009, p. 768). On the other hand, “a policy is power” (Ball, 2006 p. 43). Being a policy itself implies an obligation on the practitioners’ part and thus has major implications for the teachers’ actual assessment practices in classrooms, as evident below.
5.2 TEACHER LEVEL RESPONSES

Understanding of NKU teachers’ enactment of the institutional assessment policy is derived from an analysis of data sources which include two individual teacher interviews and three observation sessions of their classes, a focus group interview and some collected artefacts. It starts with the teachers’ profiles (Section 5.2.1). The major themes on teachers’ perceptions (Section 5.3.2), understandings of the CERP policy change (Section 5.2.3) and their accounts in practice (Section 5.2.4) follow. A discussion (Section 5.2.5) summarises the main findings that address the research question relating to teachers’ responses to the policy changes.

5.2.1 Teacher profiles

The profiles of teacher participants are presented with an acknowledgement to the sociocultural emphasis on the critical role of historical and cultural backgrounds in explicating people’s behaviours (Wertsch et al., 1995). The presentation follows an order from individual teachers to the focus group.

NKU-Teacher 1 (T1)

NKU-T1, a female in her mid 30s, is a lecturer with over nine years’ College English teaching experience. She received her Master’s degree in Education several years ago, which has equipped her with a broad perspective on College English education and the program’s inherent problems. She has not undertaken study outside China. She teaches College English to two classes of around 100 first year students, which is the minimal teaching load for College English teachers at NKU.

NKU-Teacher 2 (T2)

NKU-T2, a female in her late 30s, started teaching College English in 1992 when she graduated from university. Over this period, she has built herself a reputation as one of the best English teachers in the university, and enjoys popularity among colleagues and students. She had a one-year stay at an institute in Singapore, where she took some Master courses. More recently, she had a short term visit to America as a visiting scholar. Her professional title is now Senior Lecturer and she was teaching three classes of 150 students when the data collection was in progress. She also undertakes some administrative work at the College English Department. Her extensive working and studying experience has established her as a confident and independent practitioner.
Focus group

The demographic information of the teachers involved in the focus group is presented in Table 3.3. As clearly shown, the five teachers’ College English teaching experience ranges from five to 23 years. Their professional titles range from lecturer to professor. They are all local Chinese. Only NKU-T7, the professor, had one year’s study experience in an English speaking country as a visiting scholar. All these participants teach College English courses, which include Reading, Writing & Translation (RWT) and Speaking & Listening (SL) lessons to the first year undergraduate students.

5.2.2 Teacher perceptions

The teachers’ perceptions are presented in order from their general views on English language learning (ELL) to the more specific themes related to assessment. This organisation mirrors the arrangement of the perceptions of the teachers at the Key University (KU) presented in Chapter 4.

English Language Learning

A communicative view of ELL surfaced in the comments of NKU-T2, NKU-T4 and NKU-T7. The view was that “English is a language” (NKU-T2) and should be taught and learned as such. Class transmission and reception in this view, though essential for establishing a base such as grammar and vocabulary was “only part of the story” (NKU-T2). Given that the “ultimate purpose of ELL is for communication” (NKU-T7), “practical use in daily life and thinking is more important” (NKU-T2). Therefore, a learner should be involved more in activities that have practical use. With such a view, NKU-T2 actively involved herself in English corner\(^2\) organisation at the NKU campus and strongly recommended it to her students. Despite their value of the language environment as a key supporting condition for effective ELL, NKU-T4 and NKU-T7 did not seem to regard the English corner as highly. NKU-T4, for example, argued that broken English did possibly more harm than good to one’s English learning.

Culture was valued in the teachers’ view of ELL. NKU-T2, for example, maintained that English should be regarded as “a window to know a different

---

\(^2\) English corner is an out-of-class English learning activity popular on the Chinese mainland, which can be self initiated by learners or organized by universities. It is used by English learners for the purpose of improving spoken language (Gao, 2008).
culture” so as to deepen one’s thinking. Therefore, the culture in which a language was imbedded should be included as an important aspect of ELL. NKU-T1 went further in this regard. She viewed ELL as a process in which learners were exposed to a full understanding of another culture, another way of thinking and then associated what they learned with their daily lives. She held strongly to the view that ELL was by no means just a matter of language.

The teachers also regarded the students’ approaches to ELL as related to their goals, which could be pure interest (NKU-T5) or practical need (NKU-T4) in a foreign country or even exams. The utilitarian attitude that regards learning English as a means to go to college or gain a degree was disapproved of by teachers such as NKU-T2. This mirrors debates around the role of motivation and types of motivation being more effective for ELL (Dornyei, 1994). The view that instrumental motivation was somehow less effective in promoting ELL has been noted in research although findings suggest that second language learners are motivated differently at different junctures in the learning experience (Dornyei, 1994). This will be discussed further in Chapter 6.

**The roles of teacher and learner in ELL**

The six teachers’ views regarding the role of teacher and learner in ELL were divided. One view, held by the majority of teachers, perceived the role of teacher as secondary. Teachers in this view were “guide, helper, and consultant” (NKU-T2) who tried their best to guide the students to think and provide advice when they needed it, and as an “auxiliary figure” (NKU-T7) who tried to arouse the students’ interest or trigger their motivation to learn. The active and central role of the learner was highlighted in this view, which considered that “learning [was] mainly up to the students themselves” (NKU-T7) and that internal and intrinsic motivation was the priority. A saying that “extrinsic factors work through intrinsic factors” was used to support this view. Therefore, as NKU-T1 stated, ELL was not only about lecturing but also about involving the students’ engagement; not only about imparting knowledge about what but also about showing the students how to study, about provoking the students to think and explore new meanings in life. By identifying the teachers’ role as auxiliary, this view placed the learner in a central position, and seems to be aligned with the role shift associated with the constructivist view of learning (Shepard, 2001).
The other view was held by the least experienced NKU-T6, who maintained that to achieve best learning results, the teacher should be a central figure responsible for imparting knowledge to the students and administrating the students’ learning in and out of class. By positioning the teacher as a predominant knowledge dispenser and organiser, this view seems to align with roles defined in the Chinese traditional values (Biggs, 1996).

**The roles of teacher and learner in assessment**

There is a similar complexity in the teachers’ perceptions regarding the roles of the teacher and the learner in the assessment. The scaffolding role of the teacher in assessment was identified by the teachers. NKU-T1, for example, referred to the teacher in assessment as “an instructor” who helped the learner to see areas in need of improvement and how to achieve a certain goal, and “a supervisor” who supervised the student’s ongoing learning process and provided advice when needed. This view was echoed by NKU-T2, who maintained that assessment was to be used by the teacher to help learners to improve rather than “control” their learning.

While the teacher’s role in assessment was agreed upon by the teachers, their understandings about the role of the learner in assessment showed a different perspective. They could see the multiple functions that were intended from involving students in assessment which included: improving students’ communicative skills as part of learning (NKU-T2); motivating students more effectively (NKU-T1); and making the assessment results more “convincing” to students (NKU-T7). However, not all of them thought it feasible. Some showed a concern about its enactment given the big class sizes, different learning styles (NKU-T2) and the lack of a systematic plan for professional training, individualized design and the allocation of class time for effective implementation (NKU-T1). In sharp contrast, some other teachers such as NKU-T4 and NKU-T7 articulated support for the policy initiative through involving students in assessing tasks such as writing and presentation. In NKU-T7’s opinion, the students could be trusted and they took the trust put on them responsibly, as she made clear below.

Even if you let the students to give themselves a grade, they will not do it on no grounds. I do like this to give a convincing grade. You know sometimes students are not clear about why I get a B while they think they deserve an A. After the student gives the grade to him/herself, I will ask the other students
to comment. Students do not necessarily keep silent or give consent because of face or something. They might utter their voices if they don’t agree, like “too high”...

She reported the positive impact of such practice and noted the trust was more effective than restraint in cultivating students’ self management ability. These views reflect the teachers’ different perspectives regarding assessments, which are elaborated upon below.

Assessment and its purposes

The teachers saw various purposes of assessment, which included “support[ing]” (NKU-T2, NKU-T4) and “motivat[ing]” (NKU-T1) learning so as to “help realize the ultimate purpose of practical use in their career life” (NKU-T6) or “promot[ing] both learning and teaching’ (NKU-T7). The view that regarded assessment as only a means to increase the students’ test results was of less interest (NKU-T2); rather, assessment was valued as an important way to let the students know where they were in their learning so that they could identify the areas requiring more effort. All this seems to reveal that the teachers regarded assessment as supporting learning. The teachers also saw that summative assessment, which emphasised the end product of learning, was not sufficient to ensure ongoing learning. Given that “there is no end in one’s study” (NKU-T1), the teachers thought that formative assessment, which focuses on the process of learning, should be “complementary”. That is, the teachers generally perceived a complementary relationship between summative and formative assessment.

Formative assessment

While the teachers saw the purpose of formative assessment as learning, they unanimously reported a lack of systematic understanding about formative assessment, which they attributed to a lack of relevant training and professional development. None of the seven teachers had formal or long-term training in assessment although there were reports about a short term conference (NKU-T2), pre-service observation (NKU-T4), and a teleconference (NKU-T5). The teachers seemed to regard these short term training courses as contributing little to their understanding of assessment and formative assessment in particular. They unanimously attributed their assessment skills and knowledge to experience, as shown in the following excerpts.

NKU-T4: My skills are mainly from experience.
NKU-T7: Mostly experience for me.

NKU-T5, NKU-T6: Yeah, experience, no theory.

Analysis of relevant data shows that the teachers’ understanding of formative assessment is complex. Primarily, it was related to when the assessment occurred, that is, they saw formative assessment as more focused on the process (NKU-T1, NKU-T5) and was continuous rather than “a once and for all business” (NKU-T4). Furthermore, there is evidence that indicates individualised feedback was valued as better for learning. This is shown in NKU-T2’s comments and in an example that NKU-T1 gave to demonstrate her understanding of formative assessment. As well, the teachers were aware of the subjective nature of formative assessment. NKU-T6 said:

Teacher is subjective, so the formative assessment can’t be absolutely objective. Even if two students get the same grade finally, they can be different in many ways. It is impossible to be exactly exact in numbers.

Moreover, the teachers saw formative assessment as having more of a learning enhancement function than summative assessment. However, they tended to associate this function with better summative results. This was evident in NKU-T2’s statement: “formative assessment is a necessary insurance for effective summative assessment”. This same point was iterated by NKU-T1 as “if the process is good, the result will take care of itself”. That is, the priority of summative purpose was still obvious, which is not surprising given the institutional assessment policy’s prescription of the assessment results for reporting uses. Tension appears to exist in the teachers’ understanding of formative assessment and this intended function because of these conflicting messages from the institutional policy. This tension became more obvious later when teachers tried to give objective grades to the formative assessment tasks.

The perception is one of the identified factors that influence the teachers’ teaching and assessment practice (Murphy, 2008; Rogers et al., 2007). However, it should be noted that this picture about the teachers’ understanding of formative assessment is not uniform across the teachers. It is also recognised that what is revealed in this section is only the teachers’ perceptions of what they are doing and what they are supposed to do rather than what they actually do. The existence of possible disparity between perceptions and practices was acknowledged by the
teachers. NKU-T2, for instance, admitted that on many occasions she was a leader and authority figure doing what she thought was the best. It needs to be recognised that the apparent constructivist view does not necessarily mean it will be reflected in classroom practice.

5.2.3 Teacher understandings of the policy change

How the teachers interpreted the meaning of the policy closely relates to their uptake and affirmation of it in classroom practices (Murphy, 2008). This understanding guides the examination of the teachers’ understandings of the policy change in this section. The teachers’ responses confirmed the changes made to the institutional assessment policy such as the introduction of an optional CET4 since 2005 and the incorporation of 10% process assessment. Presented below are the findings that relate to the teachers’ understanding of the two policy changes: the rationale and the impact on their assessment practice.

External test becoming optional

The NKU teachers reported failures to communicate this policy change to them from the institutional level. Nonetheless, the NKU teachers somehow seemed to know of this change. This was manifested in the excerpts below.

NKU-T7: … Later on, the China Ministry of Education (CMoE) issued a document, which demands the CET and degree to be disconnected. I searched on the web and found the document there. It was never communicated to us.

NKU-T5, NKU-T4: No.

The rationale for making the CET4 optional was understood by NKU-T2 as aiming to “divert teaching and learning from testing orientation”, which aligns with the policy intent. This intent, however, did not appear to be understood by the other teachers. They generally saw the policy initiative as well-intended overall, yet retained concerns about the change’s possible negative influence on the students’ motivation to learn. NKU-T1’s perspective below was representative:

I think it’s good. The external test should be optional. But when it becomes optional, more students don’t want to study English hard any longer. Most of them study for the sole purpose of passing CET4 or 6. If there were no exams, it is not necessary for them to study hard…
The teachers saw a disjuncture between the policy intent and its practice at the top level, which resulted in confusion about how to implement it for the individual institution and the teacher. NKU-T7 articulated this resultant confusion as follows:

The policy-maker, like the Chinese Ministry of Education (CMoE), is confusing in a way. You see, on one hand it states it is definitely not appropriate to measure all the students with one test paper; on the other hand, the CET4 test is maintained in a very high position. That tends to make institutions and teachers at loss as to how to link their teaching practice with the changing assessment. There seems to be a lack for such a link.

Accordingly, the teachers saw the administration of CET4 remained largely unchanged at NKU. They maintained that despite its optional status, CET4 was still the yardstick for ELL success (NKU-T2, 2, 3), which directed their teaching practice. This was made manifest by NKU-T2:

Yes, the CET4 test is no longer the only goal of College English education, but it definitely is a yardstick. Its orientation is supposed to be the orientation of College English teaching…

Nonetheless, most of the teachers did abandon the test preparation practice which used to occupy several weeks of teaching time. NKU-T6 was an exception. She admitted she still prepared her students for the test and she “hope[d] they can pass, and as many as possible”. While the abandonment of the test preparation in most teachers’ cases could be regarded as a change, the change of the CET4’s status seems to have exerted no impact on individual teachers such as NKU-T6’s teaching practice.

Two major insights emerged from the teachers’ comments about the impact of this policy change on students and their learning behaviours. First, the CET4 was still important to students, in whom a strong drive to sit for and pass the test was obvious. This is apparent from these excerpts.

NKU-T4: A student in my class is overjoyed to know she got registered for the CET4 successfully. Most of the students in my class will sit for the test. Only one did not register.

NKU-T5 and NKU-T6: Right, they will try every effort to pass.

NKU-T7: They all want to take the test, required or not.
NKU-T4: And happy when they pass. After all, the certificate is kind of a proof of their English level.

It can be seen from these accounts that the NKU students seemed to retain an overwhelming enthusiasm for CET4.

Second, the teachers saw a strong testing orientation in the students’ learning practice. The teachers reported witnessing students drilling themselves towards the test despite being informed of the supposed reduction in the importance of the CET4 (NKU-T2); asking constantly when they were allowed to take it (NKU-T2, NKU-T4); and expecting teachers to prepare them for the test (NKU-T3, NKU-T4). The orientation as the teachers understood was due to a combination of the students’ desire to prove their English level (NKU-T3), the social consequences such as employment needs (NKU-T1, NKU-T3, NKU-T7), and the established testing tradition (NKU-T2). Further information about NKU students’ interest in CET4 is provided in the student data in Section 5.3.4.

Based on these insights, the teachers argued for no “substantive” change in students and their learning practice. NKU-T2 commented: “Actually I think for most students the CET4 is virtually compulsory rather than optional”. NKU-T7 added to this point: “There are cases when students give up because of failure to pass. But few students give up willingly without even trying or do not care at all”. This finding that the testing environment and testing-oriented practice remained largely unchanged regardless of the optional nature of the external test and most teachers’ abandonment of test preparation relates to the needs of the students in response to their understandings of the benefits of taking the test. I explore these in more detail in the student data.

**Incorporating formative assessment**

The teachers saw in the incorporation of formative assessment a positive intention, which included catering for individualised learning (NKU-T2); directing more attention to the learning process (NKU-T1, NKU-T7); lowering the stakes of the test (NKU-T5) and tackling the validity issue of the standardised testing format (NKU-T7). However, the teachers shared a cautious attitude towards the use of formative assessment in practice. Several contributing reasons were identified. First, formative assessment was a complex issue; accordingly, the teachers felt that a
practical and feasible system needed to be fully established for it to be effectively carried out.

NKU-T5: The assessment is not fully designed; it did not play a role as it is supposed to. I think it is mostly because the criteria and procedures are not specified and established. Specific details are needed for practicability and feasibility…

NKU-T7: The concept is good, but the practicability…

NKU-T4, NKU-T6: I agree.

Given that no such system was available, the conditions for incorporating formative assessment were not considered mature.

Second, formative assessment, with focus on the individual student, demanded more time and energy input on the teacher’s part. Given the contextual constraints like big class sizes (NKU-T1, NKU-T6), students’ poor English levels (NKU-T7) and the tight schedule (NKU-T2, NKU-T4), it could be difficult to carry out effectively.

Furthermore, the fact that the heavier workload for teachers was ignored in the teaching appraisal system seems to discourage the teachers’ attempts to use it. The view was that a lack of feasible procedures and supportive logistic conditions would lead to disorder and the failure to successfully undertake assessment as envisioned by the policy. These understandings reflect the teachers’ view of the assessment policy as lacking specified criteria, support, training, and misalignment with other policies.

As well, given the small weightings that process assessment was given in the assessment framework, this policy change at the institutional level was scarcely visible. NKU-T4 commented: “Actually I am thinking we don’t have formative assessment for its own sake. Ten percent doesn’t mean anything; it can’t possibly play a role it is supposed to be”. Some teachers were not even aware of the policy change. NKU-T7, for example, commented: “To me, the institutional policy has not made any changes over these years”. With all these stated reasons, the teachers admitted rarely using formative assessment in practice, and hence saw minimal impact from the incorporation of process assessment on their assessment practice. NKU-T4 made this point clear by saying: “Actually we did like this before the process grade is demanded”.

166 Chapter 5: Case study two
The impact of the incorporation of process assessment on students’ learning was similarly perceived as minimal. This is clearly shown in the excerpts below.

Researcher: How about its effect on students?

NKU-T7: Not much.

NKU-T5: No.

NKU-T4: Only 10 points, they don’t quite care.

The teachers mostly believed that even if they would like to take the trouble to use process assessment, the students’ indifference to it because of its small weighting might invalidate the effort. But there was always an exception. NKU-T2 disagreed with this view and predicted that the students would participate consciously when they saw the consequences.

To summarise, the NKU teachers generally understood the intention of the two changes to the policy of making CET4 optional and incorporating process assessment. However, they perceived minimal impact of the changes on their teaching as well as the students’ learning practice. Rather, they saw a testing orientation as still strong and dominant in NKU’s learning environment, which virtually negated the demotion of the CET4 to voluntary status. The lack of established procedures and training, complicated by multiple contextual constraints, made the implementation of the formative assessment ineffective. How the teachers interpreted the policy initiatives might influence their translation into classroom practice, which is mapped out below.

5.2.4 Assessment practices and procedures as enacted by the teachers

Analysis of the teacher assessment practice-related data brings to the surface themes about what and how assessment was practised, which include assessment procedures, criteria use, student involvement, feedback and grading. The focus of the analysis is to detect if and how formative assessment has been enacted in the classroom.

Alignment with the institutional policy

The procedures that the teachers reported to have taken for assessment during the process are attached in Appendix I. As can be seen, the assessment framework in the institutional assessment policy (see Figure 5.1) was mostly taken up by the teachers. Specifically, all the teachers assigned some tasks, used students’
participation in class and attendance for the purposes of process assessment. The alignment between policy and practice was acknowledged. The teachers mostly considered that they were adhering to the institutional requirements about process assessment and followed the suggested distribution. However, the ways that the teachers practised assessment were varied and primarily up to a teacher’s individual preferences, as can be seen from the analysis below.

First, the teachers all assigned tasks, but they did so differently in terms of task types and frequency. NKU-T1, thinking writing as a task type was too hard for the first year students, kept it until the second year. She currently assigned only translation tasks and required her students to hand in three assignments minimally to secure the three points allocated for assignments. Other teachers assigned not only translation but also writing and other tasks; they also reported to be flexible in terms of the number of submissions (NKU-T2, NKU-T4, NKU-T7).

Attendance too was checked in different ways by different teachers. NKU-T2 regarded calling the name lists as a waste of time and never called the roll. She recorded the absence only when the absentee was appointed to answer some questions. Others such as NKU-T1 reported to have called the roll sometimes. Regardless of the different checking practices, one absence when recorded would lead to one point subtracted from the 10% process grade, which was stipulated by the institutional policy. With absences leading to the subtraction of points, attendance appears to have been used to control students. In this sense, it has a punitive function.

In terms of classroom participation, students’ responses to teachers’ questions in class were referred to by all the teachers. But those who used presentation tasks such as NKU-T1, NKU-T3 and NKU-T7 also looked at students’ performance of the task. NKU-T4 reported to have taken students’ attitude and attention into consideration. There seems to be no specific criteria but the teachers shared a concern about the reliability of using classroom participation for assessment, which is discussed later.

The teachers’ assessment practice could be summarised with a quote from NKU-T7: “It [assessment] is conducted according to the need of classroom practices, and the flexible choices and the likes of individual teachers”. In other words,

---

3 The online learning, which was still experimental, was scarcely mentioned by the teachers and hence has not been discussed.
although the teachers mostly took up the assessment framework as stipulated in the institutional assessment policy, they also made use of the flexibility allowed to accommodate the needs in real situations in relation to their professional expertise and experience in the assessment area. The following sections look more closely at how teachers practised assessment in the areas of criteria use, student involvement, feedback and grading.

The use of criteria

This section focuses on what the interview data reveal about the teachers’ use of criteria. Specifically, the practice of the teachers who had used criteria was analysed to identify the criteria and how these criteria were implemented. Of the seven teachers involved, one (NKU-T3) frankly admitted he had relied on his “professional judgement” rather than some specific criteria for an evaluation of students’ performance; two (NKU-T5, NKU-T7) did not clearly state if they used criteria; four (NKU-T2, NKU-T1, NKU-T4, NKU-T6) reported using criteria when evaluating assignment tasks.

The four teachers who gave affirmative responses reported slightly different criteria for different tasks. NKU-T1, for example, stated using “grammar and logic” for translation tasks. For the writing task, NKU-T4 and NKU-T6 reported on “language, structure, and logic”, whereas NKU-T2 referred to language and the requirements she communicated to students when assigning tasks. While these criteria were mostly not specified, NKU-T6 made it clear that she had referred to the CET4 criteria for writing, which include grammar and expression in terms of language and coherence and pertinence to the given topic in terms of content. The communication of these criteria was usually done orally in class. NKU-T2 was the exception in that she had shown students samples of student assignments to indicate her standards for “good” or “excellent”.

Criteria are crucial in that they explicate the requirements of the assessment task and can scaffold students’ engagement with a task (Orsmond et al., 2000; Rust et al., 2003). Yet there seemed to be no common approach to the seven teachers’ design and use of criteria. Some teachers used criteria while others did not; some had the criteria specified while others had it implied and regarded it as commonsense. The teachers’ varied use of criteria is related to the reliability issue of the teachers’ grading practice, which is discussed later.
**Student involvement**

Students’ involvement in assessment and how they were involved is the focus of this section. The data shows that five teachers involved students in assessment processes while two did not. Of the five teachers who gave an affirmative reply, two (NKU-T4, NKU-T6) stated they had tried it only once. As referred to above, NKU-T6 was frustrated to find that students failed to undertake the role of peer assessor properly and gave up the trial based on her view that the students were used to being assessed rather than assessing. Another two teachers, NKU-T2 and NKU-T7, reported prior student involvement experience. NKU-T2 encountered peer feedback when she studied overseas and used it in 2004 when she came back to China for writing assignments with both undergraduate and postgraduate students. But she did not use it with the current students, that is, the students involved in this study. NKU-T7 described her involvement of students in presentation tasks in the previous term and stated that she had learned it from a journal article. She described her practice in detail:

I did not give them grades directly myself. I would let them say for themselves what they think of their report and how many points they think their reports deserve. Then I asked other students in the class to assess and gave a grade. Lastly I gave them a grade according to my criteria…

The other teacher, NKU-T4’s experience was current. He reported “ask[ing] students to rank and choose the best group or best presenter” in writing and presentation tasks. Yet, he admitted doing this only when he could find time. He reported doing this only once that term up to the time of the interview, which was conducted in midterm. NKU-T4 involvement of students in assessment was also revealed in an episode he recounted. A student wrote about his love of smoking in response to the assigned writing topic “love and to be loved”. The essay was well written but NKU-T4, being unhappy about the issue the student chose, gave it a low grade. The student provided his reasons; NKU-T4 subsequently upgraded the results. This shows that NKU-T4 involved students in the negotiation of the assessment results. The episode meanwhile indicates that criteria for assessment were apparently not made clear from the outset, which will be discussed further in the student section.

Another view was heard from NKU-T6 who reasoned: “Actually I think even if the students do not speak out, they are assessing in a sense. They know who is doing
well and who is not…” Also, it is worth noting that during the last observation session of NKU-T2’s class, the teacher did introduce peer review for a writing assignment. The peer review was not able to be followed up at the time of the study because of the teacher’s sudden hospitalisation. However, this attempt did indicate this teacher’s increasing awareness of formative assessment.

All these provide evidence to suggest student involvement in teachers’ assessment practice, though these experiences were occasional and had happened mostly in the past. Professional development and exposure either overseas or through self-development at home appears to be the major contributing factor. Interestingly the teachers made little or no explicit reference to the association between student involvement and formative assessment, which indicates in a way the lack of awareness in this regard.

**Feedback**

The feedback the teachers gave and how they gave it is the focus of this section. All the teachers reported providing comments as feedback, which was usually oral for in-class activities such as presentations and written for translation and writing assignments. However, the ways the feedback were given varied and seemed to relate to the teachers’ values about the students’ proficiency levels and learning approaches. NKU-T2, with a concern for individualism, tended to give different feedback for low achieving and high achieving students. To the former, her comments were more encouraging than negative. This she made succinct in the interview:

> I do not give much negative feedback though, seldom, nearly none… Even when some students do not do well, I will say like this: “your idea is not bad. With better wording, this would be an excellent essay” – in an encouraging tone, you know…

To the latter, however, she would give extra comments like “good” “excellent” but tended to be “pickier” with extra information on specific areas to improve. NKU-T1, differing from NKU-T2’s individualistic style, adopted a collective approach. She preferred to communicate in class the shared problems via some chosen samples. NKU-T4 and NKU-T6, in contrast to other teachers’ overall comments gave detailed comments according to the categories they specified like language, structure, and logic, but they admitted giving feedback like this only when they had enough time.
In the varying feedback practices, the teachers shared an emphasis on evaluation and tended to link it with giving grades. The alignment with the institutional policy is evident therein.

While a general tendency for judgement was apparent in these comments, there appeared also evidence of a valuing of student improvement and learning, for example in the advice NKU-T7 provided to her presenting students:

I usually gave them some advice so that the presenter and other students can do a better presentation. Say, I would suggest them to provide background or even difficult words if necessary so as to ensure the report understood by others…

Similar evidence was shown in the feedback that NKU-T1 provided to the class via assignment samples (see Snapshot One). Whether or not the teachers were aware of the link between feedback and formative assessment as it was defined in the theorists (Black & Wiliam, 1998) and the CECR policy (CMoE, 2007) was not clearly shown. However, the data does indicate that the teachers did utilise feedback to promote learning, which demonstrates a formative nature.

**Grading**

In accordance with the institutional requirements, the teachers made records of students’ performance in the areas of classroom participation, assignment and attendance. Also as required, they produced a process grade according to the records they made, which constituted 10% of the final grade for reporting purposes (a record sheet by NKU-T2 is attached in Appendix J).

The teachers all took it for granted to give students an “objective” grade, which was summative. Nonetheless, their grading practices were diverse. Take the assignment grading for example. Some teachers like NKU-T2 mainly counted the number of assignments that each student submitted and used the number as the grades. She did not give a ranking grade as other teachers did, which she reasoned: “I do not want my students or any students to think that by giving them an A+ or B−, I have put them in a certain ranking”. She chose to leave evaluation of assignment quality to the feedback. Some other teachers such as NKU-T5 and NKU-T4 stated to take both the quality and quantity of assignments into consideration when grading. That means they not only looked at how many assignments each student handed in
but also how well the assignments were completed. While some teachers recorded only three of the eight assigned tasks for each term, which was the institutional requirement, some reported to take in as many as were submitted and focused on encouraging students’ enthusiasm (NKU-T4, NKU-T7, NKU-T6) or with a respect for individual learning styles (NKU-T2). Still, one teacher (NKU-T6) reported to have attached an encouraging value to her grading practice: “To avoid frustrating them, I give C at the least and an A− the most”. The grading practice was summarised by NKU-T7 as “different teachers handle this [grading] differently”.

The teachers shared a concern about the reliability of their diverse grading practice, which resides not only with assignments but also with classroom participation. NKU-T2 expressed her dilemma below:

Say, some students volunteer to answer the questions in class, but give a wrong answer. Would you give a score for that? If yes, a high or a low one? Should the students be regarded as active or not? It is really hard to say. And if you pose strict requirements about participation in classroom activities, some active students might dominate. I have been trying my best to organise as many as possible classroom activities, but still I cannot guarantee fair opportunities for every student. That is virtually impossible within the limited class time and within the limited opportunities…

Based on these grounds and her respect for students’ individual learning styles, NKU-T2 allowed for flexibility within the seven points of classroom participation and assignments. In other words, her students could choose between actively responding in classroom and submitting assignments. For example, one student who kept quiet in the classroom could equally gain the seven points if he or she submitted seven assignments rather than the required three. NKU-S4 is a case in point (see Section 5.3.3).

The diversity of grading practice raised a critical issue about the maintenance of standards which relates to the intervention of the teaching administrative sector as aforementioned by the senior administrator in Section 5.1.3. While understanding that the TAO’s intervention was intended to keep standards, the teachers stated three major reasons to justify their practice. The small percentage allocated to formative assessment was listed the number one reason by the teachers. According to them,
given the fine-grained judgement needed for the 7%, the teacher did not necessarily abuse the power. The excerpt from NKU-T4 interview illustrated this point.

> Of the 10%, only 7% is of the teachers’ professional judgement. You know, 3 percent is for attendance. If the student is active in class and hands in all the three assignments and is not absent as you might notice, he/she will surely get all the 10.

The teachers related the slim weightings further to the institutional control, which seems to have affected the teachers’ grading practice in a negative sense, as is evidenced below.

> NKU-T7: The day before, we proposed to the dean that 10% was too little. It seems the authority does not trust us teachers. They are afraid the teachers will abuse the power. On this point, I strongly disagree. I think since you distrust me, I will not bother to weight between 10 and 9.

It seems the distrust from the administrative body has aroused resistance in teachers, which as indicated by research (Carless, 2009; Wang & Cheng, 2005) could be detrimental to any innovation attempt including this one.

The reliability issue was also regarded as related to the lack of practical procedures and specified criteria in the assessment system, as NKU-T5 expressed:

> What is lacking is a practical implementation procedure. If the specifications are prescribed in details, well, it is easy to carry it out. But if not, it is hard to say. 4 points for classroom participation – then what a 4 points should be like and what a 3 points is like. It is sort of abstract rather than concrete.

She maintained that given the lack of specified criteria the teacher’s behaviour was understandable. With a need for criteria, the teachers expressed a view that attempts to use formative assessment for objective measurement. This aligns with the institutional policy that prescribed the use of process grades for reporting purposes but contradicts the perceived formative purpose of assessment in enhancing learning (Black, et al., 2003; Black & Wiliam, 1998). While the institutional and technical reasons such as the teachers’ individualised grading practice, institutional control and regulation, and the lack of specified criteria are responsible for the tension over reliability, the attempted use of formative assessment for objective and summative measure purposes can be identified as another major factor leading to the results.
Snapshots of classroom assessment practice

Three snapshots of classroom observation were reproduced in this section for better understanding of how assessment was practised at NKU. The first is NKU-T1’s feedback practice on a translation task, which also showcased her use of criteria and student involvement. The second was of a change that NKU-T1 introduced in response to the changes in the new CET4. The third captures NKU-T2’s feedback practice on an assigned task.

Snapshot One

Time: 8:00 – 10:00am (Beijing time), 18th March, 2009

NKU-T1 shows students some samples with projector and comments while reading:

Can “worthwhile” be used this way? Try to make sure if it can be used this way before you actually use it, ok? What’s the difference between “worthy”, “worth” and “worthwhile”? ....

NKU-T1 gives a detailed explanation between the meanings and usage of the three words then she comments on:

Now this is another sample. Look, “largest number” is alright, but “the” is needed here.

The present is talked about, not the past. So use “gets” rather than “got”

I have made the following comment on this assignment: “try to make your English correct grammatically and acceptable socially”. Grammatical correctness is not enough, it is also very important to make sure your English is acceptable according to social conventions...

Please ignore the marks you’ve got. Marks are by no means the most important. The point is you attend to the errors and make sure you know why and how to improve so that you will not make the same mistakes again.

When NKU-T1 commented on the chosen sample, she referred to word families and vocabulary learning and grammar such as tenses, articles, and sociolinguistic appropriateness in language use. She used the criteria she mentioned in the interview: grammar and logic. Her feedback was intended for better learning,
which she clearly stated in the concluding comment (A writing assignment sample with NKU-T1’s comments is attached in Appendix K). She did refer to a grade but also provided the students with more extended feedback with an emphasis on learning. The intention to direct students’ attention from scores to learning itself was thus clearly expressed. She tried to engage students by constant questioning. However, no effective interactions actually happened and the feedback was primarily the teacher’s monologue.

Juxtaposed with NKU-T1’s conscious effort to shift the focus to learning and its process is the occasional testing-oriented practice, which is captured in Snapshot Two.

<table>
<thead>
<tr>
<th>Snapshot Two</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time: 8:00 – 10:00am (Beijing time), 12th April, 2009</td>
</tr>
<tr>
<td>NKU-T1 is going through the after-text exercises with students in class, when she introduces a change to the way to do the exercise. She also communicates to students the rationale for such a change and some specific strategies.</td>
</tr>
</tbody>
</table>

As to this exercise, I used to let you do it yourself, then raise questions if you have any. Now we need to change a little bit because a new item in the new CET4 requires test-takers to choose 10 from the 15 provided words. Moreover, this item was reported to be the one that most students lost the most of the scores. So we’ll take this exercise out and deal with it specifically. Actually, you have ways to follow in this kind of exercise. You need first of all to scan the words provided. Knowing the meaning of the words is not enough. You also need to know of the parts of speech of each word to get it right. This is very important. Now let’s try and do it to get some sense out of it.

Then NKU-T1 went through the exercise with students.

The two snapshots seem to indicate that NKU-T1 was involved in juggling an interest in promoting students’ learning approaches and preparing them for the specific features of the CET4 test. On the one hand, she emphasised learning and tried to direct students’ attention to the learning process, which she personally valued (see Section 5.2.2 Teacher perceptions). On the other hand, she made the change as
the situations called for it so as to prepare the students for the CET4. This she did to come to terms with the students’ expectations for good test results. This shows that the pressure from the external test, which was now optional in policy, still existed and tended to affect the classroom practices at least to a certain degree. It also showed her acknowledgment of the continued importance of the test for her students, despite the fact that it had been made optional at policy and institutional levels.

A similar dilemma can be seen in Snapshot Three, which showcased NKU-T2’s feedback practice.

**Snapshot Three**

Time: 8:00 – 10:00am (Beijing time), 24th March, 2009

The week before, NKU-T2 asked students to find out the definitions of some terms such as “civil rights”. Now NKU-T2 was giving feedback on this task:

Some students have handed in their assignments for unit 2. I looked through some and found that many students wrote the definitions themselves. What you are supposed to do is to find the definitions. The skills you are supposed to use include looking up in the dictionary or searching websites for what the authoritative explanation of the concepts. Do not produce what you understand about it. If you write in your own way, you are definitely not going to get high scores.

As can be seen, the feedback that NKU-T2 provided was to clear the students’ misunderstandings so that the task could be completed as intended: to find the definition via various resources rather than try to produce one by themselves. It was intended for better accomplishment of the assignment, and hence appeared to be focused on learning. The last sentence, however, revealed another purpose that NKU-T2 intended for the feedback: better assessment results.

### 5.2.5 Discussion

Analysis of the teachers’ perceptions showed that most of the teachers seem to have taken the role shift that resides with a constructivist view of learning at least in understanding but not quite with assessment practice. The teachers seem to be conscious of the function of assessment for improvement and some of them were aware of the benefits of using feedback and engaging students. But the awareness of
the link between formative assessment and feedback and involving learners in assessment, as commonly defined (Black et al., 2003; Black & Wiliam, 1998) were not demonstrated.

The teachers mostly seem to understand the intention of making CET4 optional and incorporating process assessment and were able to see the multiple potential benefits. However, they shared concerns about the implementation of both. For the former, they were concerned about its possible negative impact on students’ motivation to learn English. For the latter, they were concerned about its lack of established procedures and criteria and hence shared a cautious attitude to its use in practice.

Above all, the teachers perceived minimal impact from both policy changes on their practice except the abandonment of test preparation practice. Actually, they perceived a continued “yardstick” role of the test in their teaching and students’ learning practice. The data also showed one teacher (NKU-T6) still actively taught for the test and some teachers such as NKU-T2 and NKU-T1 managed to juggle learning and testing because of their awareness of its continued importance for students. The teachers saw the compulsory status of CET4 virtually unchanged in their context.

As to the incorporated process assessment, they admitted doing as regulated by the institutional policy and reserved a cautious attitude towards it because of a concern for its reliability. Nonetheless, the analysis of teachers’ assessment practice revealed the formative approaches to assessment in multiple aspects. For example, there was evidence of using criteria, the standards of which however were not specified and uniformly understood by the teachers. There was evidence of student involvement in assessment although it was mostly self-initiated by some teachers rather than in response to the initiated policy change. There was evidence of focus on learning and its process which was related to an emphasis on learning outcomes in the tests and their results. Finally, there was evidence of feedback for improvement purposes. The point was that teachers were undertaking many of the tasks and activities aligned with formative assessment and its function of promoting learning but did not recognise the term or its role in the policy reform. The fact that the teachers were actually doing assessment for learning but without awareness of the theoretical nature of their assessment practice or an informed understanding of its
role and enactment indicates a strong need for professional development in the area of assessment and formative assessment in particular, especially as it is envisaged in the CECR policy.

Whether these apparent formative assessment practices actually enhanced students’ learning or not needs to be substantiated with data from the students. The next section examines the data from students to shed more light on the enactment of the policy changes and the students’ experience.

5.3 STUDENT LEVEL RESPONSES

Informed by a sociocultural perspective, the investigation of the students’ responses to the College English assessment policy change starts with the informants’ profiles (Section 5.3.1) and their perception regarding ELL and assessment (Section 5.3.2). This is followed by an examination of their experience of College English assessment (Section 5.3.3) and their response to this change (Section 5.3.4). The research question that relates to the impact of the policy change on students - their views and responses - is addressed in the discussion (Section 5.3.5). The findings as presented below are based on individual interviews conducted with five students.

5.3.1 Student profiles

NKU-S1

NKU-S1, a female student, came from Shijiazhuang, the capital city of Hebei Province, a level three region (China Investigation & Assessment Centre, 2008). Her ELL started in primary school and she retained a strong interest in it. According to NKU-S1, her English has been “standing high above her peers from the very beginning and remained ever after”. However, she did not spend much time on it. What she did was to concentrate in class and take detailed notes and then memorise new words afterwards. She did consciously make efforts to cultivate a feel for the language though; for example, she watched English movies and TV programmes trying to imitate the pronunciation. And she was always “happy” to address any gap she became aware of. She described learning English as “easy”.

The high school that NKU-S1 attended was a key school at the provincial level, which means it was one of the best within the province. She also attended several English training programs, which provided her with extracurricular knowledge of
aspects of English. She identified herself as “an active responder” in class and had the habit of “taking the front seat”, which made it convenient for her to “communicate with teachers in both verbal and body language”. She appeared confident and was proud of her English level.

**NKU-S2**

NKU-S2, a male student, came from Wenzhou, a mid-sized but developed city of Zhejiang Province which was a level two region according to the 2008 China Development Index (China Investigation & Assessment Center, 2008). He stated that he learned English out of interest rather than to get high grades in tests. His ELL started from year three in primary school when he learned some basic vocabulary for daily use. He gained greater confidence in English from learning phonetics in a training program that he attended in the vacation between primary and junior high schools. NKU-S2 attended a key high school at the municipal level. During high school years he was acknowledged as “one of the top students in English”. However, he did not have the enthusiasm to participate in College English classroom activities and tended to regard himself as “not quite active”. The factors that influenced his participation included “laziness”, “interest in the provided topic” and lack of opportunities. Irrespective of his inactivity in classroom, NKU-S2 devoted a lot of time to ELL with a goal of studying further abroad.

**NKU-S3**

NKU-S3, another male student, was from a small city in the north Shanxi Province. He attended a non-key high school, where the students were streamed into “experimental or parallel class” system according to their test results in the entrance examination for the purpose of higher university entrance rates. The students could be either promoted or demoted according to their test results at the end of an academic year. NKU-S3 started in a parallel, that is, a non-key class, and was promoted to an experimental class in the second year. According to NKU-S3, “The whole system was oriented towards examinations and only the test results talk”. Specifically, vocabulary and grammar explication were the focus of English teaching and English classes were mainly conducted in Chinese. Test drilling in cloze and

---

4 This is a literal translation of the terms used to express euphemistically the key and non-key stratification of classes within the school. Usually the high achieving students, who are regarded as having more potential to go to university, attended the experimental classes, to which more experienced teachers are assigned in contrast to parallel classes, which are staffed by less experienced teachers.
reading was practiced when exams approached. And content not tested in the EEHE was not covered, as he said:

We did not ever have any speaking and listening classes at all because the speaking and listening test results are not included in HEEE except for those who want to learn English as a major. Our teacher told us not to put any effort into it.

This practice, combined with the contextual constraints such as lack of listening material and speaking opportunities, led directly to his current weakness in speaking and listening. He described his English level as “so so”, but he regarded himself as “an active learner” on the grounds that he organised his study himself and “never need[ed] to be urged by the teacher”.

**NKU-S4**

NKU-S4, a female student, came from a small city in the south of Shanxi Province. Although she did not attend a key high school, she was included in one of the elite classes\(^5\). She started to develop a keen interest in ELL and liked writing particularly. Her frequent writing and detailed feedback from the teacher had helped improve her English level, which in turn brought her good grades. She said:

I have been quite at ease in doing all kinds of exercises and taking various tests. I have a relatively sound knowledge foundation, so did not find much difficulty in gaining good marks or worry about failure in any test.

Although she achieved good results in tests, NKU-S4 she did not think tests could assess what was learned; rather she held writing highly in this regard. Limited opportunities to use the language were mentioned as a contributing factor to her relative weakness in speaking and listening. Regardless, NKU-S4 still identified herself as “an active learner”, whom the teacher “did not need to worry about”.

**NKU-S5**

NKU-S5, female, was also from Yuncheng City of Shanxi Province. She started her ELL in junior high when she moved to Yuncheng City from a small county. This relatively late start placed her at a disadvantage compared to her classmates. Nonetheless, she managed to enrol in a key high school at the provincial level, which was well-known for its high EEHE admission rates. The teaching mode

---

\(^5\) An alternative expression of the key and non-key classification of classes within a school.
there was open and free until recently when the admission rates dropped and a new
principal came to power. Then the students were streamed into a key and non-key
system, within which emphasis was directed towards the EEHE. NKU-S5 was in one
of the non-key classes. She described her English speaking and listening as “poor”
and attributed this to their unimportant positions in the EEHE and limited access to
listening opportunities in the learning context. This perceived weakness, combined
with her “introverted” character, led her to participate in classroom activities only to
an “average” degree.

As the above profiles make clear, the five NKU students’ ELL experiences
were similar in some ways. Firstly, they all had formal English lessons at school.
However, the English education they experienced and their feelings towards it varied.
Students such as NKU-S1 and NKU-S2, who were from more developed regions,
started their ELL early with an interest in it and seemed to have access to more and
better learning opportunities. Hence they did well in ELL and felt confident about
their English proficiency. NKU-S3 and NKU-S5 started to learn English late and
found ELL a struggle in high school as well as in university. The other student,
NKU-S4, did well in the non-key school she attended but lost confidence in her EL
ability at university when communicative skills became more important. These three
students, all from under-developed areas, unanimously mentioned the negative
influence of contextual constraints such as the availability of facilities and access to
authentic learning opportunities.

Secondly, all five students have gone through an EEHE-oriented practice;
however, the washback effects of the EEHE on their ELL again differed and tended
to relate to their educational backgrounds. For NKU-S1 and NKU-S2, the impact did
exist but was not as severe as for NKU-S3, NKU-S4 and NKU-S5, the other three
who were from under-developed regions. For them, the EEHE-oriented practice
which emphasised test-wise skills at the expense of speaking and listening, which
resulted in weak oral communicative skills. It seems the ELL experience had exerted
an influence on the students’ achievement and the ways they perceived and
approached College English learning and assessment, which is the focus of this study.

5.3.2 Student perceptions

Themes that emerged about the students’ perceptions relate to ELL, assessment,
and the roles of teachers and learners.
English Language Learning

Several insights emerged from the interview data about the students’ perceptions of ELL. First, the students all regarded ELL as very important to them, but they valued it for different reasons. ELL was a subject to tackle for some students like NKU-S3, whereas NKU-S2 viewed it a means by which he could communicate with people from other countries. Second, the students saw an interest in ELL as crucial for gaining a high level of proficiency (NKU-S1, NKU-S2, and NKU-S4). This point was evident in the excerpt from NKU-S2’s interview: “I think interest is important. Otherwise you can’t put your heart in it. You see, those who are not interested in English are scarcely good at English”. Third, vocabulary was a big concern for most of the students, and memorisation and recitation were widely used for the purpose of developing vocabulary; however, the students differed in what and how to memorise: NKU-S3 was used to rote learning from the textbook, whereas NKU-S2 tended to rely on a dictionary and other resources. Last, opportunities to use English were valued by NKU-S2 and NKU-S5, who went to the campus-based English corner occasionally or went to chat online with English speakers, but not by others. NKU-S1 and NKU-S3, for example, did not like to go to English corner at all. Just like some teachers, these students’ perceptions of the value of communication in English corner as a way of ELL influenced their behaviours.

The roles of teacher and learner in ELL

The data revealed two different understandings about the role of teacher and learner. The first, held by NKU-S3 and NKU-S5 placed teachers in an authoritative position and they tended to rely on them for knowledge and guidance. What NKU-S5 said was representative of this understanding: “They guide us. Most knowledge we get is from the teachers. They are just like live dictionaries for me”. What a learner needed to do in this view was to follow the teacher. This understanding reflects the Chinese traditional value of the teacher as an authoritative figure and the hierarchical teacher-student relationship (Biggs, 1996; Ho et al., 2001a; Hu, 2002).

The teacher’s authority was recognised by the second understanding, which was held by NKU-S1, NKU-S2 and NKU-S4. They regarded the teacher’s role as important, especially for beginners. However, this view listed its importance as “second to” that of the learner him/herself. That means while recognising the teacher especially a good teacher’s importance in scaffolding students’ learning, this view...
also highlighted the active and even decisive role of learner in the process. This point is illustrated below.

   NKU-S1: To me, teacher plays a role only second to learner himself/herself. After all, teacher is equipped with richer knowledge and has experienced ELL him/herself. Both their knowledge and experience are quite helpful for us beginners.

   Those who held this view tended to take up the teachers’ guidance critically. This was evidenced when NKU-S2 said: “I do not exactly follow teacher’s guidance”. This view assumes an active role for the learner, and thus indicates a change of the stereotyped teacher-student relationship (Cortazzi & Jin, 1996). The students’ understanding of the roles that the teacher and learner took in ELL influences the approach they adopted to learn it.

   The roles of teacher and learner in assessment

   The role of teacher as assessor was taken for granted by the students. This was evidenced when NKU-S3 said: “We students used to be assessed by teacher and the examinations. Actually we are quite used to this. I personally think the teacher can give us more objective assessment…” It seems the prior experience has impressed on these students that assessment was mainly the responsibility of the teacher not the student.

   While the students agreed that assessment was mainly the teacher’s responsibility, their opinions were divided over the role that the learner plays in assessment. The role of the learner in assessment sounded new to students like NKU-S3 and NKU-S5; they did not seem to think they as learners had the proper expertise or were in the right position to assess their or other students’ work. When asked about engaging in assessment, NKU-S3 replied: “I have never thought of it before… I have never been put into that position…” A different perspective on the role of the learner in assessment came from NKU-S2 and NKU-S4. They highlighted the importance of self reflection and stated that they had used it to regulate their ELL. This was reflected in NKU-S2’s comments: “You know yourself the best. Other people see only some aspects of you, like how you study. You know clearly where you are and what you are like”. These students were aware of the active role that a learner could play in assessment and its value in monitoring their study. Nonetheless, supporters of both views valued the feedback from teachers. It seems despite the deep-rooted
ideology of teacher as an authoritative figure (Biggs, 1996; Ho et al., 2001a; Hu, 2002) in learning as well as in assessment, some students began to recognise their autonomous role and used to regulate their ELL. How the students perceived the roles that the teacher and the learner play in assessment appears to be reflected in their engagement with assessment activities.

**Assessment**

The five students interviewed at NKU identified testing as the major means that had been used to assess their ELL and hence tended to view assessment as synonymous with testing. The impact of tests including the EEHE on their ELL was well recognised, which was evidenced in NKU-S3, NKU-S4, and NKU-S5’s weakness in listening and speaking. It is also clearly shown in the excerpts below.

NKU-S1: Well, I have been learning this way all these years. Influence [of tests] is quite strong, but whatever you say and do, tests are there and “grades are students’ life”. So, we have got to pay special concern for tests and the test content.

Acknowledging the washback effect of testing on ELL, NKU-S1 and NKU-S2 reported adjusting their ELL approaches to cope with tests, yet they both managed to keep their focus mainly on learning. Comparatively, the impact on other students seems to be more severe. NKU-S3 stated that testing had influenced not only what they learned but also how they learned it, as he made manifest by saying: “…I spent most of the time on textbooks, reciting, memorising … The materials I choose are mostly that of testing materials, seldom go beyond that”. There were also reports that testing had influenced how much they learned. NKU-S2 said, “To me, I can scarcely learn anything when I study for tests”. By saying this, he recognised that test-preparation was not learning. With high stakes usually attached to the test results, the testing context, in which the students were immersed affected also their feelings and identity. This is illustrated in the quote below and will be elaborated upon in detail in the discussion chapter.

NKU-S2: Tests hang over everything. It is something that cannot possibly be changed. Test results decide everything as a consequence and make the difference of confidence and frustration for one. I don’t like it but good test results make me happy and proud of myself.
In contrast to the general disapproval of the testing culture and its impact, there is one student who thought positively of exams. NKU-S5 for instance, asked: “I am puzzled. Is testing harmful in any possible ways?” Being immersed in a testing culture context so far, this student seems to have taken it for granted and did not even give it a second thought.

To summarise, the students seem to feel the washback effects of a testing culture, which tend to influence not only what they learned, how they learned it but also how much they learned and their identity. Despite some individual student’s positive attitude toward tests, there seems to be a general doubt about the effectiveness of testing as the only means to measure as well as enhance ELL. Prior experience, as an influencing factor of one’s later behaviours (Vygotsky, 1978), will influence their engagement with assessment activities.

5.3.3 College English assessment as experienced by the students

Analysis of the five students’ experience of institutional assessment highlighted five major themes, which are about assessment procedures, criteria use, student involvement, feedback and grading. The key findings on each of the themes are presented below from the student data. As the five students were from NKU-T2’s class, the part of the experience that concerned teachers’ practice was limited to NKU-T2 only. It must therefore be acknowledged that the findings may not be representative of the experience of all students.

Assessment procedures

Analysis of the data first of all revealed that the students were quite clear about the assessment framework, which included process assessment and the final term exam, the specific assessment procedures to be taken and the weightings allocated to each procedure. NKU-S1’s response below was an exemplar in this regard.

The assessment included the grades during the semester plus the grades from the final exam at the end of the semester. We usually gain process grades during the semester from writing assignments and participation in class as our teacher judges it. At the end of semester, we take a final term exam, which includes speaking, listening, reading, writing, cloze, grammar and structure etc. The standardized format, you know. The speaking test accounts for 10%. The rest is written.
As can be seen, the assessment procedures that the students experienced agreed with the institutional assessment policy (see Figure 5.1) and confirmed the incorporation of process assessment that the teachers reported (Section 5.2.4). However, the students were also conscious of the teachers’ varied adoption of process assessment procedures such as attendance. To illustrate:

Researcher: Does the overall grade include your attendance rate?

NKU-S1: To be frank, other teachers do, but our teacher does not. She doesn’t keep records, she knows who is absent who is not though.

The teachers’ varied adoption of the assessment policy did not seem to be a major concern to these students as it did to the KU students. This could be because of the minor weightings of process grades (10%) as compared to 60% at KU and will be discussed in Chapter 6.

The use of criteria

In contrast to the teachers’ reports about their use of criteria (Section 5.2.4), the students demonstrated a lack of knowledge about the criteria used to assess their College English. This was reiterated by four of the five students in their interviews.

NKU-S2: No, I have no idea about it at all.

NKU-S3: No, I have never thought of it before. The teacher has never told us either.

NKU-S4: No, I don’t know anything about it. Teacher told us about when and how and what is to be tested and we follow. That’s all.

NKU-S5: We are not quite clear about how each mark comes about at all. Sometimes we wonder if the teacher has given wrong marks. But think more probably it is because of our carelessness.

NKU-S1, the only student who gave an affirmative reply to this question, confused assessment criteria with the assessment framework and tended to see scores as synonymous with criteria.

In my understanding, the criteria are the process grades and the final term exam grades. The procedure is also process grades and final term exam grades. That’s it.

It seems it is justified to say that the assessment criteria were kept in “a black box” (Black & Wiliam, 1998) rather than made transparent to the students. They
were not informed of the criteria; some of them like NKU-S3 were not conscious of the need to know. When doubts about their assessment results arose, they tended to attribute them to their carelessness. In short, these students assumed a passive position of the one being assessed rather than being actively involved in the process of assessment.

**Student involvement**

Similar to the findings from Chapter 4, the data analysis firstly made apparent an ambiguous understanding of concepts. Self and peer-assessment as defined in Chapter 4 did not appear to be used in the students’ prior and current assessment experience. Nonetheless, some students (NKU-S2, NKU-S4) spoke of self-reflection which they used to manage their own study, and also mentioned instances of peer feedback which they used on informal occasions to communicate with each other. NKU-S4, for example, reported to have helped her roommates with editing their writing in some informal occasions and was able to see the value of it. She commented: “It is a review of my previously attained knowledge, and also a reflection of my own writing. If I write, I will try to avoid making the same or similar mistakes”. But overall, it seems these students’ experiences of being involved by the teacher in assessment were limited.

**Feedback**

Several insights emerged about the feedback that students received from their teacher. First, the teacher tended to use a lot of marking when giving feedback to writing assignments and language accuracy was the focus. A representative remark was from NKU-S3, who said: “She [the teacher] will usually cross the good sentences and correct the spelling or grammatical mistakes”. The feedback on language was taken up by some, such as NKU-S5, who said she would usually correct the errors and try to perfect her writing further, but others did not report further actions.

Second, comments were brief and used for differentiating purposes.

- NKU-S1: Very brief comments like “good” “excellent”, no more.
- NKU-S5: The teacher usually praised the students who write well. As for me, I only got my weaknesses.
The students seem aware of the differentiating value attached to the positive comments. That means assignments, if accomplished well, will receive extra comments like “good” or “excellent”, but not these kinds of comments when not well accomplished. This differentiating value was taken up by some like NKU-S4, who expressed her delight in receiving one:

I am especially fond of the assignment that teacher gave me a “good”. That is kind of recognition of my writing. You know in high school I got always “very good” while now “good” is a rare thing. It made me happy.

This role of feedback in motivation resonated with NKU-S2, who expressed his valuing of teachers’ comments in influencing his feelings and identity:

Teacher’s comments are critical to me. Good comments will boost my confidence and interest while bad comments will make me low and trigger me to re-evaluate myself and think of what to do to improve.

Third, the feedback was seen to lack instructive value. NKU-S2, for example, was disappointed to find that the brief comments he received were not instructive in any way, he therefore ignored them. The disappointment was shared by NKU-S1, who expressed a wish for more and constructive advice for improvement, saying: “I hope the teacher can give us some comments on our oral participation in terms of the pronunciation, the appropriateness and give us more advice on how to improve …” She also called for more feedback that reflected the quality of assignments:

I am hoping for more focus on the quality rather than the quantity of assignment tasks. I know teacher means to encourage everyone to participate, only if we participate, things will get better. But I still hope teacher can also value the quality. That can be encouragement for the good to work for the better.

It seems for those for whom English was a struggle, the feedback that the teacher gave served the purpose of improving language accuracy and hence was valued. But for those whose English level was more advanced, the feedback could not serve the same purpose and was thus regarded as insufficient. They asked for more constructive advice. Therefore, the teachers’ feedback seems to meet only the needs of the students of lower level but not that of students of more advanced level of English proficiency.
Grading

The students reported having experienced a grading practice that mainly focused on quantity rather than quality. This was evidenced in the excerpts below.

NKU-S1: We get one point for each time we submit the writing assignments and one point for each time we answer the teacher’s questions in class.

Researcher: Do you mean only the quantity matters, not quality?

NKU-S1: Yeah, it seems so. She will encourage if it is good; but she will give the points only if we submit or only if we respond. After all, there are 10 points altogether for the process evaluation, not much.

This corresponds with what NKU-T2 reported in Section 5.2.4. This kind of grading practice did not seem to motivate the students much to produce quality work. NKU-S1 made this point manifest.

NKU-S1: After you get your 10 points you will find you are no longer motivated to submit or respond, since you will not get more.

Researcher: Do you mean if more points are allocated….?

NKU-S1: Of course. We would work harder for the points’ sake at least. Also, if the criteria for points are more of quality than quantity, we will work more for quality.

It seems the teacher’s emphasis on quantity rather than quality while giving grades has led to a similar focus in the students’ learning behaviours. Actually, seen from the above quotes from NKU-S1, the emphasis on quantity in grading seems to have de-motivated rather than motivated the students to learn. In this sense, the grading practices did not seem to effectively enhance learning; rather it hindered improvement. It did not seem to measure learning either, since what counted was only the number of assignments rather than how well the assignments were completed. This goes against the teacher’s intention and expectation (see Grading in Section 5.2.4). This finding in a way reflects the fact that the scores are still the predominant concern, directing the students’ responses.

The above documentation of students’ lived experience of assessment provides confirmation for the teachers’ account of their enactment of the policy initiatives. It also draws a picture of the students’ understanding of the assessment as practised. How the students responded to cope with the enacted assessment is now the focus.
5.3.4 Student responsive behaviours

Students’ responsive behaviours to cope with the now optional CET4 and the internal assessment framework are analysed in this section. The focus is to find out if and to what degree the implemented assessment policy has impacted on the students’ learning behaviours.

The external test

Analysis of the data showed that the students were not well informed about the changed status of the external test. First of all, they were not aware of the optional nature of the external test. NKU-S1, for example, appeared to still believe that the test was compulsory.

Researcher: Now that CET4 is no longer compulsory …?

NKU-S1: Is that true? I don’t know at all.

The disconnection of CET4 with qualification requirements was not known by the students either. This is evidenced with the excerpts below.

Researcher: So you think CET4 is still linked to the graduation or degree?

NKU-S4: Yes.

Their understanding about English requirements came from senior students rather than institutional policy. NKU-S5 replied when asked why he took the CET4: “Yes. We have to take it. We are told by senior students that we must pass the CET4, even the CET6”. The lack of communication of this policy change provides a good reason for the fact that the five students registered and would take the test in the coming June.

However, this reason could not fully explain the students’ taking the test. All the students expressed a wish to take CET4 regardless of its required status. More reasons for their taking the test elicited from the interviews include firstly, the fact that students regarded the CET4 score as indicative of their English proficiency from the authority. The second reason related to the social need for employment (NKU-S1, NKU-S2, NKU-S5). A typical expression of this point was from NKU-S5, who said: “Of course I will take it. Even though it is separated from the degree, it is not when you look for your job”. The third reason related to the need for taking the test for further study. This was stated by NKU-S2 and NKU-S3, who noted: “I am planning
to do higher degree studies after my graduation. A CET4 or CET6 certificate will bring me closer to that goal”. Expectations from teachers were also mentioned by NKU-S3: “Our English teacher keeps reminding us that CET4 is approaching...” The pressure from the students themselves was an overt influence too, as can be seen from the excerpts below.

NKU-S4: I have been thinking of taking CET4 in high school. At that time CET4 was a goal that I set for my ELL in university. But now to be frank, I don’t think much of tests these days. But I have to take it.

Researcher: Why?

NKU-S4: I am compelled to.

Researcher: Anybody coerced you?

NKU-S4: (Laughing) Anyway, I sure will pass CET4 during university years. It is kind of settle-down of the business. If not, I will not be able to put my heart at ease.

The students’ lack of knowledge about the policy change and registration for the test prior to the interviews were major factors influencing their responses. Nonetheless, the lack of communication does not seem to explain the students’ enthusiasm in taking the test regardless of its optional status. This could be due to the examination-oriented system and the testing culture they were in. As well, the pressure from the demands of social background, expectations of the teacher and even the students themselves seems paramount and more likely to influence their behaviours. Under these pressures, the students did not seem to regard the external test as something that they could choose to do or not to do. Rather, they saw taking the test as essential. It is justified to say CET4, the external test, is still a dominant goal of ELL for the five students interviewed at least. This finding, along with what the teachers observed (see Section 5.2.4), indicates that making CET4 optional at the policy level exerted virtually no impact on the students’ test-taking situation and seemingly little impact on the learning environment at NKU.

**Process assessment**

The students’ responsive behaviours to the incorporated process assessment are analysed in this section. Also included is the perceived impact of process assessment on the students’ ELL. The data first of all highlights a concern that the five students
shared over the assessment results. The students reported adopting a strategic approach to cope with process assessment for the purpose of high grades. However, the degree of purposefulness varied. It was obvious in NKU-S3, a student struggling to catch up and weak in oral communicative skills, as was shown in his words below.

    The classroom participation involves speaking mainly. As I did not have speaking and listening classes in high school, I am weak in this aspect... but the teacher usually ticks off if students speak out in classroom. So I sometimes try to respond when the teacher asks questions.

    NKU-S3 admitted that his participation in the classroom, a struggle as it was, had “in a way turn[ed] out to be kind of practice for my oral communicative ability”. NKU-S4, who shared the weakness in speaking, responded differently. She chose to “sit and listen quietly”, but she managed to “hand in more assignments than required so as to secure a good grade”.

    The purposefulness was not quite obvious in NKU-S1 and NKU-S2, two students whose English proficiency level was higher. They reported minimal impact of the current process assessment on their learning approaches.

    NKU-S1: Not much difference. I participate too much as always. Actually participating in classroom is a habit for me.

    NKU-S2: I hand in assignments as required. That’s all.

    Considering the weightings of process assessment were only 10%, the students were asked about the possible impact of larger weightings on their learning for a better understanding of the students’ responsive behaviours. The findings are as follows.

    The possible impact on learning approaches from larger weightings was refuted by an active participator like NKU-S1. But those who did not participate actively such as NKU-S2 thought more weighting of process assessment would exert an impact on their in-class learning behaviours.

    NKU-S2: I am not active in involving classroom activities. If more weight of the overall evaluation is given to classroom participation, I will try better to participate, for the sake of grades. As we are told from early years on, grades are what matters at least it is so in China. And I did not see any sign of change. It is no exception for me.
The excerpt made it again manifest that the scores were still the students’ biggest concern; it was this concern that triggered the possible change in their classroom participations. But the impact was limited to the in-class behaviours only. The possible influence on the students’ out-of class approach to ELL was not recognised.

NKU-S5: In class, it will. But out of class, it has little influence. Since we have learned English for about ten years, we have formed our own ways of learning English. We might just adjust our learning ways a little.

This student appeared to keep two types of learning approaches to balance the different needs for testing and learning. This resonated with NKU-S2, who said: “In classroom I study for the sake of test. But I learn for my interest’s sake after class. That’s it … My English proficiency mainly comes from interest-ignited and self-instructed study”. Overall, a concern over the scores appears to dominate the students’ responsive behaviours to the implemented assessment policy. Larger weighting to process assessment seems to potentially influence the students’ learning behaviours only to a certain degree.

5.3.5 Discussion

The above analysis reveals several major findings that address the research question about the response to the policy changes at the student level in this context. First, the external test-taking practice remains largely unchanged. All the students expressed a desire to take CET4, the external test that was now optional and disconnected from the qualification requirements. The lack of communication of the policy change to the students, who mistakenly regarded it as compulsory to their qualifications or further study, was only one of the many reasons accounting for this finding. The pressure of social consequences such as the employment, the expectation of teachers and even themselves, and the influence of a culture that values the product of learning as opposed to the process were also responsible for this result.

The impact of the incorporated process assessment was not evident generally and varied between students. Those who were active and focused more on learning did not seem to be influenced by the process grade. But the in-class learning behaviours of those who were not quite active or focused more on scores appeared to be more likely to be influenced. The minor weightings that was allocated to the
process and the teachers’ feedback and grading practice, which emphasised quantity rather than quality, were problematised for leading to this result. Moreover, the learning habit and the focus on learning results that the students had established previously were also found to be responsible. This finding indicates that a testing culture which emphasises the summative assessment results still dominates the students’ learning practices and the learning environment at NKU. However, there are students who reported to have focused on the learning itself (NKU-S1) and students who reported to balance both learning and testing purposes via differed approaches (NKU-S2, NKU-S5). This indicates the existence of a learning orientation together with the testing oriented mainstream practice.

5.4 SUMMARY

Analysis of the implementation of the assessment policy change at the local context of NKU again showcases a reshaping process that was closely related to the local actors’ appropriation of the meaning of the policy within the situated cultural, institutional, and historical context (Wertsch etc., 1995).

The appropriation at the institutional level resulted in the adoption of the policy initiative of making the external test optional as advocated in the CERP. The formative assessment initiative, however, was implemented in the form of allocating weightings of the assessment grade to the process. Students’ involvement in assessment and feedback, which were imbedded in formative assessment and encouraged in the national policy, was not emphasised. The meaning of the formative assessment was narrowed. The minor weightings allocated to process assessment made this response more like an attempt rather than a serious implementation.

With little training provided prior to the implementation of the policy change and little previous expertise in assessment, the appropriation at the teacher level relates to their tacit understanding and experiences. It is found that the teachers largely took the institutional interpretation of formative assessment as process assessment and enacted most of the procedures prescribed in the institutional assessment framework. However, flexibility in regard to specific assessment practice was also shown. Emergent themes illustrated use of criteria, which was not specified and unified at the policy level and not always communicated to the students at the
teacher level. There was evidence of involving students in assessment, mostly conducted with students of previous years or only when time permitted. There was evidence of constructive feedback for learning purposes, which focused on language and was taken up only by those students who had this need. The emphasis on quantity rather than quality when giving grades indicated process assessment not only fell short of the measurement purpose, but also detriment to students’ learning motivation. A need for professional training that clarifies the concepts and procedures of formative assessment thus surfaced. Control and distrust in teachers from the administrative sector was found to influence the teachers’ assessment practice negatively too.

Analysis of the student level responses firstly points to a strong testing orientation that dominated the learning environment of NKU. All the interviewed students would take the CET4. Not knowing the policy change was only part of the reason. Social factors such as employment need, expectations from the teacher and even themselves, and previously established habits also contribute to this result. The incorporated process assessment influenced only some but not others. Some students seemed to retain a double orientation of testing and learning, they also seemed to adopt different approaches to achieve respective purposes. The students’ responses appeared to relate to their educational background and prior experience in ELL and assessment. However, given the limited sample size in this case study, this claim is tentative and needs to be confirmed in large scale quantitative study.

Overall, these findings seem to point to a conclusion that the impact of these initiated policy changes on the NKU teachers and students is minimal. The policy intent to measure as well as improve learning via formative assessment was not fully achieved in this context. Once again the examination of the College English assessment policy at the NKU proved that policy implementation is by no means a linear process from the text to practice; rather, it is an appropriating and recreating process in “situated locales and community of practice” with unexpected outcomes (Levinson et al., 2009, p. 768).
Chapter 6: Cross-case Analysis

This chapter provides a cross-case analysis of the two individual cases of the Key University (KU) and the Non-key University (NKU) that focus on the unique characteristics of each in the implementation of the assessment policy change initiated in the College English Reform Program (CERP) (CMoE, 2004). The key themes that emerge from the analyses of the two cases are compared, synthesised and discussed. This comparative analysis focuses on identifying the tensions that exist within each particular case and across the two cases. To begin, the two institutions’ approaches to the policy change are compared (Section 6.1). The findings are discussed in relation to the first research question, which relates to the institutional interpretation of the policy change. Then, the similarities and differences in the teachers’ understandings of the assessment policy change (Section 6.2) and their assessment practices (Section 6.3) follow. Students’ perceptions (Section 6.4) and responses (Section 6.5) to the implemented approaches to assess their English Language Learning (ELL) are also compared across the two cases. Section 6.6 brings these four sections together and discusses the findings in relation to the second research question that relates to the translation of the policy intent into practice. The discussions also integrate relevant assessment and ELL literature. The final section (Section 6.7) summarises the findings and discussions and links to the next chapter.

6.1 INSTITUTIONAL APPROACHES TO THE POLICY CHANGE

The first question of this study examined how the College English assessment policy change was interpreted at the institutional level. The local interpretation of the national policy is significant in that the resulting institutional response prescribes the specific procedures for the initiatives within the local context. Therefore, the institutional interpretation is the pivotal link between the policy initiatives and classroom practices, and indicates the way in which the policy intent is translated into practice (Fullan, 2004). The question about policy initiatives at the institutional level in relation to each case has been addressed in Chapters 4 and 5. Now the two are brought together to present a broader picture of institutional approaches to the policy change.
6.1.1 Interpreting formative assessment

Through the guideline document, College English Curriculum Requirements (CECR) (CMoE, 2007), CERP advocated the incorporation of formative assessment within the original summative framework for the purpose of measuring and promoting learning. Comparison of the KU and NKU institutional assessment policies indicated that the meaning of the formative assessment initiative was interpreted similarly in the two local contexts. This can be seen in the shared intention, terminology, procedures and purposes. First, the incorporation of formative assessment for the purpose of improved ELL, which was stated in CECR (CMoE, 2007), was articulated by the two respective senior administrators in the interviews. That is, awareness of this intention was shared at the institutional level. Second, in both contexts, formative assessment was also referred to as process assessment. The use of the term *process assessment* represented its emphasis on timing and frequency. That is, this mode of assessment was conducted during the teaching and learning process rather than at the end of a certain period, and it was conducted more than once. Third, the procedures specified in the two institutional assessment policies were similar. These procedures refer to the internal assessment framework that included process assessment components of student participation in classroom activities, assignments and attendance (see Figures 4.2 and 5.1). These procedures incorporated alternative modes of assessment, which relied on teachers’ judgement and decision-making. Lastly, both institutional policies required records to be made of student performance in the three specified areas of (a) student participation in classroom activities, (b) assignment(s), and (c) attendance in order to generate a process grade. The process grade was combined with the final exam result and used for reporting. This meant in both cases the so-called process assessment was aggregated into a summative grade and had a summative purpose.

The two institutional policies also shared two other similarities. The first related to the link between formative assessment and feedback, which has been emphasised in much of the assessment for learning literature (ARG, 1999; 2002; Black, 1999; Black & Wiliam, 1998; Broadfoot & Black, 2004) and clearly stated in the CECR (CMoE 2007) (see Section 1.4.3). In these two universities, however, the link was implied rather than emphasised. The second concerns the central roles of students in assessment, an essential feature of formative assessment (Broadfoot,
CECR (CMoE, 2007) took up the feature via encouraging the use of self-assessment and peer-assessment, which, however, was absent in both of the two institutional policies. Specifically, involving students in assessment was not mentioned in the two senior administrator interviews.

To summarise, the meaning of formative assessment as interpreted in the institutional policies seemed to focus on the term *process* in terms of the timing and frequency of assessment. It involved a shift from the one-off summation of learning at the end to a collection of multiple sources of evidence during the teaching and learning process, and a shift from the sole reliance on objective standardised testing to the incorporation of teachers’ professional judgement. However, the importance of involving students as agents in the assessment process and the link between feedback and formative assessment were missing. In addition, the evidence collected from the process was graded and used primarily for summative purposes. Taken together, the two institutions seemed to regard the use of teachers’ professional judgement of students’ performances during the process as formative assessment. In other words, process assessment, though used interchangeably with the term formative assessment in these two Chinese institutions, does not convey the meaning of formative assessment as understood in other cultural contexts (Black & Wiliam, 1998; Broadfoot, 2007). It is a culturally situated adoption of the concept of formative assessment and will be elaborated upon in Chapter 7.

6.1.2 Weightings for process assessment

Despite the similar interpretations of process assessment as formative assessment in both universities, the two institutions differed significantly in the weightings given to this form of assessment. At KU, the weighting ratio between process assessment and the final term exam was 60:40, whereas at NKU the ratio was 10:90. The different weightings that the two institutions gave to process assessment were revealing in three aspects. First, the weightings suggested that each institution emphasised the change differently. The ten percent that was allocated to process assessment at NKU demonstrated much less emphasis on the change compared with 60% at KU. It should be noted, however, that NKU’s lower weightings did not necessarily mean a valuing of the change to a lesser degree. Rather, it was a matter of trust in teachers’ professional judgement. KU, by giving 60% to process assessment, entrusted teachers to take more control of their
classroom practice. In contrast, NKU, by allocating only 10% to process assessment and refusing to increase the percentage as requested by teachers, seemed to show a lesser degree of trust. This was shown in the excerpts below, when the NKU senior administrator commented on the reason for the minor weightings at NKU.

  Researcher: Is there any plan for further change recently?

  NKU Senior Administrator: Not as I see it. We have argued for more weighting for the process grade. Some teachers raised this question too. We did try, but the Teaching Administrative Office (TAO) rejected the proposal. The rejection was grounded in the finding that some teachers did not take the matter seriously, like giving all his/her students full scores… Given the irresponsibility of these teachers, we really cannot say anything when the TAO brought up this.

This means that TAO, the institutional policy-making body at NKU, on finding that some of the teachers’ grading practices were not accountable, chose to trust the objective final exam. This reflected an institutional culture that valued tests as a form of accountability.

  Second, the weightings suggested varying degrees of space allowed for the change at the institutions. Process assessment was acknowledged by both the senior administrator and the teachers at both cases as “[assessment] in teachers’ hands”. The weightings given to process assessment actually indicated the space that the institutions gave teachers to try assessment that was not a traditional test and use their professional judgement to assess students’ learning. It is justified to say that, compared with 60% given to process assessment at KU, the 10% given at NKU left the teachers with limited possibilities for assessing using means other than testing. This was recognised by NKU-T4:

  To be frank, I don’t think we teachers have much chance to assess at all. The lion’s share (90%) is not in our hands. Only the 10%, only orally in classroom, I … really don’t think we teachers can do much. Assessment or not, it is not up to us teachers.

  Taken together, the different weightings for process assessment at the two institutions demonstrated two different institutional cultures. One (KU) valued the change, trusted practitioners and gave them space to try assessment means other than
testing. On the other hand, the other university (NKU) still held on to testing and emphasised accountability in assessment.

6.1.3 Flexibility allowed to teachers

The degree of flexibility allowed to teachers is related to the acknowledgement of teachers’ voice in assessment, which, according to Hamp-Lyons (2007), is one of the characteristics that distinguish an examination-oriented assessment culture from a learning-oriented assessment culture. Flexibility of the institutional policy was demonstrated first of all in the weightings that the institutional policy gave to the process grade, as discussed above. Flexibility was also demonstrated in that both institutional policies prescribed the assessment procedures, but did not specify how they should be carried out but left it to the teachers. This was articulated by the KU senior administrator:

Yes, actually our teachers have a lot of flexibility except for the settled principles like the weightings for final term exam and the other parts. Say, for the share in teachers’ hands, they can do whatever they think the best...

While the flexibility as to how to carry out process assessment was given to the teachers in both cases, the flexibility that the two institutional policies allowed for the adoption and adaptation of the weightings differed. At KU, the 60:40 distribution between process assessment and the final term exam in the institutional policy was adopted and adapted differently by different teaching groups. For example, the Reading, Writing and Translating (RWT) teaching group adopted a ratio of 40:60, whereas the Speaking and Listening (SL) teaching group used 50:50. This flexibility was affirmed by KU-T5, a group leader, as she said: “Different teaching groups give different percentage to it. Sixty percent for our group (RWT)…” KU-T1, a SL teacher, adjusted the 50:50 ratio within the teaching group and used 40:60. He affirmed the adjustment he made when he said: “Yes, I make some adjustments and make one [syllabus] that both I and my students are comfortable with”.

Of the 60% given to process assessment, the percentage for attendance, participation and assignments was adjustable too. This was stated by the senior administrator:

The first 20% is attendance and participation. Another 20% is for assignments or quizzes. Some teachers combine the two together. Teachers
are quite flexible in the weightings for attendance, participation, and assignments.

In sharp contrast, the NKU 10:90 ratio between process assessment and final term exam was strictly regulated and implemented uniformly throughout the institution. There was only one teacher (NKU-T2) who reported allowing her students the flexibility to choose between classroom participation and assignments within the 10% allocated for process assessment. On the other hand, the six other NKU teachers abided by the institutional policy’s regulation regarding the ratio between process and the final term exam and that between assignment, participation and attendance.

The greater flexibility that the KU teachers were given suggested the institutional policy acted as a guideline only. The teachers, as a group or individually, had the freedom to adjust the suggested ratio as they thought best and conduct assessment in their preferred ways. In contrast, the minimal flexibility allowed to NKU teachers indicated that they had little chance for their voice to be heard in relation to assessment. Thus, the assessment culture at KU seemed to have more of a learning orientation in that attempts were being made to involve teachers more in the decision making process, whereas the NKU’s limited room for the teacher’s voice in assessment indicated adherence to a testing culture (Hamp-Lyons, 2007).

6.1.4 The positioning of CET4 at the institutional level

The external test, CET4, was made optional in both universities. However, comparative analysis showed that this initiative was approached differently in the two cases. KU made the CET4 optional and removed the high stakes attached to it several years before the initiative (CMoE, 2005) was launched. This pre-emptive action was prompted more by an internal need arising from the high English proficiency levels of the KU students, who were top students enrolled from all over the country rather than a mere response to the policy. Instead of using CET4, KU developed a proficiency test for gate-keeping purposes, whereby students had to pass this test successfully in order to obtain their degree. It could be said that CET4 had lost its significant position as an external test at the institutional level.

In contrast, NKU made CET4 optional after the policy release (CMoE, 2005) and thus this was a direct response to the national initiative. CET4 was no longer a qualification requirement for the bachelor degree, yet its link to the awarding of
scholarships was retained. In this way, CET4 remained a high-stakes test at NKU. This was mostly due to the institutional need to use the test for motivational purposes. NKU’s aim was to motivate students to learn by using scholarship as an incentive. The institutionally-based College English achievement test (band four), which is currently used as a baseline qualification test was generally regarded as much easier than CET4. In this sense, CET4 was still regarded as significant at NKU. This positioning of CET4 again reflected an institutional assessment culture, which valued objective testing and hence was more examination oriented.

The discussion in this section so far has examined institutional responses to the assessment policy change in the two universities. The findings link to the first research question for this study: How is the College English assessment policy change interpreted at an institutional level?

In summary, formative assessment was adapted in both cases as process assessment. Process assessment had incorporated the national policy’s directive to encompass both the process and the product of learning in assessment (CMoE, 2007) but implied meanings that differ from the concept of formative assessment as defined by theorists from other contexts (ARG, 1999, 2002; Black & Wiliam, 1998; Broadfoot, 2007).

While appearing to share an understanding of the formative assessment initiative, the two institutions differed significantly in the weightings given to process assessment, the flexibility given to teachers, and the approaches to the external test. These differences suggested different institutional cultures (support or control) and assessment cultures (learning-oriented or examination-oriented) within the two institutions. Overall, KU appeared to offer a policy context that was supportive of teacher initiatives and adaptations within the process component of the assessment framework. It was a policy context more favourable for the policy changes to be translated into classroom practices.

In the following four sections, I turn to a cross-case analysis of the teachers’ and students’ approaches to the assessment policy change. I first examine teachers’ understandings of the assessment policy change and then their practices in the classroom. I follow this with an examination of the students’ understandings and responses to the implemented assessment practices. These four sections correspond
to the four sub-questions designed to address the second research question that relates to the extent that the policy intent was reflected in practice.

6.2 TEACHERS’ UNDERSTANDINGS OF THE ASSESSMENT POLICY CHANGE

Comparison of the teachers’ understandings across the two cases revealed that the two teacher groups shared an awareness of the policy initiatives. Both groups also indicated a lack of training and professional knowledge about formative assessment and instead relied on a tacit understanding of what it was and how it could be realised in the classroom. However, significant differences appeared in the two teacher groups’ views of the roles of CET4. These similarities and differences are now discussed.

6.2.1 Awareness of the formative assessment initiative

Mixed findings were revealed in both cases regarding the teachers’ awareness of the changes inherent in the formative assessment initiative. The majority of teachers reported knowledge of the incorporation of process assessment, although, as noted earlier, they used the term process assessment interchangeably with formative assessment without seeming to understand the distinctions between the two terms. However, one teacher from each university was not aware of the changes, as evidenced below.

KU-T6: [The previous and the current institutional policies are] roughly the same. I do not see much change.

NKU-T7: To me, the institutional one [policy] has not made any changes over all these years.

Quite possibly, these two teachers, both having extensive experience in teaching College English (25 years for KU-T6; 23 years for NKU-T7), had developed fatigue with policy changes, as has been found in research on teachers’ responses to educational change (Muller, Norrie, Hernandez, & Goodson, 2009; Walsh, 2006). Or, it might reflect their impression that this policy change was insignificant. More evidence is needed to confirm exactly what led to this lack of awareness of the change. However, their lack of awareness raises the question: If the change in assessment policy was not recognised by teachers, how could it possibly be translated into practice?
6.2.2 Awareness of the changed status of CET4

The teachers were aware of the optional status of CET4. A difference existed in the communication of this change to teachers in the respective institutions. NKU teachers did appear to know about the change to the status of the external test. However, they unanimously stated that the institutional policy-making body had not communicated this change to them. NKU-T7, for example, reported she accessed information about this change from the internet. NKU-T5 got to know about this change via informal discussion with colleagues. In contrast, the KU teachers knew about the optional status of CET4 and did not seem to see the need to engage with it. This is understandable given that the change to the status of CET4 at KU was made over a decade ago and the CET4 pass rate at KU was nearly 100%.

Despite the fact that the teachers from both contexts shared the knowledge of the changed status of CET4, the varied ways in which they were notified about the changes seemed to lead to the differing outcomes. Specifically, the institutional approach to disseminating information about the initiative at NKU led to haphazard knowledge of the reform and a level of disinterest and/or a lack of understanding about the significance of the changes and their justification in terms of student ELL. This approach seemed to impact the NKU teachers’ uptake of the policy change.

6.2.3 Training and professional knowledge

Of the thirteen teachers involved in this study, only KU-T1 had received specialised assessment training, but this had not incorporated formative assessment. The NKU teachers, however, did mention some training programs in the interviews. An example was the training that CMoE provided for teachers who were involved in the piloting of CECR, which NKU-T2 alone attended. However, NKU-T2 was reluctant to recognise it as kind of ‘assessment training’ since, according to her, assessment was only a minor area covered. The NKU focus group mentioned a recent teleconference program, in which teachers from other institutions communicated their experience of College English assessment. However, only some NKU teachers were able to participate because of the cost associated (the teachers were asked to pay 100RMB to participate). Overall, the teachers of both institutions seemed to receive little or no systematic training in assessment. This supports the findings of Cheng et al. (2004) that only a minority of tertiary English as a Foreign
Language (EFL) teachers in China have received any specialised training in the area of assessment.

In response to this lack of training in assessment, the teachers in both institutions admitted a lack of “theoretical” understanding of formative assessment. They reported an understanding of the concept as informed by “tacit knowledge” (KU-T3, KU-S4) or “experience” (NKU-T4, NKU-T5, NKU-T7). It is important to note that these teachers, all local Chinese, had spent most of their educational experience, both as students and teachers, in the Chinese educational system which was dominated by an examination culture (Cheng & Curtis, 2009a; Qi, 2007). The assessment they experienced as EFL learners and teachers was mostly limited to testing. The lack of training and experience in areas other than testing was illustrated by KU-T5.

[We have] no perception of [assessment at] theoretical level. Assessment is really not our speciality. We come in touch with it mostly when it comes to examinations. There is training, twice a year, on the tests. That is the training relevant to student assessment, as we can see it.

The importance of training and sustained technical support cannot be overemphasised in innovation (ARG, 2008; Fullan, 2004), otherwise, teachers tend to fall back on their experience (Snow, 1989). In this study, teachers’ prior experiences meant that they would possibly fall back to their experience of test-based assessment. The fact that teachers did not have professional knowledge of formative assessment and that they were not provided with effective training surfaced as the major factor that most affected their realisation of the policy changes in their practice. This will be discussed further in Chapter 7.

6.2.4 Understanding formative assessment

With little training and professional knowledge of assessment in areas other than testing, the teachers in both universities reported an intuitive approach to understanding formative assessment. Comparative analysis showed that the teachers’ understanding of formative assessment emphasised process as opposed to one-off tests. This emphasis was illustrated by KU-T1: “It [formative assessment] sounds to me like assessing the students’ learning by looking at the learning process rather than the results from a one-time test”. The teachers seemed to be aware of the need to focus more on individual students (KU-T2, NKU-T2, NKU-T1, NKU-T4) and what
they regarded as a subjective (KU-T2, KU-T4, NKU-T2, NKU-T6) nature of formative assessment which involved teachers’ judgement. Nonetheless, they all perceived the need for an objective result for assessment – summative and formative alike (KU-T2, KU-T3, NKU-T2, NKU-T5). The emphasis on objectivity and results was expressed in the excerpt from a focus group with NKU teachers below:

Researcher: Then what do you think formative assessment should be like?

NKU-T6: I wonder whether the grades from the process like quizzes should be integrated. That way, the assessment could be more objective.

NKU-T5: Sure.

NKU-T7: Yes, I think scores from the quizzes, mid-term exam and also the oral test should all be included to make it more objective.

NKU-T6: Yeah, all those that make assessment objective.

The emphasis on process and objective results as described in the excerpt indicates the teachers’ understanding of formative assessment, which could be more accurately described as progressive mini-tests. Such an understanding reflects cultural and institutional influences as discussed in Section 6.2.1. It also marks a departure from the existing literature on formative assessment – its purposes for ELL and its classroom realisation. The question, which I address in the following section, is how this culturally situated understanding was translated into classroom assessment practices in the two universities.

6.2.5 The role of CET4

The roles that CET4, the external test, played in the two contexts were regarded differently by the KU and NKU teachers. Given the high pass rate of the KU students in the test, the CET4 was seen as having little relevance by KU teachers in terms of measuring and motivating students’ learning. Moreover, because the test had been optional for a decade at KU, the changed status of the test exerted minimal impact on teaching practice, which was largely learning-oriented.

The NKU teachers provided a contrasting position. They agreed that the choice of taking the external test should be up to the students. However, they also subscribed to the view that the test was needed for quality assurance and motivation and that the students needed good test results for practical considerations such as employment after graduation. It seems the teachers become very committed to existing practices,
particularly those that have lasted for a long time. They accrue value and have an investment in them. Acknowledging the official status of CET4, they recognised the test was still an important benchmark in their teaching practices. The significant role of CET4 was echoed in the three teacher interviews (two individual and one focus group) and illustrated by NKU-T3 below.

> In my opinion the CET4 is just like a yardstick, which directs both the College English teaching and learning practice. Despite CET4 being disconnected from the qualifications, the mainstream practice remains the same.

The teachers’ understanding of the role of CET4 in their teaching appeared to align with the institutional positioning of the test in the policy and relate to their understanding of the needs of the students. The data shows the teachers in the two universities exhibiting totally different views on the role of CET4 which followed through into their orientations to the test in their teaching. Specifically, in contrast to the learning orientation that the KU teachers stated as guiding their practice, the NKU teachers indicated a more obvious testing orientation, with CET4 continuing to be a presence in their teaching. This difference in practice is to be discussed in detail in Section 6.3.3.

### 6.3 TEACHERS’ ASSESSMENT PRACTICES

The KU and NKU teachers are compared in this section in terms of their practical classroom approaches to the assessment policy changes. Their classroom practice of process assessment in alignment with the institutional policies is analysed first. This is followed by an examination of their practices in terms of criteria use, student involvement and feedback. The two teacher groups’ practical engagement with CET4, the external test, is also compared and contrasted.

#### 6.3.1 Alignment of process assessment practice to the institutional policy

Comparative analysis of the two teacher groups’ assessment practice did not show significant differences in terms of alignment with the institutional policy. Both KU and NKU teachers reported assessment procedures that reflected their respective institutional policies. Specifically, they assigned assessment tasks or quizzes to students, although the task types and frequencies varied between teachers. They referred to the students’ participation in classroom activities, but they adopted
different means to engage students. While some teachers asked students to give an oral presentation (KU-T1, KU-T2, KU-T3, NKU-T1, KU-T7), other teachers looked at students’ participation in terms of their responses to teachers’ questions (KU-T4, KU-T6, NKU-T2, NKU-T3). The teachers checked students’ attendance; only the means of checking differed from one to another. Some checked by occasionally calling the roll (KU-T2, KU-T3, NKU-T1, NKU-T3, NKU-T4); some checked via the weekly assignments (KU-T1, KU-T4); and some monitored attendance by asking students to answer questions (KU-T6, NKU-T2). Moreover, the teachers, in accordance with institutional policies, made records of the students’ performances in these three areas and generated a process grade for summative use (see Appendices G and H for example). There is evidence of some teachers’ altering the institutional ratios between process assessment and the final exam, as was discussed above in Section 6.2.3. The majority of teachers in both universities, however, adopted the institutionally suggested or prescribed ratio or the ratio suggested within their teaching groups. In general, the teachers’ assessment practices adhered to what was stipulated in the institutional assessment policy, which was principally the implementation of process assessment seen as a series of progressive mini-tasks that were used for accountability and summative purposes. The comments of NKU-T2 below are illustrative:

Part [of my assessment practice] is conducted as required by the school authority, as demonstrated in the record sheet I showed you just now [attached in Appendix J]. It requires 3% classroom participation, 3% assignments and 4% attendance for the process grades, which altogether accounts for 10% of the overall assessment [for reporting purpose]. Therefore I have to come up with a record and hand it in by the end of the term.

Overall, the teachers in both contexts, by aligning their assessment practices to the institutional policies, adopted the institutional interpretation of formative assessment as process assessment (see Section 6.2.1). This approach, evident in comments from teachers in both universities, appeared to agree with the teachers’ understanding of formative assessment (Section 6.2.4).
6.3.2 Assessment for formative purposes

Informed by the principles of formative assessment as set out in the literature (ARG, 1999, 2002), I investigated the teachers’ assessment practices in three areas: use of criteria, student involvement in assessment, and feedback practices, which were respectively presented in Chapters 4 and 5. The analysis of the data reveals that the teachers in both cases approached these areas in their individually preferred ways. Differences between the two teacher groups’ practices in these areas were not obvious, as I explain below.

Criteria use

At both universities, a majority of the teachers reported that they used criteria. But the criteria were self-developed for some (KU-T1, KU-T2, NKU-T4), borrowed from the CET4 syllabus for others (NKU-T6), or based on tacit understanding for some (KU-T4, NKU-T6). Of those who used criteria, some teachers stated that they shared the criteria with students (KU-T1, KU-T2, NKU-T2, NKU-T6) whereas others utilised the criteria implicitly and assumed that students knew these criteria (KU-T4, KU-T6, NKU-T4). The variety of the teachers’ practices in this area pointed to the finding that there was no uniform use of criteria. This was the situation both in and across the two universities.

The lack of specific and uniform criteria raised the issue of the objectivity of the teachers’ judgement, which brought about different consequences in the two contexts. At NKU, the TAO’s questioning of the teaching grading practices was mentioned by the two individual teachers and led to a heated debate in the focus group interview. The NKU teachers found themselves experiencing a dilemma by being required to report an objective grade based on their professional judgement of the students’ performances in process assessment. This situation was made more difficult as no specified criteria were provided. This is illustrated in the statement below.

NKU-T5: What is lacking is a set of practical implementation procedure. If the specifications of [criteria and procedures] are prescribed in detail, well, it is easy to carry it out. But if not, it is hard to say. Four points for classroom participation – then what a 4 points should be like and what a 3 points is like. It is sort of abstract rather than concrete.
In contrast, at KU the questioning of the teachers’ professional judgement of students’ participation from the administration did not happen. The KU teachers were similarly concerned about monitoring by the TAO, as KU-T6 observed: “the institution will check on the details of assessment. Not a bit should be missing.” Nonetheless, the KU teachers showed a high level of confidence in their professional judgement and tacit understanding of the standards as they applied at KU. The confidence, demonstrated in the statement below, indicated a tension of a much lesser degree at KU.

KU-T4: You know it [if the assignment is good or not] only by tacit knowledge and tacit knowledge is more exact mostly in the case of say speaking and writing… It is more reliable than the test results from a standardised test.

The tension over objectivity, on the one hand, reflects the technical issues that needed to be addressed at the institutional policy level when using teachers’ judgement for summative purposes, such as reliability, validity, and the necessity of training (Harlen, 2005a, 2005b) and practicality (James, 2006; Taras, 2005). On the other hand, it mirrors a test-based perception of assessment which emphasised objectivity and measurement. Also, the different degrees of the tension in the two contexts again pointed to different institutional cultures: whether the institution entrusted the teachers and gave them the confidence to try new means of assessment or not.

**Student involvement**

Student involvement in assessment practices was reported in both cases. There were occasions that students were involved in negotiating assessment arrangements (KU-T1) and assessment results (NKU-T4). There were also reports of involving students in assessing peers’ assignments such as oral presentation in English (KU-T1, KU-T2, NKU-T4, NKU-T7) and writing an English passage (KU-T1, NKU-T2, NKU-T4). However, it should be noted that these instances of student involvement were self-initiated by the individual teachers and based on their understanding of the role of students in assessment. For instance, KU-T1 used a writing blog and NKU-T7 invited students to assess their own and others’ oral presentations. These examples were based on these teachers’ understanding of what was good for students’ ELL. These practices were also in line with their understanding that students should and
could take an active role in assessment. In contrast, there were other teachers (KU-T3 and NKU-T1) who did not include students in assessment activities such as self- or peer-assessment. This did not appear to be a valued activity by these teachers.

The ways these teachers involved students differed and appeared to be related to their experience and professional development. KU-T1 and NKU-T2, for instance, both related their practices to their work (KU-T1) or study (KU-T1, NKU-T2) experience abroad. NKU-T7 attributed her practice to ideas gained from a journal article. In contrast, the majority of teachers who did not report that they had such professional experiences did not actively engage students in the assessment of their own or others’ work. NKU-T4, who had not been abroad, was the exception. He reported that his student-oriented assessment practices originated from his reflections on his own teaching experience. To illustrate: “my [assessment] skills are mainly from experience”.

The extent to which the teachers had experienced professional or academic development either afforded or constrained their assessment practice and influenced the degree to which they involved students in peer- or self-assessment. The majority of teachers lacked professional expertise in formative assessment. Professional development or study of formative assessment seemed to provide the necessary knowledge and skills for only four teachers to engage students in assessment. The involvement of students by these teachers aligns with the view in the literature that learners’ involvement in assessment can enable them to take control of their learning and develop metacognition for learning in the future (Broadfoot, 2007). The focus from this view shifts to an emphasis on learning and away from a teacher-oriented testing culture.

It should also be noted that the use of assessment activities experienced when teachers were abroad did not necessarily mean a change to the teachers’ perception of their role in assessment. KU-T1, for example, had introduced peer review to a writing blog program that he learned when he was abroad. Nonetheless, he still upheld the teacher’s role in assessment as dominant. For example:

[Teachers play] a leading role [in assessment], I suppose. He/she designs the assessment tasks and does the assessment him/herself. It is especially so in the Chinese classroom. At least I am a leader and do all the assessment myself, mainly according to their results.
The examples of student involvement in assessment practice that have been noted indicate a change for these four teachers, in that they are sharing with students the responsibility for assessment. For these teachers in particular this development suggests the beginnings of a dialogical rather than a monological teacher/student relationship. The latter is characteristic of the Chinese teacher/student relationship (Ho et al., 2001; Hu, 2002).

**Feedback**

Comparative analysis of the varied feedback practices of the teacher groups at the two universities showed some similarities and differences. To begin, there were teachers in both cases who appeared to share a preference for positive feedback (KU-T1, KU-T2, KU-T5, KU-T6, NKU-T2, NKU-T6). They related it to the purpose of motivation. This preference seemed stronger at KU given that more KU teachers identified their preference for using positive feedback in the interviews (four out of six KU teachers as compared to two out of seven NKU teachers). This could also be due to the KU students’ higher levels of English compared with the NKU students.

Second, the feedback tended to be brief and evaluative for most teachers across the two cases. For example, the feedback was given once only rather than progressively. This meant the feedback given was a summation of the assessment process and hence tended to be evaluative. Whether the feedback was acted upon, to what extent this occurred and whether the feedback was used to inform learning was left up to the students themselves. While a general tendency to evaluate how well the tasks were accomplished was obvious in most teachers’ feedback practice, there were teachers from each case who reported focusing on learning as well and gave detailed and constructive feedback to students about how to improve (KU-T1, KU-T2, KU-T4, NKU-T2, NKU-T4). They differed in that the NKU teachers reported less frequency of such practice and this was reported as related to the teachers’ workload.

NKU-T4: That’s why I can’t possibly do this [giving detailed feedback using criteria] many times. I did it only twice last semester and once this semester.

But even so, it can give them a kind of feeling about where they are and what they need to do for the next step. It is really exhausting.

The heavy workload associated with giving feedback as mentioned by NKU-T4 and the perception that this was a constraint was stated by seven NKU teachers and only one KU teacher (KU-T1). The greater number of classes and the larger class
sizes at NKU were possibly contributing factors to this response. As shown in Tables 3.4 and 3.6, there were 40 students in KU classes, but in NKU classes the number of students ranged from 45 to 64. The KU teachers taught only Reading, Writing and Translation (RWT) or Speaking and Listening (SL). Two of them taught one class only with the maximum number of classes taught by KU teachers was three. In contrast, the NKU teachers taught both RWT and SL lessons. The minimum number of classes taught by NKU teachers was two and could be as many as four. Workload then was a problem for teachers at NKU and impacted on their feedback practices, whereas this problem was not apparent at KU.

Another constraint that was shared by the teachers of both cases related to the time limits. According to the teachers the weekly session for RWT and SL, which was 90 minutes in all, impacted on their feedback and assessment practice in general. This constraint was illustrative in the following two quotes from the teacher interview data:

- KU-T1: We have only a session per week, little time is left to check.
- NKU-T4: … Too much content needs to be covered in class; time is just not enough. I can’t possibly spend too much time in assessing.

The teachers tended to see teaching and assessment as separate activities rather than view assessment as integral to the teaching learning cycle. Assessment for formative purposes tended to be marginalised when the teachers needed to meet the teaching agenda.

To summarise, although the teachers from both universities aligned their assessment practices with their respective institutional policies, they did not necessarily assess only for summative purposes. The analysis of assessment practices showed that teachers in both contexts were engaged with principles associated with formative assessment, that is, assessment and evaluation that prioritises learning. The data showed that although the criteria used to assess student work were not uniform and explicit within or across the cases, some teachers were utilising criteria. Also, despite the authoritative position of the teacher in regards to assessment, students were involved on some assessment occasions and were given an active role in some teachers’ assessment processes. Although feedback was mostly evaluative and brief, some teachers used feedback for learning purposes. Assessment practice that could
be considered formative existed in some classrooms of both cases although these practices were not formally recognised by teachers as such. They were more the product of the teachers’ pedagogy, overseas experiences and own academic endeavours and were directed at enabling the students’ English Language Learning.

Also, the two teacher groups’ largely similar responses to the formative assessment initiative reveal a finding worth further consideration. At NKU, the minor weightings for process assessment and distrust from the administration were identified by the teachers as major problems constraining their assessment practices. At KU, where the two problems did not occur, the teachers’ assessment practices did not differ significantly. This means that something more than larger weightings and trust is needed for a substantial change in the teachers’ assessment practice.

6.3.3 Engagement with CET4

The cross-analysis of the KU and NKU teachers’ classroom engagement with CET4 revealed major differences. The KU teachers did not prioritise the students’ performance on the tests including CET4. They shared an interest in fulfilling the assigned teaching tasks and developing students’ English proficiency albeit with little or no focus on CET4. At NKU, the teachers were likewise interested in enhancing student learning; but they also saw CET4 as important for their students. Despite the abandonment of test preparation practice in response to the changed policy, many NKU teachers continued to balance their interest in learning with their views about students’ need for the test. This is illustrated by NKU-T2 and NKU-T1 in the snapshots presented in Chapter 5. NKU-T5 saw the test as so important for her students that she maintained testing preparation and attempted to help “as many [students] as possible” to pass.

While the KU teachers’ assessment practice in terms of the external test largely aligned with the institutional policy initiative, some NKU teachers interpreted the policy such that they continued with test preparation even though CET4 was made optional. This was because they acknowledged the NKU students’ continuing interest and need for the test results. In other words the teachers’ interests and beliefs about what was important seemed to be the major mediating factors within their teaching practice. The different interests and the corresponding behaviours of the two teacher groups confirmed the different assessment cultures (Hamp-Lyons, 2007) of these two contexts. That is, the assessment culture at KU seemed to have more of a learning
orientation while that at NKU was more examination-oriented. Put differently, the teachers’ engagement with CET4 reflected the ongoing impact of the testing culture (Cheng et al., 2007), which appeared stronger at NKU than that at KU. The attributing factors to this result are further elaborated upon in Chapter 7.

6.4 STUDENTS’ UNDERSTANDINGS OF THE ASSESSMENT CHANGE

I start this section by examining the KU and NKU students’ awareness of the two policy initiatives: the optional status of the external test and the incorporation of formative assessment that was also referred to as process assessment. The students from the two cases shared an awareness of the incorporated process assessment but differed significantly in their knowledge of the changed status of CET4. This discussion is followed by a comparison of the two student groups’ understandings of process assessment and their views of the role of CET4. CET4 appeared to mean something different for the KU and NKU students; however they both understood process assessment similarly.

6.4.1 Awareness of the incorporation of process assessment

The KU and NKU students demonstrated knowledge about the incorporation of formative assessment as process assessment, and understood how this occurred. They were aware that process assessment comprised assignments, classroom participation and attendance. They were also aware that the result of process assessment was used summatively. One minor difference related to the students’ understanding of the allocated weightings. The majority of KU students seemed to have a general understanding of the weightings of process assessment whereas the NKU students knew the exact weighting for each component of process assessment. The NKU students showed a greater concern for the process grades and appeared to value the grades per se.

6.4.2 Awareness of the changed status of CET4

A major difference emerged regarding the KU and NKU students’ awareness of the changed status of the external test. Specifically, the five students interviewed at KU knew that CET4 was no longer compulsory and did not see it as high-stakes. The NKU students seemed unaware of the optional status of the test and its disconnection from the qualification requirements. It seemed that the NKU students were not well informed of the institutional policy. They reported that they relied on
senior students for information about qualification requirements. This pointed to the possibility that the way that policy was disseminated at NKU had resulted in a different interpretation about its significance by NKU students as opposed to the KU students’ interpretation.

6.4.3 Understanding of process assessment

Comparative analyses showed that the two student groups shared a common understanding of process assessment. KU and NKU students saw process assessment as different from their previous assessment experience which was primarily dominated by exams. Their understanding seemed to relate to the specific tasks, as illustrated by KU-S2 who used the term formative assessment when interviewed: “I regard formative assessment as assignment tasks I need to complete”. They appeared to understand the intent of the change to distribute the students’ efforts more equitably across the process. To illustrate, KU-S4 commented: “For good or bad, some percentage is given to the process; we devote some effort to the process rather than keeping it for the final exam”. The purpose of using process assessment to improve learning also seemed to be understood by most students across the two cases. The students also saw the summative use of the assessment results, unanimously referred to as process grades, as a concern. NKU-S1’s response illustrates this for, when asked about her understanding of how English was assessed, she replied:

NKU-S1: Just the process grades during the semester plus the grades from the final exam at the end of the semester. We usually gain grades during the semester from writing assignments and participation in class as our teacher judges it. At the end of semester, we take a final term exam...

Knowing that a process grade was what was demanded of them, process assessment to them seemed to be perceived as no more than a different means of assessment of learning. In this sense, the students from both contexts took up process assessment as implemented by the teachers and as interpreted by the institutional policy.

Students from both contexts were also aware of the teachers’ flexible and varied practices regarding process assessment and what they perceived to be its subjective nature. Some students were in fact doubtful about the reliability of the teachers’ judgement. KU students, for example, seemed to show a higher level of
concern over “objectivity” and “fairness” of the teachers’ professional judgement than the NKU students. KU-S5 stated:

KU-S5: Most students in some classes got more than ninety. Only one or two students in our class got that high. Many thought it unfair.

Researcher: How come?

KU-S5: It seems different teachers used different standards…

This doubt regarding standard maintenance was not apparent among NKU students. This could be due to the weighting of 10% for process assessment. NKU-S1 stated: “After all, there are only 10 points altogether for the process, no more”. In this way, it seems the different weightings for process assessment at the two universities influenced the students’ responses to the change. While the smaller weighting at NKU did not arouse major concern for the teachers’ judgement, the heavier weightings at KU did.

The KU students’ concerns for consistency of standards and greater reliability of teachers’ professional judgements, the teachers’ needs for objectivity and the NKU administration’s questioning of the objectivity and reliability of the teachers’ professional judgements, reflected the influence of a product-focused and testing-oriented assessment culture. This perception seems to go against the learning-oriented culture that is needed for formative assessment (Hamp-Lyons, 2007; Sherpard, 2001). It posed a major barrier to the implementation of the formative assessment initiative and is discussed in Chapter 7.

6.4.4 The roles of CET4

KU and NKU students offered different views regarding the role of CET4 in ELL. NKU students saw the test as significant to them. Compulsory or not, the CET4 for most of them was still a yardstick that the teachers referred to in and out of class (NKU-S3). It was a means to prove their English proficiency level to themselves and others (NKU-S1, NKU-S2, NKU-S4); and an important aspect of cultural capital (Cheng, 2008) that they needed to invest for present scholarship or for employment in the future (NKU-S2, NKU-S3, NKU-S5). The significance of CET4 was stated by NKU-S1:

Anyway, CET4 certificate is kind of recognition of your English level. And the recognition is authoritative and from the society. Whatever you said
about your English proficiency means nothing when you try to find a job. What they want is a proof of your English level. CET4/6 is a good one. Otherwise, you never know [what the consequences will be like]…

It seems the NKU students’ high regard for CET4 was related, but not limited, to the academic stakes that were attached to it by the institutions. The students’ understanding of the high-stakes consequences to them in a social context such as employment opportunities seemed to carry more weight. Four of the five NKU students (NKU-S1, NKU-S2, NKU-S4, NKU-S5) clearly expressed that the optional status of CET4 did not affect their views of its significance and their subsequent decisions to take the test.

The KU students did not seem to view CET4 as significant. They knew the optional nature of the test and that the stakes associated with qualifications were attached to a different test – the English proficiency test that KU had developed. In addition, they also knew that the test posed no challenge for them given their high English levels. Moreover, they indicated that they would take tests of higher level such as CET6, Graduate Record Examination (GRE) or Test of English as a Foreign Language (TOEFL), which could provide proof of their English proficiency when they needed it for employment purposes. Taken together, CET4 represented much less significance for the KU students. It was no more than a goal that they used to regulate their study and facilitate the long term goal they set for themselves – to gain high scores in TOEFL or GRE so as to study overseas (KU-S2, KU-S4) or gain higher proficiency for use in a future career (KU-S3, KU-S5). Making CET4 optional was not very significant for them.

The KU and NKU students’ different views of the significance of CET4 seemed to be consistent with the different positioning of the test at the institutional level and the teachers’ different approaches to it. The students’ English level was an influential factor.

6.5 STUDENTS’ RESPONSES TO THE ASSESSMENT CHANGE

The KU and NKU students’ responses to the assessment changes are now compared and contrasted. Their approaches to process assessment as implemented are discussed first. This is followed by an examination of their responses to teachers’
practices in terms of criteria use, student involvement and feedback. Their approaches to CET4, the external test, are also included.

6.5.1 Experience of process assessment

The KU and NKU students’ responses to process assessment indicated a concern over scores and a desire for high grades. With this motivation it is not surprising to find that the students from both contexts took strategic approaches to cope with process assessment procedures that were implemented to assess their College English Language Learning. The detailed representation of the students’ pragmatic approach to assessment of the KU case (Chapter 4) and the strategic approach to assessment of the NKU students (Chapter 5) reveal some similarities. Specifically, the students made efforts to gain the highest possible scores by: seeking cues from teachers and students of peer classes or senior years (KU-S2, KU-S4, NKU-S1, NKU-S3); trying to present an assignment task as best as they could (KU-S2, KU-S5, NKU-S1, NKU-S4) or responding purposefully to teachers’ questions (NKU-S3) in the classroom. They were playing the “examination game” (Miller & Parlett, 1974) and undertook process assessment from an instrumental perspective (Swann & Arthurs, 1998), that is, pragmatic and grade-focused.

Nonetheless, this finding is far from conclusive. There were students in both cases (KU-S2, KU-S3, KU-S5, NKU-S1, NKU-S2) who reported that coping with the implemented assessment for the highest possible scores was only a minor part of their English Language Learning experience. They did study strategically for high scores, but they maintained a concurrent approach to learning for improvement. KU-S4 and KU-S5 were also aware of this learning-oriented approach, but did not allow time for this emphasis on improvement for learning.

The five students who reported this learning-oriented approach (KU-S2, KU-S3, KU-S5, NKU-S1, NKU-S2) were from the developed regions. Although NKU-S1 and NKU-S2 were enrolled at NKU they had been recruited from developed regions. It appears that the learning orientation was greater for these students of higher English proficiency levels and who were from developed regions than for those students from the under-developed regions and who were not as proficient in ELL. Success, higher ELL proficiency, in this case, generated motivation. An implication is that the achievement of a certain level of proficiency in ELL as evident from testing and high scores was the first priority and that learning improvement
only became a priority once a certain level of proficiency in ELL had been reached. In other words, testing was more likely to shape the learning behaviours of those whose English level was less proficient. The students were more likely to orient toward learning if they did not feel threatened by assessment. This finding, together with the anxiety that most of the students from under-developed regions showed and the confidence that the students from the developed regions demonstrated in the interviews, confirm the disparity that Hu (2003) found between students from economically more developed regions compared with those from less developed areas. There appears to be a pattern of difference between the students’ ways of experiencing assessment. This is only a tentative assumption based on this small scale study and needs to be further explored on a large scale quantitative investigation.

6.5.2 Formative assessment

The two student groups’ responses to the teachers’ assessment practices of criteria use, student involvement and feedback are now compared. The extent to which formative assessment brought about positive effects on the students’ ELL is the focus of this comparison.

Criteria Use

In contrast to the teachers’ reports about criteria use and sharing, the students reported a lack of knowledge of the criteria that were used to assess their assignments. This was stated by nine of the ten students from both universities. The only one who gave a positive response (NKU-S1) mistook the process grade as criteria. It was concluded that the students were not aware of the criteria used by teachers. When asked about the PowerPoint that KU-T1 shared with the students, which KU-T1 regarded as criteria sharing, KU-S3 replied:

[Through the PowerPoint] the teacher (KU-T1) just gave you a clear goal for you to work towards. But actually that is not a very clear concept. For example, you need to figure out how to achieve it yourself. [The criteria] are not clear to me.

Similarly the NKU students stated that they had “no idea” of how they got their scores. The students seemed to perceive criteria differently from their teachers. KU-T1’s attempt at sharing the goal was not understood by the students as criteria. The
examples and the criteria as described by NKU-T2 were also not seen as criteria for assessment by the NKU students. It could be said that criteria, even if used, were not made explicit enough for the students. A precondition for effective formative assessment practice as identified by Black and Wiliam (1998) and Orsmond et al. (2000) is the sharing of assessment criteria and the standards by teachers. This precondition was not met in both contexts.

**Student involvement**

In line with NKU-T2’s report that she did not involve the current class (the one I observed and chose student interviewees from) in assessment practices, the five NKU students did not report any experience of being involved in assessment activities. But there was evidence of the students’ use of self-reflection in monitoring their ELL (NKU-S2) and reviewing classmates’ writing on informal occasions (NKU-S4). With little experience of being involved in assessment, the NKU students overall showed a reserved attitude to express any opinions about student involvement in assessment. This was represented by NKU-S3, who said: “I need to try myself before I can tell how”.

In contrast, the five KU students reported experience of being involved in assessment in KU-T1’s class or in others. Nonetheless, their views and responses to being involved in assessment differed. KU-S3, in line with her positive view of the role of assessment in her ELL, highly regarded her peer-review experience in KU-T1’s writing blog program. KU-S2 and KU-S5 spoke positively of the peer-review they experienced previously in high school, but they did not seem to think their involvement in KU-T1’s writing blog really helped to improve their ELL greatly. In addition, the KU students seemed to share a sceptical attitude to the formal use of peer-assessment and self-assessment. Drawing upon their personal experience the KU students raised doubts regarding the reliability and accountability of these forms of assessment. They suggested it was “tokenism” (KU-S4) or “an administrative thing” (KU-S2). KU-S1 expressed concerns about the possibility of exposing her weaknesses to peers, which, she worried, would affect her grades. These doubts and concerns related to the summative use of the results and reflected the students’ perception of assessment, which was framed as objective and suitable for summative use. The sceptical attitude and the perception of assessment for summative use were barriers to the implementation of assessment for formative purposes.
This sceptical attitude may reflect the possibility that the self- and peer-assessment activities in which these students were involved were not conducted effectively. This suggests a need for teachers to participate in professional development to develop skills and understandings about formative assessment. Another implication is that if the teacher does not fully understand the assessment procedures this could impact on students’ attitudes towards the changes which could prove detrimental to future policy efforts.

**Feedback**

The KU and NKU students’ responses to questions about feedback confirmed the variety of teachers’ practices, as revealed in the teachers’ section (6.3.2). While NKU-T2 provided feedback that distinguished the qualities of good assignments from others, KU-T1 provided feedback that was always complimentary and undifferentiated. Nonetheless, the feedback from both teachers was brief and lacked constructive advice about how to improve.

Although the positive comments did boost some students’ confidence (KU-S4, NKU-S4), the students from both cases shared a general disappointment over the lack of constructive advice. Consequently, most of the students tended to ignore the feedback. Two students (NKU-S5, KU-S2) were the exceptions. NKU-S5 reported that she corrected the grammatical errors that the teachers had marked in her writing assignments. KU-S2 once acted upon the comments the teacher made on her oral presentation and sent a revised PowerPoint to the teacher for fear her grades might be affected. Overall, voices for more critical, individualised and constructive feedback were loud and clear (KU-S1, KU-S3, NKU-S1, NKU-S2, NKU-S4). To illustrate:

> I hope he [KU-T1] can point out my weaknesses. I find that [KU-T1] always said “very good” to everyone and then pointed out a small mistake. We hope he can point out more. For Chinese students who had experienced much frustration, we needn’t so much praise. We hope for some beneficial advice.

The above comparison of the students’ responses to the teachers’ assessment practices in the areas of criteria use, student involvement and feedback revealed largely negative findings. The students reported that the criteria used by the teachers for formative purposes were not always clear nor easily understood. Some students viewed their involvement in assessment activities as positive while the majority did not view these activities as helpful to their ELL. Overall, students reported that the
teachers’ feedback was brief, complimentary and lacked constructive advice so was seldom acted upon. Taken together, the teachers’ assessment practices for apparent formative purposes (Section 6.3.2) did not seem to function to improve the students’ ELL as intended. Teachers require professional development when policy change such as this is implemented and this and other implications will be elaborated upon in next chapter.

Also, the students in both cases recognised that the incorporation of process assessment did impact their time and effort distribution in a way, which, in most students’ view (KU-S3, KU-S4, KU-S5, NKU-S1, NKU-S2, NKU-S3, NKU-S4, NKU-S5), should lead to positive results in learning. However, the possible change to learning approaches that some NKU students (NKU-S2, NKU-S5) expected in case of larger weightings to process assessment did not occur in the KU students, who did have large weightings. The KU students refuted the possibility of changing their learning approaches owing to process assessment. To illustrate:

KU-S4: As to the effect of changes in assessment methods [incorporation of process assessment] on us, I don’t think the changes made so far are enough [to result in major changes in our ELL approaches and results].

Researcher: Do you mean more weighting to the process?

KU-S4: Not only giving more weighting to the process, I think more can be done. Only writing a formatted essay does not mean much.

In other words, the weighting, big or small, was not pivotal for changes to happen in students’ learning. Similar to the teachers, something more substantial than larger weighting is needed for the intended change to happen.

6.5.3 Responses to CET4

Ten students from KU and NKU reported sitting for, or registering for, the external test. It appears that whether the students were informed of the policy change or not (see Section 6.4.2), and whether they viewed it significant or not (see Section 6.4.4), did not bring about a difference in their responsive behaviours to the test.

Three reasons appeared to contribute to the students’ decision to take the test. For most of the NKU students (NKU-S1, NKU-S2, NKU-S3, NKU-S4) and KU-S4, disconnecting CET4 from the qualification requirements within the institutional context did not remove the high-stakes that were attached to the test in the social
context. These students took the test because they understood that their results would have consequences for them when it came to employment opportunities. NKU-S5 made this point manifest by saying: “Although it has been separated from the degree, it is not when you look for your job”.

This social reason however was not applicable to NKU-S4, who bluntly stated “I don’t think it [CET4 certificate] would be of any use for getting a job”. She related her test-taking practice to the expectations that she had for herself, which related more to cultural factors. NKU-S4’s reaction was echoed by KU-S1 and KU-S3. These three students related their test-taking practice to a habit that developed from conditioning within the historical culture of examinations of the Chinese education system. The influence of this historical cultural context is illustrated in the following:

KU-S1: Other purposes? No. You know, we study and take tests… kind of habit…

KU-S3: From very young on, we have had countless tests and we have been accustomed to that…

It seems these students who had been immersed in an examination culture, had also been positioned into a testing orientation. Even when testing was not a requirement of the school, the students expected it of themselves – to take the test and pass. It could be argued this cultural context influenced the students’ responsive behaviours.

Similarly, social consequences could not fully explain NKU-S2’s behaviour, who explained that he took the test not only because the certificate was needed for higher degree study and employment but because he was interested in improving his English. This learning orientation resonated with KU-S2, KU-S3 and KU-S5, who articulated their desire to take the test so as to keep some time on ELL and for improved learning purposes.

To conclude, whether the students took CET4 or not seemed to have little relevance to the optional status of the test within the institutional context. The social consequences and cultural context positioned them such that they felt that they had to do it. There were two differences identified. First, the social consequences appeared stronger in the NKU context (with four students indicating this reason), whereas only one KU student (KU-S4) mentioned it. Second, the learning orientation seemed more
influential for the KU site, with three students stating that they use the test formatively as compared with one student at NKU. These differences appeared consistent with, and relevant to, the roles of CET4 perceived by the students (see Section 6.4.4), and the learning or testing orientation of the teachers’ practice (see Section 6.3.3) and the institutional culture (see Section 6.1.4). The small sample size of this study means that this finding must be confirmed by a large-scale study before any significant claims can be made.

6.6 DISCUSSION

The comparative analysis of the teachers’ and students’ approaches to the assessment change across the two cases in the four sections above elicited the following findings.

First, given the different positioning of CET4 at the institutional level, the KU and NKU teachers held different views regarding the role of test. They adopted different approaches to the external test in classrooms. It seems that KU teachers highlighted the learning orientation in their teaching of College English while the NKU teachers tried to balance learning and testing in their teaching practices.

Second, there was a lack of relevant professional skills and experience in assessment for the teachers. Consequently the teachers from both contexts adopted the institutional interpretation of formative assessment as process assessment. That meant giving weight to the process but retaining the emphasis on the summative function. Despite the apparent summative approach to process assessment there was some evidence of an emerging use of formative assessment. For example, some teachers from both contexts reported making use of criteria, giving feedback for improved learning purposes, and involving students in the assessment of their own and their peers’ English Language Learning.

Third, the students from both universities continued to sit for CET4 regardless of their different awareness about the optional status of CET4 (KU students knew about the change; NKU students did not know) and their different views regarding the roles of the test (KU students did not think the test very significant but NKU students thought it still very significant to them). The social consequences of taking the test and the historical cultural contexts seemed to impact on the implementation of the policy initiative. Nonetheless, some students used the test formatively for
learning improvement and were not focused on it as only a summative evaluation of their proficiency.

Fourth, students from both universities adopted a strategic approach to process assessment to achieve high scores. However, some students incorporated a learning orientation in their ELL. Also, the students’ responses to the teachers’ use of criteria indicated a disjuncture in relation to what was reported by the teachers. This made apparent the lack of transparency and explicitness in terms of assessment criteria used by the teachers. The students’ responses to the teachers’ use of feedback, on the one hand, revealed a score orientation and on the other hand, called for more constructive feedback. Students’ attitudes and responses to being involved in assessment raised caution against the practice of self and peer-assessment without sufficient understanding and professional development on the part of teachers.

These findings are now discussed in relation to the second research question: To what extent does the assessment practice that is implemented reflect the policy intent? The change to the assessment policy involved making the external test optional and incorporating assessment during the process of College English teaching. The aim and purpose of this initiative as espoused by the CMoE was to shift the focus from testing to learning and to achieve improved ELL. This study found that students still focus on grades yet an orientation towards learning is apparent in some teaching and learning practices. There seem to be elements of the policy intent being reflected in the teaching and assessment practices. For instance, some teachers began to use assessment tasks such as oral presentations and engaged students in peer review, in addition to the continued use of testing. The teachers’ professional judgement in the assessment of these tasks was combined with the results of the objective tests. There is also some evidence that individual teachers practised assessment for formative purposes. Some students also spent more time and effort on ELL during the term and some began to participate in classroom activities more actively as a result of the inclusion of process assessment. The students’ participation in a way indicates a changing role. However, no fundamental change to the teachers’ assessment practice seemed to have been implemented as a direct consequence of the assessment policy change. Teachers such as KU-T1 and NKU-T7 tended to attribute their learning orientation in teaching to their changed pedagogical beliefs arising from professional development rather than the policy initiative. The
individual-based formative assessment practices were referred to as something that they “have been doing all the way” (NKU-T4, KU-T4). This was the intent of the policy but was not attributed to it. According to Pryor and Crossouard (2008), “the development of formative assessment entails finding new tools and changed classroom practices” (p. 2). The largely unchanged assessment practices thus meant that the policy intent, if reflected in the practice, did it only to a limited extent.

### 6.7 SUMMARY

This chapter compared and contrasted how the two assessment policy changes initiated in CERP (CMoE, 2004) had been interpreted and implemented in two universities. Comparative analysis of the approaches adopted at the institutional, the teacher and the student levels showed similarities as well as differences, as now illustrated in Tables 6.1 and 6.2.

Table 6.1
External Test (CET4) Initiative Being Approached at the Two Local Contexts

<table>
<thead>
<tr>
<th></th>
<th>KU</th>
<th>NKU</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Institutional</strong></td>
<td>CET4 made optional</td>
<td>CET4 made optional</td>
</tr>
<tr>
<td></td>
<td>No longer high stakes</td>
<td>High stakes partly disconnected</td>
</tr>
<tr>
<td></td>
<td>Positioned as not significant</td>
<td>Still significantly positioned</td>
</tr>
<tr>
<td><strong>Teacher</strong></td>
<td>Aware of the optional status of CET4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CET4 no longer the sole measurement</td>
<td>CET4 should be optional but still needed</td>
</tr>
<tr>
<td></td>
<td>Little perceived impact on teaching</td>
<td>Still a benchmark and incorporated in</td>
</tr>
<tr>
<td></td>
<td>practice</td>
<td>teaching practice</td>
</tr>
<tr>
<td><strong>Student</strong></td>
<td>Aware of the optional status of CET4</td>
<td>Unaware of the optional status of CET4</td>
</tr>
<tr>
<td></td>
<td>A mini goal to facilitate long term goal</td>
<td>Significant capital for scholarship and</td>
</tr>
<tr>
<td></td>
<td>More students tended to use the test</td>
<td>employment</td>
</tr>
<tr>
<td></td>
<td>formatively</td>
<td>More students tended to use to the test</td>
</tr>
<tr>
<td></td>
<td></td>
<td>summatively</td>
</tr>
<tr>
<td></td>
<td>All sat for or will sit for the test</td>
<td></td>
</tr>
</tbody>
</table>

228 Chapter 6: Cross-case analysis
**Table 6.2**  
Approach to Formative Assessment in the Two Local Contexts

<table>
<thead>
<tr>
<th></th>
<th>KU</th>
<th>NKU</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Institutional</strong></td>
<td>Formative assessment interpreted as process assessment</td>
<td>Formative assessment interpreted as process assessment</td>
</tr>
<tr>
<td></td>
<td>Similar assessment procedures</td>
<td>Similar assessment procedures</td>
</tr>
<tr>
<td></td>
<td>Similar summative use of the assessment results</td>
<td>Similar summative use of the assessment results</td>
</tr>
<tr>
<td><strong>Teacher</strong></td>
<td>60% to process assessment</td>
<td>10% to process assessment</td>
</tr>
<tr>
<td></td>
<td>Flexibility allowed regarding changing ratios</td>
<td>Little flexibility allowed regarding ratios</td>
</tr>
<tr>
<td><strong>Student</strong></td>
<td>Most teachers aware of the change; one teacher in each case not</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lack of training and professional knowledge in assessment and formative assessment in particular</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Take up the interpreted meaning of formative assessment as process assessment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Process assessment procedures aligned to the institutional policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessment of formative nature visible in some teachers’ practice</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aware of the incorporation of process assessment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Process assessment understood as summative process grade</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Strategic approach to process assessment tasks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recognised no criteria; disappointed in feedback; reserved or sceptical attitude to student involvement in assessment</td>
<td></td>
</tr>
</tbody>
</table>

While the differences listed in the two tables indicated different orientations in terms of assessment culture and even institutional culture at KU and NKU, the similar responsive behaviours to the formative assessment initiative deserves further explanation. This will be the focus of the next chapter, in which these similarities and differences will be further analysed, discussed and explained from a sociocultural perspective.
Chapter 7: Discussion

The comparative analysis in Chapter 6 addressed the research questions from a cross-case perspective. It concluded that the policy intent of CERP assessment policy change (CMoE, 2004) was reflected in assessment practices to a limited extent. Chapter 7 explains, from a sociocultural perspective, what contributed to this result. A community perspective of learning (Lave & Wenger, 1991; Wenger, 1998, 2000) is adopted to explain the outcome. In line with the research design, this discussion foregrounds the implementation of the formative assessment initiative, whilst backgrounding the external test initiative. The discussion aims to identify the factors that afford and constrain formative assessment practice in the Chinese higher education context. The insights are elicited from the perspective of English Language Learning (ELL).

The discussion starts with identifying the communities of practice to which the two cases belong (Section 7.1). The findings at the institutional (Section 7.2), teacher (Section 7.3) and student (Section 7.4) levels (as revealed in the previous chapters) are then discussed. A concluding section (Section 7.5) summarises the discussion and links it to the conclusion and implications for policy and practice in the final chapter.

7.1 COMMUNITIES OF PRACTICE

As reviewed in Chapter 2, community of practice is where learning and identity transformations happen through participation and interaction with social partners (Lave & Wenger, 1991; Wenger, 1998, 2000). However, a community of practice is much more than this. According to Wenger (1998), a community of practice is

…the locus of engagement in action, interpersonal relations, shared knowledge and negotiation of enterprise and it holds the key to real transformation – the kind that has real effects on people’s lives. (p. 85)

A community contains the “competences of a learning system that shares cultural practices and reflects a collective learning” and holds “the very core of” (Wenger, 2000, p. 229) meaning making within its given context. Learning at the community level, also referred to as apprenticeship metaphorically (Rogoff, 1995,
p. 143), focuses on the mutual embedded relationship of the individuals and the communities. The community level of analysis is used in sociocultural studies to examine “the institutional structure and cultural technologies of intellectual activities” (Rogoff, 1995, p. 143), such as the negotiation of the meaning of formative assessment, as in this study. Based on this theoretical basis, the two case studies are located within the communities to which they belong. The findings and issues are discussed from this dimension of community.

Wenger (1998) defines a community as constituted by three dimensions of practice. These are mutual engagement, joint enterprise and shared repertoire (p. 73), as represented in Figure 7.1.

![Figure 7.1. Dimensions of practice as the property of a community (Wenger, 1998, p. 73).](image)

A community of practice is, therefore, defined by the shared practices and the relations of mutual engagement (Wenger, 1998). Membership of a community means access to the resources, competence and opportunity to participate in a shared endeavour (Lave & Wenger, 1991; Wenger, 1998). One belongs to a community because of what one does. The two universities involved in this study differ in terms of location and key or non-key status (see Sections 4.1.1 and 5.1.1), but share engagement with College English education. This shared engagement entitled the two universities to membership within the College English community of practice, which has over 2000 other higher educational institutions. Also included in this community are the administrators, the College English teachers, and the undergraduate students who major in fields other than English. However, it needs to be noted that the College English community of practice is not self contained. It is situated within a larger Chinese cultural community. Moreover, it is involved in an
international disciplinary community of English Language Learning (ELL). The multi-membership of the two cases and the interrelationship of these communities are illustrated in Figure 7.2.

Figure 7.2. KU and NKU as situated in communities of practice.

The three communities of practice provide shared repertoire, resources and competence that mediate the local actors’ negotiation of policy meaning. To begin, ELL as a discipline had a tradition of psychometric testing (Kunnan, 2005), that was characterised as large-scale, standardised and objective (Davidson, 2004; Davison & Cummins, 2007). These qualities were encompassed in the College English community of practice. It had a history of utilising CET4/6, a standardised test, as an external test, and a testing culture which highlighted objective measurement (Cheng, 2008; Jin & Yang, 2006; Yang, 2003). The testing culture was complicated by the high-stakes use of the test results in the institutional and societal contexts (Niu, 2001; Shu, 2004), which led to consequences, as discussed in Chapter 1 (Cai, 2004, 2005; Han, 2002; Jia & Yang, 2005; Tang, 2005b; J. Wang, 2007; Zhang, 2005). The traditional values embedded in the Chinese culture of learning and assessment functioned as cultural tools for thinking within the Chinese cultural community. Examples include the utilitarian view of learning, the hierarchical teacher/student relationship (Biggs, 1996b; Ho, 2001a; Hu, 2002), and the emphasis on the summative function of assessment (Han & Yang, 2001; Tang, 2006; Tian, 2004). These traditions and cultural tools formed the resources and repertoire that are shared by the members and frame the members’ ways of perceiving and engaging. With legitimate access to these resources, the two universities participated in the joint enterprise of negotiating the meaning of the formative assessment policy initiatives.
The findings elicited from the analysis of the interplay between assessment principles, the CECR assessment policy, and the assessment practices within the two localised contexts are now discussed in relation to the mediation of the three communities of practice.

7.2 NEGOTIATED MEANING OF PROCESS ASSESSMENT

The comparative analysis of the two institutional assessment policies revealed that both KU and NKU implemented formative assessment as process assessment. Process assessment, although used interchangeably with the term formative assessment by the teachers and administrators in the both contexts, did not convey the full meaning of formative assessment as defined in the CECR national policy (CMoE, 2007) and the international literature (Black et al., 2003; Black & Wiliam, 1998).

The two terms (process assessment and formative assessment) have one similarity, which is assessment should be conducted during the process of teaching and learning, rather than only conducted at the end of the study period. But the two terms diverged in terms of: purposes served, agents involved, and the link with feedback. Specifically, process assessment was used within the two universities for the summative purposes of reporting (and also selection for scholarship at NKU). It was reliant mainly on the teachers and/or tests for assessment. The link with feedback was not emphasised. That is, the departure of the meaning of process assessment from formative assessment as defined in CECR (CMoE, 2007) and the international literature (Black et al., 2003; Black & Wiliam, 1998) was apparent.

However, there were considerable differences between the two universities. In terms of status in the Chinese higher education circle, one was a prestigious key university in Beijing and recruited the highest achieving teachers and students in the area of ELL from around the country (see Section 4.1.1). The other was a non-key university in a provincial area of China and had teachers and students of average achievement level in ELL (see Section 5.1.1). The two universities also differed in terms of organisational culture. KU was more flexible in the implementation of the assessment policy and entrusted the teachers to try new means of assessment. In contrast, NKU was more restrictive and gave teachers limited voice in assessment. Moreover, the assessment culture, as reflected in the institutional and teachers’
approaches to the external test, differed. The assessment culture at KU tended to orient more towards learning; whereas a testing orientation was more obvious at NKU. These differences are illustrated in Table 7.1.

Table 7.1 KU and NKU Comparison

<table>
<thead>
<tr>
<th>Factor</th>
<th>KU</th>
<th>NKU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Key university, prestigious</td>
<td>Non-key university, average</td>
</tr>
<tr>
<td>Location</td>
<td>Developed area</td>
<td>Under-developed area</td>
</tr>
<tr>
<td>Education input</td>
<td>Abundantly resourced</td>
<td>Resourced</td>
</tr>
<tr>
<td>Teachers</td>
<td>Highly qualified</td>
<td>Qualified</td>
</tr>
<tr>
<td>Students</td>
<td>Top-level in EEHE</td>
<td>Average-level in EEHE</td>
</tr>
<tr>
<td>Organisation culture</td>
<td>Trust and empowerment</td>
<td>Distrust and control</td>
</tr>
<tr>
<td>Assessment culture</td>
<td>Learning-oriented</td>
<td>Learning and testing oriented</td>
</tr>
</tbody>
</table>

The question that arises is: what contributed to this similarity in interpretation in the two institutional policies despite the multiple differences in the two universities? A sociocultural perspective posits to explicate human actions within situated “cultural, institutional, and historical” contexts (Wertsch et al., 1995, p. 3). Process assessment, when put into its communities of practice, revealed a marked influence from the Chinese cultural community of practice. The summative use of the assessment results, for instance, paralleled the Chinese assessment tradition that values the product of learning more than the process of learning (Han & Yang, 2001). Further, the marginalised role of students in assessment mirrors the hierarchical teacher-student relationship that is embedded in the Chinese culture of learning (Biggs, 1996; Ho et al., 2001a; Hu, 2002). It seems the values and the perceptions that were embodied in the shared cultural community functioned as cultural tools for thinking and mediated the meaning negotiation process at the institutional level (Levinson & Sutton, 2001).

The mediating role of a situated cultural community of practice was highlighted by Levinson et al. (2009):

When non-authorised policy actors – typically teachers and students, but possibly, too building administrators [whom, Levinson et al. (2009) referred to as “cultural animals”] – appropriate policy, they are in effect making new policy in situated locales and community of practice. (p. 768)
Wenger (1998) too stated: “It is only as negotiated by the community that conditions, resources and demands shape the practice” (p. 80). According to Wenger (1998), the mediating role of a community of practice is no less significant when it comes to an external mandate such as policy:

The enterprise is never fully determined by an outside mandate, by a prescription or by an individual participant. Even when a community of practice arises in response to that mandate, the practice evolves into the community’s own response to that mandate. (p. 80)

In this sense, the two universities’ individual responses to the formative assessment initiative are indeed a communal practice. It is this communal response that negotiated the meaning of formative assessment as process assessment. This communal perspective explains why the two universities similarly implemented formative assessment as process assessment despite the differences in institutional situations.

This cultural influence was verified by the NKU senior administrator, who noted: “This [doing assessment this way] is a cultural thing, more or less.” Informal personal communications with colleagues in other universities and published articles (Guo & Li, 2007; Tang, 2007) also confirmed that this approach was adopted by other universities as well. In this sense, process assessment is a culturally situated adoption and adaptation of formative assessment as defined in international literature.

It is well recognised (Ball, 2006; Fullan, 2004) that policy is by no means only text; rather it is also discourse that embodies power per se. A sociocultural approach regards educational policy as a social practice of power (Levinson & Sutton, 2001; Levinson, Sutton, & Winstead, 2009). The policy makes governing statements about what can and should be done; on the other hand, it is shaped and recreated by local actors and dynamics (Levinson et al., 2009; Spillane, Reiser, & Reimer, 2002). That is the case with process assessment. KU and NKU, using the agency allowed by the national policy, recreated formative assessment in their situated cultural context.

It should be noted that the departure of meaning as intended in policy from what was implemented were reported in other contexts too. In the UK, formative assessment, even with training provided, is implemented by many teachers to the letter without the spirit (Marshall & Drummond, 2006) or even distorted (Klenowski, 2009a).
Process assessment, the results of which were prescribed for summative use also involved some feedback that was given to guide future learning. This use of process assessment corresponds with the fourth and last possibility in the typology of Davison (2008). Process assessment as identified in this study focused mainly on the product of learning, but was also used formatively to guide future teaching and learning. The teachers used the results for reporting, which was sometimes accompanied with teachers’ feedback and students’ self-evaluations. This departs from the formative emphasis of assessment for learning purposes. Process assessment, nonetheless, is a stepping away from the summative end in the local context and a move towards the formative end along the assessment continuum. In this sense, process assessment also seems to resemble the informal summative assessment that Harlen (2006) identified in the assessment continuum (see Figure 7.3). It is a shift from a long established summative testing tradition and a move towards the purpose of improved learning.

Carless (2010) suggested a “contextually grounded approach” for the formative assessment policy initiatives in a Confucian Heritage Culture (CHC) to cope with the contextual reality. Process assessment then is a culturally situated interpretation of formative assessment within its particular local contexts. It embodies cultural values, and reflects contextual reality and history.

7.3 FACTORS MEDIATING TEACHERS’ NEGOTIATION OF MEANING

The comparative analysis in Chapter 6 revealed that teachers from both cases reported minimal changes to their assessment practices. No significant differences emerged from the two teacher groups’ assessment practices regardless of the disparities within the two cases. However, the teachers’ assessment practices varied on an individual basis, with some incorporating assessment for learning purposes
while others did not. The focus of the discussion, then, is on the factors that are
decisive in teachers’ assessment practices and the factors that enabled these
individual teachers to conduct assessment formatively. Contextual factors that
influenced the teachers’ assessment practices were identified in Chapters 4, 5, and 6.
A comparison of these factors in terms of the two teacher groups is illustrated in
Table 7.2.

Table 7.2 Contextual Factors Influencing Teachers’ Assessment Practices (differences)

<table>
<thead>
<tr>
<th>Factor</th>
<th>KU-Teachers</th>
<th>NKU-Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy delivery</td>
<td>Informed orally and in written documents</td>
<td>Informed only orally</td>
</tr>
<tr>
<td>Flexibility</td>
<td>Greater in terms of weighting (60%) for process assessment</td>
<td>Less in terms of weighting for process assessment (10%)</td>
</tr>
<tr>
<td>Trust</td>
<td>High level of trust from administration</td>
<td>Low level of trust from administration</td>
</tr>
<tr>
<td>Assessment culture</td>
<td>More learning-oriented</td>
<td>More testing-oriented</td>
</tr>
<tr>
<td>Teaching workload</td>
<td>Average</td>
<td>Heavy</td>
</tr>
</tbody>
</table>

The comparison shows the contextual differences experienced by the teachers
in the two cases. At NKU, the teachers identified five factors that hindered their
uptake of formative assessment. First, the institutional policy was disseminated
informally and did not seem to arouse teachers’ attention. Second, the NKU teachers
were given limited weightings for process assessment and little flexibility in the
ratios, which suggested they were controlled rather than empowered. Third, the
teachers’ professional judgement did not appear to be trusted. Fourth, the assessment
culture at NKU was more testing oriented. Moreover, the NKU teachers were
burdened with a heavier teaching load.

These obstacles correspond with the findings in assessment literature, as
detailed below. To begin, the ways that policy is disseminated are regarded as having
the potential to influence the practitioners’ interest and attention to the change (ARG,
2009b; Fullan, 2004). Second, power relations within the workplace such as the
degree of trust and empowerment are viewed as major reasons for resistance (Carless,
2009). Teachers’ positive attitudes to changes are perceived as vital to the success of
an innovation (Carless, 2009; Fullan, 2004; Rogers, Cheng & Hu, 2007; Wang &
Cheng, 2005). Empowering teachers is a recognised precondition for teachers to
conduct formative assessment (ARG, 2009b; Swann & Arthurs, 1998). Also, a
testing-oriented assessment culture is identified as the key nexus that militated
against formative assessment classroom practice in Hong Kong and other contexts (Carless, 2008; Hamp-Lyons, 2007; Kepple & Carless, 2006; Shepard, 2000). Workload too is reported in many studies as a negative element to teachers’ classroom assessment practices (Guo & Li, 2004; Kepple & Carless, 2006; Morris, 2002; Tang, 2006).

When these factors are taken into account, the minimal impact of the formative assessment initiative on the NKU teachers’ assessment practice seems understandable. However, at KU, where informal dissemination, negative power relations, testing-orientation, and heavy workload were not an issue, the teachers’ assessment practices regarding formative assessment did not show salient differences from the NKU teacher group. Therefore, it can be said these issues, though important, were not decisive factors for substantial change in the Chinese ELL teachers’ assessment practice. Other factors such as teacher qualifications and student EL levels did not seem to be decisive factors either. More decisive influences on the teachers’ assessment practice came from contextual factors that the two cases shared.

The comparative analysis revealed five similarities that the two teacher groups shared, as indicated in Table 7.3.

<table>
<thead>
<tr>
<th>Table 7.3 Contextual factors that influence teachers’ assessment practice (similarities)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process assessment policy</td>
</tr>
<tr>
<td>Lack of training in assessment, particularly formative assessment</td>
</tr>
<tr>
<td>Prior experience as EFL learners and teachers in the Chinese context</td>
</tr>
<tr>
<td>An assessment culture that demands objectivity</td>
</tr>
<tr>
<td>Hierarchical teacher/student relationship</td>
</tr>
</tbody>
</table>

The similarities relate to (1) the interpretation of formative assessment as process assessment in the institutional policies; (2) the teachers’ lack of professional training in formative assessment; (3) teachers’ shared experience as EFL learners and teachers in the Chinese context; (4) a demand that assessment must be objective; and (5) the hierarchical teacher/student relationship. These five factors reside within the communities of practice that the two universities shared: the College English, Chinese cultural, and ELL disciplinary communities. They seem to be more influential on teachers’ assessment practices and are discussed from a community perspective (Wenger, 1998, 2000) in this section.
7.3.1 Process assessment policy

The teachers of both universities implemented the assessment procedures and summative use of the results as prescribed in the institutional policies. That is, process assessment was practised by the teachers of both cases because of the influence of the institutional policies. This aligns with Murphy’s (2008) argument that curriculum and policy were a major shaping force of teachers’ classroom assessment practices.

A sociocultural approach views policy as embodying the power to shape and to be shaped (Levinson et al., 2009). However, in the present study, it appears that the shaping force of the institutional policies regarding process assessment is intensified by three factors. First, this policy change is situated in China, a sociocultural context that values authority (Biggs, 1996; Zhu, 1992). Moreover, the assessment policy change followed a top-down approach. In such a society, it is taken for granted that a policy from “above” is to be complied with. The third, and most important, factor is that the institutional policies prescribed specific procedures. When the teachers implemented the prescribed procedures such as quizzes and assignments, classroom participation and attendance, they actually took up process assessment. Based on these findings, it could be argued that the two institutions’ interpretation of formative assessment as process assessment is a major reason for the minimal difference in teachers’ assessment practices across the two cases.

7.3.2 Professional training in assessment

Lack of professional training in assessment was reported by all but one teacher interviewed in the two universities. Moreover, lack of expertise in formative assessment was reported by all teachers. These findings align with those of Cheng et al. (2004) and Wang (2003) that most EFL teachers in China did not have training in assessment. The influence of this lack of professional knowledge on assessment practice was articulated by the teachers. Confusion as to how to conduct formative assessment was articulated by NKU teachers, who, accordingly, asked for specific procedures and training.

NKU-T5: Give the practitioners something specific to refer to. Something practical and procedural, and if possible, training.
NKU-T7: Yes, practical procedures plus training. If you want to make the assessment accountable, these two conditions need to be satisfied.

The teachers’ confusion reflected the technical challenges and dilemmas that teachers faced within the changing assessment climate (Arkoudis & O’Loughlin, 2004; McNamara, 2001; Rea-Dickins, 2004).

Sociocultural theory, while highlighting the learner’s active role in meaning construction, also emphasises the significance of scaffolders, who facilitate meaning construction within the Zone of Proximal Development (ZPD) (Vygotsky, 1978). The concept of formative assessment was new to the teachers. Their lack of professional training meant that they did not have experts to scaffold their assessment practice when they tried to make sense of the meaning of formative assessment.

Sociocultural theory acknowledges that, on many occasions, the “master-apprentice” relations (Lave & Wenger, 1991, p. 93) or the “expert-novice dyads” (Rogoff, 1995, p. 143) are not available within a community of practice. Under such circumstances, sociocultural theory emphasises the significance of an informal newcomer and old-timer relationship. That is, it acknowledges that learners, on many occasions, configure their learning with other apprentices, who are more advanced in learning (Lave & Wenger, 1991; Rogoff, 1995). The reality for the teachers in this study was, however, that none of them had any theoretical training in formative assessment. Moreover, their experience in assessment as ELL learners and teachers was mostly immersed in a testing culture and examination tradition. That meant the possibility to learn from peers was absent for them as well. From a sociocultural perspective, it can be said that lack of professional training was another major factor that constrained the teachers’ assessment practices. It is, therefore, another major factor for the limited impact of the policy change on teachers’ assessment practice.

The teachers’ assessment practices tended to develop in three directions. First, most teachers complied with the institutional policy and what they understood to be process assessment. That meant they aligned their assessment practice with process assessment by taking up the assessment procedures and the summative function of assessment as prescribed in the institutional policies. Therefore, it could be said that lack of professional training did, indeed, contribute to the power of the institutional policies, as elaborated on in Section 7.3.1.
Second, some individual teachers engaged themselves in practice and developed an understanding of formative assessment through imagination (Wenger, 2000). NKU-T4, for example, did not have any previous experience of formative assessment. However, through reflective engagement, he developed an understanding about the necessity of sharing criteria, involving students, and providing constructive feedback in formative assessment. Yet, in the present study NKU-T4 was the only teacher who developed his understanding this way.

Third, teachers developed knowledge of formative assessment when they went beyond the boundary of their community and engaged with what Wenger (2000) termed “boundary interactions”. The boundary interaction, which was defined as “an experience of being exposed to a foreign competence” by Wenger (2000, p. 233), provided explanations for some teachers’ formative assessment practice. KU-T1, for example, attributed his knowledge of assessment to his experiences abroad. This included negotiating assessment arrangements with students, sharing criteria, and using peer-review in the writing blog. NKU-T7, on the other hand, learned to use self- and peer-assessment from a journal article, which was also a source beyond the community she belonged to. The journal article is one of the “boundary objects” that Wenger (1998) identified as connections across different communities. The three different directions that the teachers developed their understanding regarding formative assessment provide explanations for the variations in the individual teachers’ assessment practices.

Those who developed competence from the cross-boundary experience, such as KU-T1, are referred to by Wenger (1998, 2000) as “broker[s]”, who can introduce elements of practice from one community to another. These teachers seemed to be still in the process of learning as peripheral participants. When they developed full understanding, and became full participants, they could be “masters” within the community of practice. Hence, they may function as the “exemplars and motivation” for the new-comers and the old-timers who had not developed to the same level (Lave & Wenger, 1991). This will be an important implication of the findings and will be elaborated upon in Chapter 8.

7.3.3 Teacher’ prior experience

The teachers in both cases attributed their assessment skills and knowledge to experience. This agrees with Snow’s (1989) finding that teachers tend to fall back on
their experience when they are not provided with effective and sustained support. This also confirmed the findings of research conducted in the UK (Murphy, 2008) and China (Xu & Liu, 2009) that revealed teachers’ prior assessment experience affected their current practices and even future plans for assessment.

This finding aligns with the sociocultural perspective that values the mediation of history in human actions:

As trajectory, our engagement in learning and identities [teachers’ assessment practices in response to the formative assessment initiative in this case] incorporate the past and the future in the very process of negotiating the present. (Wenger, 2000, p. 241)

Although the KU teachers and two NKU teachers (NKU-T2, NKU-T7) had overseas study experience, most of their experience as ELL learners and teachers was in the Chinese context. The legacy from the Chinese cultural and ELL disciplinary communities of practice such as emphasis on the summative function of assessment, the hierarchical teacher/student relationship, and the testing culture unavoidably functioned as mediation tools (Rogoff, 2003). The teachers’ prior experience as ELL learners and teachers, thus, is another major barrier to their uptake of formative assessment in understanding and practice.

7.3.4 A demand for objectivity in assessment practices

The teachers from both cases faced a demand for objectivity in assessment from the stakeholders. This demand, firstly, seemed to relate to the teachers’ understanding of the purpose of process assessment as prescribed in the institutional policies, which was summative and needed to be accountable for reporting purposes. This demand was made manifest at both universities by the teachers’ reports of detailed requisite records of students’ performances. As KU-T6 illustrated, “The school will check on the details of assessment. Not a bit should be missing”. This concern was intensified at NKU by the Teaching Administrative Office’s (TAO) questioning of the objectivity of the teachers’ grading practices. This finding confirms the close relationship between teachers’ assessment practices and their understanding of the assessment purposes established in literature on English language teachers (Cheng et al., 2008; Cummings, 2001; Shulman, 1986).
The demand for objectivity also came from the students across the two cases. Two KU students (KU-S4, KU-S5), for example, raised doubts about the fairness of the teachers’ professional judgements of their performances. The NKU students echoed this demand and asked teachers to pay attention not only to the quantity but also to the quality of their assignments.

Apart from the administration and students, the concern for objectivity was also articulated by teachers themselves: not only those from NKU where the testing culture was predominant, but also teachers from KU where the assessment culture was more learning-oriented. To illustrate, KU-T5 stated: “the objective part of assessment is definitely needed”.

It can be said that the demand for objectivity prevailed with the stakeholders at various levels which included administrators, teachers and students. This reflects an assessment culture that emphasised reliability. In other words, the teachers faced this demand on a day-to-day basis and interacted with social partners, who held the same perception, in the process of negotiation of the meaning. Negotiation of the meaning of formative assessment as process assessment in the teachers’ assessment practices is thus a joint enterprise.

The demand for objectivity and reliability demonstrates influences from the situated communities of practice. The demand for objectivity at the institutional level, for example, mirrors the traditional summative use of assessment (Han & Zhang, 2001) within the Chinese cultural community of practice. The students’ demand for objectivity for the sake of fairness, on the one hand, reflects the use of objective tests as gatekeepers within the College English community of practice (Gu, 2005; J. Wang, 2007) and the English language assessment disciplinary community of practice (Davidson, 2004; Davison & Cummins, 2007; Kunnan, 2005). It also demonstrates the influence of the competitive assessment climate in the Chinese education system (Biggs, 1996a; Hu, 2003; Wang, 1996; Watkins & Biggs, 2001).

The teachers’ demand for objectivity could be a result of the institutional requirement to give objective grades for students’ performance. However, it could also possibly indicate that their perceptions of assessment still resided in the psychometric regime of language assessment, which features as a standardised and objective measurement of learning (Kunnan, 2005). In other words, the teachers’
understanding of formative assessment probably did not yet encompass the assessment for learning function. Their understandings of assessment and related concepts such as reliability and validity need to be extended to include the constructs of informal, classroom-based assessment (Davison & Leung, 2009; Lynch, 2003; McKay, 2005; Stobart, 2006). This assessment culture that demanded objectivity is, thus, another major factor that constrained the teachers’ assessment practices in both cases.

7.3.5 Teacher/student relationship

As revealed in the comparative analysis, the teachers, though aware of the students’ central role in learning (NKU-T6 was an exception), did not seem to realise an equivalent role for students in assessment. The teachers’ primary role to instruct, advise, and motivate in assessment was mostly maintained. This is similar to the findings of a series of studies (Cheng et al., 2004; Cheng et al., 2008; Rogers et al., 2007) that the Chinese EFL learners’ position in classroom assessment is marginalised due to the values embedded in the Chinese culture. The values can be identified as the belief of hierarchical teacher/student relationships (Biggs, 1996; Ho et al., 2001a; Hu, 2002) in the Chinese traditional culture of learning or the Chinese cultural community. According to this philosophy, the teachers, as authority figures, are in a better position to teach and assess (Hu, 2002). This deep-rooted belief, which is a cultural tool in the local context, seemed to have mediated the teachers’ understanding of the roles of teacher and student in assessment. This belief further influenced their decision regarding whether to involve students in assessment. The teachers’ understanding of their dominant role in assessment suggests that they still posited students as “the assessed” (Hull, 1943) and retained hierarchical teacher/student relationships. This perception, shared by most of the teachers, seems to be another reason for the limited involvement of students in assessment practices across the two cases.

The close relationship between teachers’ perception and their practice (Murphy, 2008) resonates with this finding about student involvement in assessment. Two teachers who supported learners’ taking a role in assessment did put their views into practice (NKU-T4, NKU-T7), whereas three teachers who opposed it did not involve students in assessment practice (KU-T2, KU-T3, KU-T2). However, there were cases when some teachers supported the idea of student involvement in assessment, but did
not practice this (KU-T6), while some teachers involved students in assessment, but still upheld the teacher-centred approach to assessment (KU-T1). The complexity indicates that the teachers’ assessment practice is by no means determined by one or two factors; rather, it is a combined result of multiple influences.

Students, for example, can be one of these influences. There were reports that even if the teachers wanted to involve students it was not easy to have students take up a role in assessment (NKU-T6, NKU-T7, KU-T1). KU-T1 mentioned this difficulty in his writing blog program, in which he introduced peer review:

KU-T1: Good students tend to give very constructive feedback whereas they seem to have complaints about receiving little valuable feedback from others. I encouraged them, saying it was kind of interactive learning.

Researcher: How do you understand the reasons for students’ different performance?

KU-T1: Proficiency disparity in my opinion, and also cultural differences.

Researcher: Oh?

KU-T1: By cultural difference, I mean according to the Chinese culture of learning students expect to learn from the teacher rather than from others such as peers. I elaborated on this point beforehand, telling them it was not necessarily so. Last term, in the blogging program I asked them to read from a reader’s perspective, not only to spot grammatical problems, but also to have an eye on the textual coherence and cohesion, and see the gap between yours and the blogger’s, then give a comment on the textual structure, logic and contents. This way, they have no more complaints.

It seems that whether students were given a role to play in assessment is one thing, but whether they could, and would, take it up is another. KU-T1’s efforts to address the students’ complaints and success in getting the students involved in peer review reveals one major point. That is, involving students in assessment is possible; but extra work such as clarifying the purposes and providing practical support is needed. Therefore, this is also a matter of whether the teachers saw the possibility; whether they prepared students effectively; and whether they knew how to use formative assessment for learning improvement.

The five factors identified above, that is, the influence of institutional policy, lack of professional training, teachers’ prior experience, the demand for objectivity,
and the hierarchical teacher/student relationship, seem fundamental and decisive to teachers’ assessment practices. These factors overrode the multiple differences between the two universities and worked together to contribute to the minimal changes to the teachers’ assessment practice.

### 7.4 FACTORS MEDIATING STUDENTS’ NEGOTIATION OF MEANING

Comparative analysis of the two student groups’ experience, and responses to the implemented assessment policy in the previous chapter, revealed two major findings. The first, the students in both universities continued to sit for the external CET4 regardless of its optional status. Their different views of the role of the test, the different status of the test in the institutional policies, and the teachers’ different practices in the classroom did not seem to affect their responses. The second finding was that no significant case-based differences emerged in the students’ responses to the implemented process assessment, despite the different assessment cultures and the different weightings utilised in the two institutions. However, individual variations did exist, with some students tending to orient more towards testing, while others maintained equal interest in testing as well as learning. The focus of this section is on the factors that mediated students’ responses to the external test, process assessment, and assessment in general. This is now discussed using a community lens (Wenger, 1998, 2000).

The cross case and individual case analyses revealed multiple factors mediating students’ negotiation of the meaning of the assessment initiatives. These factors are illustrated in Table 7.4.

<table>
<thead>
<tr>
<th>Factor</th>
<th>KU-Students</th>
<th>NKU-Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>University status</td>
<td>Key and prestigious</td>
<td>Non-key</td>
</tr>
<tr>
<td>Teachers</td>
<td>High-qualified</td>
<td>Qualified</td>
</tr>
<tr>
<td>Students</td>
<td>Top level in EEHE</td>
<td>Average level in EEHE</td>
</tr>
<tr>
<td>Weightings for process assessment</td>
<td>60%</td>
<td>10%</td>
</tr>
<tr>
<td>Assessment culture</td>
<td>More learning-oriented</td>
<td>More testing-oriented</td>
</tr>
</tbody>
</table>

Again, it seems the factors that differed between the two student groups, though important, were not decisive. It is the common factors between the two
student groups that exerted more decisive influence on the students appropriating the meaning of process assessment in practice.

Table 7.5
Contextual Factors Influencing Students’ Responses to Assessment (similarities)

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior experience as EFL learners in the Chinese context</td>
</tr>
<tr>
<td>A focus on summative assessment</td>
</tr>
<tr>
<td>Consequences arising from the use of assessment results</td>
</tr>
<tr>
<td>Hierarchical teacher/student relationship</td>
</tr>
</tbody>
</table>

These factors, which included 1) the students’ prior experience as an English language learner, 2) a focus on summative outcomes of assessment, 3) the hierarchical teacher/student relationship, and 4) the consequences arising from the use of assessment results, are situated in the shared sociocultural context. I discuss these in detail in the following sections.

7.4.1 Students’ prior experience

Students in both contexts reported that their experience in assessment was test-dominated and they all went through an intense period when their English learning was oriented towards the Entrance Examination of Higher Education (EEHE). Nonetheless, they told different stories about this common experience. The differences mainly related to, and in a sense were determined by, the economic and developmental zoning of the regions that they were from (developed or under-developed) and the school types (key or non-key). Those from under-developed regions and non-key schools such as KU-S1, KU-S4, NKU-S3 and NKU-S4 considered that their schools were inadequate in terms of qualified teaching and learning resources. They believed that the curriculum they experienced was narrowed to focus on the EEHE so that they would be able to compete with students from developed regions and key schools. In other words, there were high levels of washback from the EEHE into the curriculum which limited the content and materials leading to higher education. As a consequence of the washback, these students identified themselves as struggling English learners at universities. They seemed to adopt very instrumental approaches to learning and assessment and tended to orient more towards testing. This finding is similar to that of Hu (2003) that the developmental zoning of the region that students come from is an important influence on the students’ approach to learning and assessment.
The sociocultural perspective emphasises the influence of prior experience on one’s present behaviours (Vygotsky, 1978; Wenger, 1998). The students’ stories about their past EFL learning and assessment experiences demonstrated how their own identity and discursive histories helped make sense of the assessment changes and their corresponding responses. The difference in the students’ stories is thus a major reason for the variations in the individual students’ responses.

7.4.2 Focus on summative outcomes

The students’ responses to process assessment featured a concern about assessment outcomes. The students were concerned about the process grade, that is, the outcomes of process assessment. As a result, they adopted a strategic learning approach to the assessment procedures. Some students, especially those who were from under-developed regions and were struggling in their English learning, took the process grades as their goal of English learning. Their responses indicated that the students understood the purposes and the nature of the assessment as summative. This aligns with the findings by Ramsden (2003) and Brown and Hirschfeld (2008) that there is a close link between students’ views of the nature, purposes and procedures of the assessment process and their subsequent responses.

The students’ responses also demonstrated their perception of the goal of learning, which was to achieve high scores in assessment. This reflects Tang and Biggs’ (1996) finding that Chinese students often take the assessment purpose as their goal of study. The students’ strategic learning approach resonated with the results reported by Tang and Biggs (1996) and Entwistle and Entwistle (1991) that when students study for an exam, they try to understand the learning materials in ways that they perceive will meet test requirements and they adopt the approaches that will maximise their grades. Overall, these findings once again confirmed the shaping force of assessment on students’ learning approaches, which has been established in research (Biggs, 1996; Black & Wiliam, 1998, 2006; Crooks, 1998; Gibbs, 2007; Hargreaves, 1997; Struyven et al., 2006).

The students’ responses seem to reflect a marked influence of the sociocultural contexts. To be specific, the traditional conceptions of learning in the Chinese cultural community, which is largely oriented towards outcomes and examination (Cortazzi & Jin, 1996; Jin, 1990; Biggs, 1996b; Ho et al., 2001a), seemed influential on the students’ behaviours. This influence is also evident in relation to the students’
decision to take the optional CET4 in order to demonstrate their learning. It was represented in the statements that studying and taking tests was a habit for them (KU-S1, KU-S3, KU-S4). It was also apparent in the unquestioned status of tests in the local context (Cheng, 2009), which was demonstrated by NKU-S5, who asked “Is there anything wrong with tests?”

Hodkinson et al. (2007) referred to the cultural settings’ influence on actors’ behaviours by stating:

Actors always operate within systems of expectations: the expectations they bring to the situation and the expectations that others have about their activities and practices…Expectations are not necessarily consciously held. They exist as ‘ways of doing’ and ‘ways of being’ that are considered to be ‘normal’. This, finally, also means that learning cultures are governed by values and ideals, by normative expectations about good learning, good teaching, good leadership, and so forth — and again, these are from ‘within’ and ‘outside’ any particular setting…(p.419)

The students’ conception of learning is influenced by their culturally-based expectations, which mediated their thinking and responses to assessment. Taking tests and wanting a summative result from their learning, therefore, was a cultural practice within this particular context. This cultural perspective explains the students’ concern for process grades as well as their continued interest in the taking CET4 even though it had been made optional.

7.4.3 Uses and Consequences

The students’ responses to assessment demonstrated an influence of the consequences arising from the use of assessment results. For instance, the students concentrated on the use of process grades for reporting purposes, which were linked to the completion of their degree. Therefore, they strove to gain the highest possible scores. This was made manifest by both the teachers and students: “everyone wants high scores” (KU-T3, NKU-T7, KU-S4, NKU-S1). For another example, some students (KU-S4, NKU-S1, NKU-S5) registered for CET4, even though it was optional because they wanted to use it for employment or scholarship purposes. This interaction between assessment and social consequences confirmed the shaping power of assessment on the students’ approaches to learning (Boud & Falchikov, 2007; Marton & Säljö, 1997; Ramsden, 2003; Struyven et al., 2006).
The influence of the social world on learning and human actions in general has been recognised in sociocultural theory. Lave and Wenger (1991) state that:

In our view, learning is not merely situated in practice—as if it were some independently reifiable process that just happened to be located somewhere; learning is an integral part of generative social practice in the lived-in world.  
(p. 35)

The influence of the social world on learners’ approaches to learning is evident in the Chinese context as a traditional utilitarian view of learning (Lee, 1996; Hu, 2002). It is also demonstrated with the examination equation (Ho et al., 2001a). The belief that success in learning, that is, successfully “pass[ing] a test” (Cheng, 2009 p. 267) would bring about a sense of achievement, substantial benefits and upward mobility in society is evidenced in the students’ responses. For these particular students, taking the CET4 test and achieving a grade was linked primarily to utilitarian aims related to jobs and scholarships. Learning English for these students was strategic and extrinsically motivated.

7.4.4 Developing a new identity as learners

The students’ responses to teacher feedback showed a tendency towards self-reliance. Specifically, although there were students who still relied on teachers for instruction and knowledge transmission (KU-S1, NKU-S3, NKU-S5), most of the students indicated the minimal instructive value of teachers’ feedback. Rather, the students valued their own self-evaluation and reflection (KU-S2, KU-S3, KU-S4, KU-S5, NKU-S1, NKU-S2). Also, they tended to use self-reflection to regulate their study. That is to say, these students had, to some extent, developed an identity as an independent learner who could take responsibility for their learning.

This finding about students as independent learners seems to contradict the stereotyped view of Chinese students as passive and dependant learners (Biggs, 1996a; Dautermann, 2005; Ginsberg, 1991; Hu, 2002 Paine, 1990). It also challenges the research that maintains Chinese learners prefer teacher feedback (Meyer, 1999; Zhang, 1995). The finding, however, provides new evidence to support the claim that stereotyped norms of Chinese culture of learning are changing (Cortazzi & Jin, 1996; Meyer, 2003; Jin et al., 2006; Shi, 2006). How these students developed this new identity as independent learners, and whether or not the development of this identity
was related to the assessment policy change, was not investigated in this study. This is an area for further research.

While those students who did not have experience in peer-assessment held a reserved attitude to it, the students who had the experience held a largely negative attitude to peer-assessment. This seems to disagree with the recent research (Hu & Lam, 2010; Miao et al., 2006; Meyer, 2003) that claimed Chinese students have an open attitude to feedback from peers. This finding, on the other hand, illuminated a potential barrier to the future implementation of formative assessment, which advocates peer involvement (ARG, 2002; Black et al., 2003).

7.5 SUMMARY

The above discussion revealed that the individual, the sociocultural environment, historical tradition and current reality, and various institutional relations, functioned together to influence the negotiation of the meaning of formative assessment in the contexts of the two universities. It was found, among the multiple influencing factors, that some sociocultural factors played a decisive role in the representations of the formative assessment as it operated in practice in the two College English language teaching and learning sites.

The first was the summative use of assessment results, which is a cultural value that is embedded in the situated context and was distributed across the shared communities of practice. It functioned as a tool for thinking and mediated the local actors’ responses to formative assessment. Through the mediation of this tool, formative assessment was interpreted as process assessment in the institutional policies; it was implemented as progressive mini-tests by the teachers; and was approached pragmatically for the sake of high scores by the students.

The second factor was the demand for objectivity which prevailed among local actors at various levels in the two sites. This appears to be a legacy of the history of the universities as situated communities of practice and the local actors’ prior experiences within the Chinese education system. Because of this perception, both institutions demanded accountability from teachers’ grading practices and the students demanded reliability from the teachers’ professional judgement of their performance. The teachers were thus presented with the challenge to provide an objective judgement for process assessment.
The third factor that emerged as a decisive influencing factor was the hierarchical teacher/student relationship, which was, again, a value transmitted from tradition and embedded in the Chinese cultural community. It contributed to the lack of recognition of the student’s role in assessment in the institutional policies and in most teachers’ assessment practices. From the data, it appeared that some of the students seemed to have developed identities as independent learners and yet were given few opportunities to take a central role in assessment.

These three interrelated factors formed the assessment environment that constrained the implementation of formative assessment in the local context. This assessment culture, combined with lack of effective training and support for the teachers, and the social consequences on the students, significantly influenced the limited policy change.

Apart from the local actors’ largely collective responses, individual variations also emerged. For instance, some individual teachers incorporated elements of formative assessment such as involving students, sharing criteria and providing constructive feedback in practice. It was found that these practices were personal, self-initiated and afforded by the teachers’ cross-boundary experiences and knowledge (Wenger, 1998, 2000) (KU-T1, NKU-T7), and personal pedagogical endeavours (KU-T4, NKU-T4).

The sociocultural perspective regards social activity as resulting from “active and dynamic contributions from individuals, their social partners, and historical traditions and material ... between the individual and the social and cultural environments” (Rogoff, 1995 p. 140). These multiple influences and relations were evident in the negotiation of the meaning of formative assessment in the two universities. The implications will now be discussed, along with recommendations for assessment policy and classroom practice.
Chapter 8: Implications, Limitations and Conclusion

This final chapter starts with a brief overview of the study (Section 8.1). The implications for policy and practice are then presented (Section 8.2). Following this, the original contributions (Section 8.3), the limitations of this study (Section 8.4), and the directions for further research (Section 8.5) are outlined. Lastly, a conclusion is drawn from my experience of learning (Section 8.6).

8.1 AN OVERVIEW OF THE STUDY

This study investigated, from a sociocultural perspective, the implementation of the recent College English assessment policy change (CMoE, 2004) within two Chinese universities. It was inspired by the researcher’s personal interest in seeking solutions for the testing and learning dilemma associated with College English and English education in China. More importantly, it was inspired by an urgent research agenda prompted by a large scale policy change.

In 2004, the China Ministry of Education (CMoE), via the College English Reform Program (CERP) (CMoE, 2004), proposed two major changes to the College English assessment. The first was to encourage the optional status of the external test, CET4/6, which had proved to be detrimental to English learning for its largely negative washback. The second was to advocate the incorporation of formative assessment into the original summative framework. Research has supported the learning function of formative assessment in the context of ELL (Brown & Hudson, 1998; Davison & Leung, 2009; Hamp-Lyons, 2007). These policy changes were intended to improve the quality of ELL via triggering changes to teachers’ assessment practice. Nonetheless, the implementation of formative assessment principles required recognition of the challenges arising from the deep-rooted local assessment tradition and culture.

Research into the interplay of the principles, the policy and its enactment in practice was a priority, so as to discover the potential of this process in the local context. The formative assessment initiative was the focus while the external testing tradition remained in the background.
Prompted by the research problem that emphasised context and interplay, this study adopted a sociocultural perspective. This approach explicates human actions through their interactions with culture, history and the social structure of the situated context (Rogoff, 2003; Wertsch et al., 1995). Accordingly, the investigation was located not only in a global changing climate motivated by the theoretical development in cognition and assessment, but also within the Chinese culture of learning which has a tradition of examination-dominated assessment.

This study adopted a case-study design and addressed the research problem by ethnographically documenting the lived experience of the key stakeholders such as the institutional policy-makers, the teachers and the students. Two universities, KU and NKU were selected to represent the two university types in the Chinese higher educational context, that is, key university and non-key university. Rich and detailed data were collected using interviews, classroom observations, and artefacts. The data were analysed through interactional and relational dimensions both within and across the two cases.

The individual case analyses in Chapters 4 and 5 and the cross-case analysis in Chapter 6 revealed similarities in the two universities in relation to the research questions. Specifically, in both cases formative assessment was interpreted in the institutional policies as process assessment and was implemented as progressive mini summative assessments by the teachers. The students were pragmatic in their approaches to assessment and focused on the summative end of attaining scores. The teachers reported minimal change to their assessment practices. The students reported an impact on their study time and their distribution of effort, but no change to their approaches and orientations to learning. Based on these findings, it was concluded that the policy intent was reflected in classroom assessment practices only to a limited extent in both cases. The major differences between the two cases, including the key or non-key status of the university, the experience and qualifications of the teachers, the English proficiency levels of the students, the institutional culture (empowerment or control), and the assessment culture (testing or learning-oriented), did not seem to affect the outcomes.

A sociocultural theory of learning was used to explain the outcomes of this research. An analysis from a community perspective (Lave & Wenger, 1991; Wenger, 1998, 2000) revealed that the sociocultural factors that the two universities shared...
within the situated community of practice were more influential than the institutional local differences. The values and traditions embedded in the sociocultural context such as a summative view of assessment, the hierarchical teacher/student relationship, and a testing culture that demands objectivity, acted as cultural tools (Rogoff, 2003) that mediated the local actors’ understanding and experience.

Overall, it can be said the formative purpose of assessment was not recognised or accepted by the stakeholders. In other words, formative assessment was marginalised by the prevailing view that assessment fulfilled a summative function. It was concluded that the key to the effectiveness of formative assessment in the Chinese context is not only how the policy is implemented, rather, it lies in the local actors’ conceptualisation and valuing of assessment. It is not the policy initiative and the implementation that decides the assessment reality; rather, it is the local actors’ view of assessment and its nature, their understanding about the relationship of assessment and learning, their understanding about who is rightly positioned to assess that are decisive factors. It is the local actors’ valuing of the learning product or process, measurement or improvement that is decisive.

8.2 IMPLICATIONS

From this exploratory inquiry, the following implications are identified for educational assessment policy making at national and institutional levels, and assessment practice for teachers and students.

8.2.1 National level policy

Assessment needs to be “fit for purpose” (ARG, 2009a), which means that policy makers need to be aware of key assessment principles pertaining to purposes and functions. If teachers’ professional judgement is to be used for accountability purposes (internally or externally), agreed criteria and standards are needed. If system level criteria and standards are not fully communicated to practitioners, reliability will be an issue, as evidenced in the two cases. If teachers’ feedback and judgement are used for the formative purposes of showing learners how, and where, to improve, then validity is important. The latter focuses on whether the feedback is constructive and whether it informs learning.

The CECR, while advocating the incorporation of formative assessment into the original summative assessment framework, designed the new assessment
approach for the purposes of measuring as well as enhancing learning. However, the discourse used to introduce formative assessment aligns with a testing regime. To illustrate: “Assessment is an important part of College English curriculum. Comprehensive, objective, scientific and accurate measurement is crucial for objectives to be fulfilled” (CMoE, 2007 p.5). The interpretation that can be made from this statement is that assessment (including formative assessment) should be objective and used for measurement purposes. This message, in the Chinese sociocultural context, communicates a number of purposes for formative assessment, which included an emphasis on measuring as well as improving learning. This message is, at least in part, responsible for how formative assessment was interpreted and implemented within the two cases. The ongoing emphasis on measurements and objectivity means that it was implemented as process assessment with smaller, ongoing tasks that were generally more interactive and student-oriented but designed to be graded summatively.

**Policy representation**

Policy should explicitly outline directions for practice (Ball, 2006; Fullan; 2004). When policy ideas are new to the local actors, such as formative assessment in the CERP, the policy needs to be presented explicitly to avoid misunderstanding and misinterpretation. To this end, policy needs to be communicated in such a way that the actors at various levels understand the goals of the change. More importantly, the policy needs to make manifest the rationale for the change. As research demonstrates, little or no full understanding of the deep underlying principles is often related with limited change (Spilliane et al., 2002).

**Training**

The role of training for practitioners cannot be overemphasised in educational innovations (Fullan, 2004; Snow, 1989). Training and technical support are essential prerequisites for the successful implementation of policy to ensure that policy messages are clearly delivered and practitioners are equipped with the requisite knowledge and skills. As Snow (1989) noted decades earlier, no change is possible without effective training. In this study, a lack of effective training in formative assessment led to the teachers falling back onto their culturally, socially and historically embedded understandings and experiences; as a response, they made minimal changes to their assessment practice.
The role of training is more crucial (Spilliane et al., 2002) for a policy change which is imported from foreign cultural contexts and involves fundamental change in conceptualisation. This is the case for the CERP. Superficial knowledge about techniques and skills are not sufficient for assessment practice to enact change from a formative orientation in the Chinese context, where there is an entrenched history and understanding about teaching and testing. As revealed in this study, the culturally situated views of learning, assessment, and the roles of learners and teachers, are still dominating and exerting powerful influence in the local context. Hence, for formative assessment to take root and flourish in such a context, teachers need to be provided with formal professional development that leads to “professional learning” (ARG, 2009b) and a fundamental change in understanding and attitude. To achieve this purpose, teachers’ professional training should start with initial teacher education (Carless, 2008) and be maintained throughout their teaching career.

Systematic and all-around training in large scale reforms, which involves a large number of practitioners, can be hard to achieve within a short period of time. In this case, effective training and continuous support in technical terms as well as theoretical terms should be made, first of all, available to policy-makers at institutional levels. As such, the essence of the policy change would be retained rather than misinterpreted in the institutional policy.

### 8.2.2 Institutional level policy

Institutional policy-makers need a clear understanding of the context and the policy message to make appropriate decisions that accommodate the local realities and the essence of the policy message. To achieve this purpose, the institutional policy-makers need to have both adequate professional expertise in ELL and assessment. First-hand, bottom-up information from the practitioners and the students is necessary, which on the one hand, can ensure the decisions are well-informed; on the other hand, can ensure commitment to the policy rather than resistance.

Institutional policies, such as the policies developed by KU and NKU, specify the policy message into specific procedures and processes, and carry great weight in the translation of policy into practice. This influence was clearly articulated by the teachers and evidenced in the teachers’ practices. The teachers’ implementation of
formative assessment as process assessment, for instance, was related to the interpretation and representation of formative assessment in the institutional policies.

Policy delivery

Institutional policy needs to be delivered so that the policy message is understood by practitioners and students, and captures their attention. If not, the impact of the policy change might be compromised. This was evidenced at NKU. The informal delivery of the formative assessment initiative, for example, meant that the NKU teachers did not see it as profound and influential. Consequently, some of the teachers did not seem to take it seriously. The negative influence of ineffective delivery was also manifested with the changed status of CET4 at NKU. Teachers, not well-informed of the change, tended to respond to it differently in practice. NKU-T6, for example, continued to practise test-preparation. The informal delivery was also, in part, responsible for the NKU students’ testing-oriented practice, which remained relatively unaffected by the changed status of the external test.

Given that much of the research on formative assessment has occurred outside of China, the institutional policies as well as the national policy need to make the rationale for the change and the underlying principles clear to the teachers. Some teachers, even when supported with training, were found to carry out formative assessment to “the letter” rather than “the spirit” (Marshall & Drummond, 2006). Without training and clear understanding of the paradigm shift underlying the policy change, the teachers will not be able to see the point of doing assessment in alternative ways, and may not attempt to do it at all. This was the case with many of the teacher participants in this study.

Policy implementation

A supportive institutional culture that trusts practitioners is important for policy changes to be effectively implemented (Carless, 2009; Wang & Cheng, 2005). This argument was evident in this study with the NKU teachers, whose attitude to the policy change and assessment practices was affected negatively by the TAO’s distrust. A sociocultural approach assumes that “the social structure of practice, its power relations, and its conditions for legitimacy, define the possibility for learning” (Lave & Wenger, 1991 p. 98). Therefore, the greater flexibility that KU gave to its teachers and the more democratic and open organisational culture at KU meant that these teachers had more opportunities to participate and learn.
To encourage teachers to work towards a formative assessment direction, the institutions also need to take some measures. The first is to acknowledge the teachers’ effort in the appraisal system. Failure to recognise teachers’ effort in appraisal was reported as discouraging the NKU teachers and their assessment practices. The second measure is to encourage and provide financial support for teachers to participate in training and communication within the community of practice. The financial issue had discouraged some NKU teachers from attending the College English assessment teleconference. Third, hands-on technical support, practical and handy tools such as *Assessment Experience Questionnaire* (Gibbs & Simpson, 2003), the check list by Hamp-Lyons (2007) could also be provided.

### 8.2.3 Teachers and teaching practices

Teachers are recognised as the key to successful educational innovation. Fullan (2004) makes this point: “Educational change depends on what teachers do and think–it is as simple and complex as that” (p.129). Teachers in this study were faced with multiple challenges. These challenges, which were logistical, technical, institutional, and cultural, individual and societal, impacted on their assessment practices. The findings of this study implied some principles that teachers can take up in their assessment and teaching practice so as to move towards real change. The principles are as follows.

**Professional learning**

Apart from the assessment training provided by the system, teachers can seek to develop literacy in assessment in their professional career. That is the case with some of the teachers in this study. KU-T1, for example, developed his skills in his overseas studies. NKU-T4 and NKU-T7 extended their assessment skills and knowledge through their own readings and pedagogical endeavours.

Teachers can also involve themselves in networking within the College English community of practice and the ELL disciplinary community at large. They can achieve the networking by reading journal articles or by participating in community activities such as conferences. In this way, teachers can keep abreast of the disciplinary development and be conscious of the change and the rationale for it.
**Constructive feedback that aims for improvement**

Feedback that highlights encouraging value such as “well done” or “a good job” can be motivating, but praise itself cannot provide any specific advice that informs better learning (Gibbs & Simpson, 2003). As evidenced in the data, this kind of feedback could not satisfy students’ need and therefore was seldom acted upon. They sought for specific and constructive advice about what to do next and how to do it to improve their ELL.

There is no denying that constructive feedback that caters to individual students takes time and effort on the teachers’ part. Moreover, logistical factors such as heavy workload at NKU and the time limits are also constraints. Teachers need to know that this kind of input is rewarding, especially when improved learning is the focus and objective.

**Transparent criteria**

Shared criteria are needed when the assessment evidence collected in the teaching and learning process is used for accountability purposes (ARG, 2009a). The lack of uniform criteria in systems, as demonstrated in the study, was found to link directly in both cases to the issue of maintaining objectivity in assessment in both sites.

If evidence is to be used for formative purposes, criteria are needed. A ranking grade such as “A” or “D” without giving further information does not inform learning. But criteria which are specified and made transparent to students do provide students with knowledge about where they are in relation to certain standards and the learning goals. The study found that the use of tacit understanding of the criteria was commonplace for the teacher participants and caused confusion for the students. This indicates a need for the teachers to use criteria that are explicit and transparent. This need is a highly regarded assessment for learning principle (ARG, 2002).

Another finding revealed in the study was that while the teachers reported they did use and share criteria with students, the students unanimously reported no understanding about the criteria their teachers used. This quite possibly was due to the criteria that were not explicitly stated or not clearly communicated. An implication of this finding is that teachers need to pay more attention to sharing criteria.
Involving students in assessment

While it is well acknowledged that students have a central role in assessment as expressed in formative assessment principles (ARG, 2002; Black et al., 2003), only a few teacher participants (KU-T1, NKU-T4, NKU-T7) involved students in classroom assessment. Also, given the new identity that some students had developed as independent learners, the teachers do need to shift their understanding about the roles of students not only in learning but also in assessment. Teachers need to see that, apart from passing the internal and external assessment, the students also need to develop meta-cognitive strategies about how to learn. Accordingly, teachers need to give students more opportunities to learn from participation and learn how to take on their central role as assessors.

Students did need instruction and support as to how to participate in evaluative practices, as KU-T1 and NKU-T7 reported. Nonetheless, it was also reported that, with support, students could play their role as an assessor well (KU-T1, KU-T6, NKU-T7). This finding, aligns with other research evidence (Orsmond et al., 2000; Struyven et al., 2008), and highlights the significance of teachers’ guidance. It calls for teachers’ attention and effort in this regard, and to make assessment part of their pedagogy.

8.2.4 Students and English Language Learning (ELL) practice

The findings also hold implications for students and their approaches to English Language Learning practice, as presented below.

Focus on ELL rather than assessment

It is agreed that when the stakes are high, students’ approaches to learning are more likely to be shaped by assessment (Biggs, 1996; Black & Wiliam, 1998, 2006; Broadfoot, 2007; Crooks, 1998; Gibbs, 2007). This shaping power of assessment was demonstrated in this study, particularly in the students’ views of the external test and the summative outcomes of process assessment. As also revealed, this view has its roots in history and the social context. It holds responsibility for the impact of the formative assessment initiative on students’ time distribution but minimal influence on their approaches to ELL.

For formative assessment to function, students, first of all, need to understand that measurement, namely, how much and how well you have learned, is important.
Nonetheless, learning and how to learn are the focus and purpose. Research has established the close relationship between making students informed about the assessment purposes and the positive impact on students’ learning (Carless, 2008; Harlen & Crick, 2003). The primary thing that students could and should do is to focus on ELL and how to learn.

This can be best achieved when assessment results collected from teaching and learning processes are not used for solely accountability purposes. Nonetheless, as the KU-S3 stated, when students focused on learning and tried their best on every task, the results would take care of themselves. The relevance of deep learning for better learning results is evidenced in research (Entwistle & Entwistle, 1991; Tang & Biggs, 1996). That means, if students can see this point and make the effort they may not need to worry about the assessment results including the external test.

**An active role in assessment**

Given that formative assessment is integral to learning, students need to take a role in assessment to bring the potential of formative assessment to enhance learning. This is not easy for students as their reliance on an authoritative figure, such as the teacher for assessment is traditionally rooted (Hu, 2002). This phenomenon may have a lot to do with whether the teacher gave students the opportunity to participate. Nonetheless, if students realise the need for the identity as an independent learner, they will seek to develop the skills by actively engaging with assessment. Some students involved in this study seemed to have already developed into independent learners and had taken control of their ELL.

Another finding revealed in this study is that the students did not seem to realise the potential benefits of learning with and from peers. This was attributable to the summative view of assessment. However, it may also relate to their perception that learning is acquired through transmission rather than participation. The implication of this finding is that students need to understand that they develop English proficiency through participation apart from by memorisation only. Therefore, it is important to develop collaborative relations with peers. This is a matter of conceptualisation. It could be said that the collaborative relations between students are more likely to happen when their understanding of learning is reconceptualised and when students see the potential to develop English proficiency via participation.
Perceiving teacher as guide and scaffold

Teachers are traditionally regarded as authoritative figures in the local context. This holds true with most of the students involved in this study, even though they did not seem to rely heavily on teachers in ELL as is assumed. While teachers’ authority is needed to provide opportunities and scaffold students’ involvement in assessment, students need to know that the point of having an authoritative figure in classroom is to facilitate their ELL. Teachers can provide guidance to develop learning and skills about how to learn. This applies not only to ELL but also to assessment. Students need to know that: they can have a say in negotiating the criteria and standards; they can go and ask for clarification of criteria if they are not clear about them; they have the right to ask for a justification for the assessment results rather than remain silent and assume they are at fault (as experienced by NKU-S5). Overall, students themselves need to realise that they have a role to play in ELL and assessment, and value the opportunities when available.

8.3 Significance

This study is significant in addressing a timely research agenda prompted by the launch of the CECR, a system-wide assessment reform. It is also important for the original contributions it makes in relation to three identified research gaps: 1) the theoretical developments of formative assessment and their application to the Chinese context; 2) the interactions between the CERP assessment policy initiative and assessment as practised in the local context; 3) the conditions necessary for formative assessment to support ELL in the Chinese context.

This study has examined the alignment of formative assessment to local policy and practice. The theoretical principles of formative assessment as disseminated in the global context (ARG, 2002; Black et al., 2003) have been considered. It was found that the definition of formative assessment of the CECR seemed to align with the theoretical principles as espoused in the international literature (ARG, 2002; Black & Wiliam, 1998; Black et al., 2003). Nonetheless, formative assessment was situated in a discourse that emphasised objectivity and measurement. In the two cases, formative assessment was adopted and adapted as process assessment. Process assessment adopted some elements of formative assessment such as maintaining a focus on process, but did not emphasise the learning enhancement function; rather
the results from process assessment were used for summative purposes. In this study process assessment was identified with the fourth formative assessment typology (Davison & Leung, 2009) and close to the informal summative assessment along the assessment continuum (Harlen, 2005). This theorisation of process assessment in the Chinese context is original and makes a contribution to the formative assessment related research in the field of ELL.

There has been research on the implementation of formative assessment in the Chinese context (Guo & Li, 2007; Xu & Liu, 2009). Nonetheless, applying sociocultural theories of learning and assessment makes this study unique. With a sociocultural perspective, the policy change and its implementation were viewed as an interactive process rather than a linear one. The dynamic agency of local actors in the negotiation of the meaning was, accordingly, foregrounded. The ethnographic documentation of their actual experience via triangulated data, such as interviews, artefacts and observation, substantiated the claims that this study made. The sociocultural perspective also enabled this study to look beyond the policy and its implementation, and to incorporate the historical, cultural, institutional and social perspectives to observe the local actors’ experiences. Consequently, this study extended its exploration from a description of what was happening to an understanding of the underlying factors that contributed to the outcomes. It was found out that the policy implementation issues such as delivery modes and control did influence the teachers’ uptake of the policy initiative, but were not determining factors. Rather, the local actors’ perceptions and values, which were deep-rooted in the Chinese culture of learning, were decisive. This finding sheds new light on the policy research and, meanwhile, highlights the importance of respecting the local realities when introducing system-wide change. Very few studies have investigated assessment policy change in the Chinese context from this sociocultural perspective.

Based on the analyses of the interactions between formative assessment principles, policy and practice within two localised contexts, practical recommendations were made from the findings. These recommendations, which relate to the policy process at national and institutional levels emerge from an understanding of the factors that influence assessment policy and practice change in the Chinese context.
8.4 LIMITATIONS

Despite its original contributions to formative assessment research in the Chinese and disciplinary contexts, I acknowledge this study has its limitations. The limitations of this study relate to the scope and duration.

The study adopted a case study approach and selected two cases to represent key and non-key university types. Nonetheless, I acknowledge the uniqueness of each key and non-key university in the Chinese higher education context. Therefore, it should not be assumed that other key universities will be exactly the same as KU and that other non-key universities will be similar to NKU. Moreover, the study focuses on 13 teachers and ten students, and although it is a detailed investigation of their experiences of the assessment policy change, its scope is limited. To further investigate how other teachers and students have adapted to the policy change, the findings of this study could be complemented and further validated by large scale quantitative studies. In the example of the finding that a stronger learning orientation is more apparent in the KU students’ attitudes as compared with a stronger orientation to examination with the NKU students, further case studies could be conducted with other and more participants. In this way, the findings of this study would be given more validity and robustness (Yin, 2003).

Another limitation of the study was the duration for the collection of the data. The conditions of the CSC scholarship mean that I was only allowed a limited amount of time away from Australia. Consequently, the data collection phase in China, apart from the pilot study, took place over a period of only four weeks at each site. A longer period of time would have allowed me time to collect more detailed and comprehensive data about the implementation and uptake of the policy change. It is possible that the data would have provided further insights for the research.

8.5 DIRECTIONS FOR FURTHER RESEARCH

This exploratory inquiry into the effect of the introduction of formative assessment into the College English assessment framework was not able to study all the areas of this large scale reform. One suggestion for further study that emerged from this research is the policy-making process at the institutional level. Given the significant influence of the institutional policies on classroom assessment practice, this area needs to be researched further to understand how the structure of the
institutions and the management of the change at the institution level impacts on classroom practice. A second possibility for further research is the theorisation of formative assessment within the Chinese context to provide the practitioners with specific procedures to follow. A third possible direction is research into how the trustworthiness and fairness of teachers’ professional judgement need to be ensured through the implementation of moderation so that valid and reliable judgements are made within formative assessment in the education system. Another question to be resolved through research relates to teacher education. While teacher training is the most apparent means for teacher development, questions remain as to how teachers within the Chinese sociocultural context can be best upskilled in formative assessment literacy. Finally, another area for future research could be investigating the equity issue associated with assessing students from different development regions.

8.6 CONCLUDING PERSONAL REFLECTIONS

As a College English teacher, I have experienced an intense professional learning experience in conducting this study. Lessons drawn from my learning experience inform the conclusion, given that:

The indigenous production of practice makes communities of practice the locus of creative achievements and the locus of inbred failure, the locus of resistance to oppression and the locus of the reproduction of its conditions; the cradle of the self and also the potential cage of the soul. (Wenger, 1998, p. 85)

Three years ago, I was granted the CSC scholarship to study at QUT for the Doctorate of Philosophy (PhD). The scholarship gave me legitimate access to QUT, a different cultural and academic community of practice. I left China, crossing the boundary of multiple communities to which I belong: College English and Chinese cultural community of practice. Engagement with research entitled me to membership of the QUT research community of practice.

I started my research program as a participant positioned on the periphery of the research community, as a novice who was keen to learn more about conducting research in assessment. The dyadic master-apprentice relationship with my supervisors provided guidance and scaffolding, which helped me to learn by
observing, by interacting (with peers and more advanced apprentices), by participating (through research and publishing), and by appropriating. Through engagement with these communal academic activities, my competence and skills as a researcher have developed.

This learning experience has been rewarding and challenging. Studying at doctoral level in a second language and in the field of assessment were new experiences for me. I had to reconceptualise my view of learning and assessment using a sociocultural theoretical perspective. The lesson learned is that reconceptualisation needs, first of all, legitimate access to the shared repertoire; it also needs hand-on support from experts, opportunities to observe, participate, and interact in the learning environment. This takes time.

Given the CERP policy context, the minimal impact of the policy change on practice is understandable. From my experience, it can be said that the road forward of implementing formative assessment in China will not be smooth. Change will probably take a considerable amount of time to be realised in practice. It needs open access to “boundary objects” such as international journal articles and other sources; it needs “brokers” such as KU-T1 or those who complete studies in different contexts. It also needs practitioners’ personal endeavours and the support of the educational and social systems. The introduction of formative assessment initiative is a first step. Much more is needed to bring about a real change towards a learning orientation, a learning culture, and improved ELL through the use of formative assessment.
References


25(3), 513-525.


Clegg (Eds.), *Innovative assessment in higher education* (pp. 37-47). London,
New York: Routledge.

*Pedagogy and practice: culture and identities* (pp. 28-39). Los Angeles, London:
Sage.

learning and society: Investigating practice study guide*. Open University:
Milton Keynes.


in Education, 14*(2), 149-170.

Teaching and Research (waiyu jiaoxue yu yanjiu), 33*(2), 140-143.

Heinle.

613.

Biddle, T. Good & I. Goodson (Eds.), *International handbook of teachers and 


importance of combining quantitative and qualitative research methodologies*

Orsmond, P., Merry, S., & Reiling, K. (2000). The use of student derived marking 
criteria in peer and self-assessment. *Assessment and Evaluation in Higher 

*Teachers College Record, 92*(1), 49-81.


motivation and achievement: Culture, motivation and achievement (pp. 73-118). Greenwich, CT: JAI Press.


References


Appendices

Appendix A
Senior administrator interview schedule

The curriculum arrangement of College English teaching in the university

- credit points
- compulsory status of College English

The College English assessment policy

- Changes made since the College English Curriculum Requirements (CECR) (CMoE, 2004, 2007) in terms of CET 4/6, the external test and the internal assessment of individual coursework
- The rationale for these changes
- Plan for further change

Delivery of the policy

Relevant training and support provided to teachers
Appendix B
Teacher Interview Schedule

Section 1 About you.

Gender:  male □
                        female □

Age range: 34 and below □, 35-44 □, 45 and above □.
1. Your professional title?
2. How many years have been teaching College English?
3. How many students are there in the class you are teaching?

Section 2 About your understandings of the current assessment policy

Q1. How do you usually go about CE assessment practice in your College English classroom?
Q2. Do you practice assessment totally according to the current College English assessment policy in your institution or…?
Q3. Could you please describe briefly the current College English assessment policy in your institution?
Q4. What changes have you seen in the College English assessment policy?
(Que描述一下贵校的现行大学英语的课程评价政策，以及相对于前政策的变化)
Q5. How do you understand the intent of CE assessment policy change?
(您是如何理解贵校现行大学英语的课程评价政策的主旨？)
   - Have you been informed of the rationale behind the College English assessment policy change? (其原因及根据您知道吗?)
     - Yes – how?
     - How have you been trained regarding the specific College English assessment procedures before the assessment policy implementation?
       (具体实施前可曾有过培训或指导?)
         - Yes – please describe briefly the focus of the training.
         - Have you received any assessment related training before?
           - Yes – please describe briefly the focus of the training.
           - No – Where do you think the knowledge and skills that you use to assess your students come from?
Q 5. Please give your opinion regarding (请阐述您的观点)
   1. Making the external assessment (CET/PETS) optional.
      (CET考试变为非硬性要求)
      - How do you understand the rationale?
      - Are there any effects on your EFL teaching practices?
      - Could you please give some details?
      - Have you seen any effects on students' learning?
- More details?

2. Making the internal assessment formative and summative.
   - How do you view the rationale for it?
   - Are there any effects on your approach to EFL teaching?
   - Some examples of formative assessment in your classroom assessment practices?
   - Have you seen any effects on students’ approach to EFL learning?
   - Some details?

3. Involving students in assessment.
   - How do view the point of it?
   - How do you involve students in the classroom assessment practice?
   - How have your assessment practice changed?
   - Have you seen any effects on students’ learning?
   - Yes, please specify.
   - No, what could the reason be?
   - Are there any salient differences in students’ participation in these assessment activities?
   - How do you understand the reasons for students’ different performance?

Q8. How do you give feedback to your students?
(Prompts: Orally or written, once or progressively)
(您给学生的反馈一般是口头的还是书面的, 一次性的还是连续反复?)

Q9. What kind of feedback do you give?
(Prompts: marks only, marks and comments, comments only, or other?)
（请描述您给学生的反馈）

Q10. Do you use criteria when giving feedback?
（您给学生反馈时有参考标准吗？）
   Yes, how? Please specify the criteria _________________________________
   No, please give the reasons ________________________________________

Q11. Do you share the criteria with the students? （该标准学生了解吗？）
   Yes, how?
   No, why not?

Q12. How have your feedback practice changed after the implementation of the current assessment policy?
（您实际教学中是如何将形成性评估和终结性评估结合应用的？）

Section 4. About your perceptions of assessment in the area of English language learning

Q13. Please describe what you understand by the purposes of assessment.
（请描述您对评估目的的理解）

Q14. How do you understand the purpose of formative assessment?
Q15. How do you understand the relationship between formative assessment and summative assessment?
(请描述您对形成性评估与终结性评估关系的理解)
Q16. Please describe your understanding of teachers’ role in the assessment practice?
(你如何理解教师在其中的作用?)
Q17. Please describe what you understand by effective assessment.
(请描述您心目中好的评估)

Section 5. About your perceptions of EFL teaching and teacher’s role
Q17. How do you think English proficiency can be best obtained?
(在你看来学好英语最好的方法或途径是什么?)
Q18. How do you understand the role of teacher in students’ English proficiency acquisition?
(你如何理解教师在学生英语学习中所起的作用?)

Section 6 Stimulated verbal recall in relation to classroom observation
- Would you do it the same way before the current CE assessment policy?
- Can you explain what you doing this for?
- Do you do it as required or as you think the best?
Appendix C
Student Interview Schedule

Section 1. About you
- Your gender, age
- Where are you from? Is that a developed region or …?
- Can you tell us about the high school you went to?
- something about your English language learning experience?

Section 2. About your views and responses to EFL learning and assessment
1. - What approaches do you use to learn English?
   - Please list the top three that are most useful for your English language learning?
   - How do you understand learner’s role in EFL learning?
   - How do you understand teacher’s role in EFL learning?

2. How would you describe your involvement in EFL learning activities in CE classroom?
   What are the reasons for your current practice?

3. - What approaches have ever been used to assess your EFL learning?
   - Does assessment influence your approaches to English learning usually?
   - If yes, how?
   - Please list the top three that affect your EFL learning the most.

4. How would you describe your involvement in assessment activities (such as self, peer-assessment, portfolio) in CE classroom? What are the reasons for your current practice?

5. Does assessment influence your choice of English learning materials?
   - If yes, how?
   - If no, please describe the materials do you usually choose for your EFL learning?

Section 3. Your views and responses to the current CE assessment policy
6. Please describe your understanding of the ways your College English learning is assessed? Any difference from the assessment you have ever gone through?

7. What are the purposes of College English assessment from your point of view?

8. How do you understand the criteria and procedures used to assess your College English learning?

(请描述你用来学习英语的具体方法并指出对你的英语学习最有效的三种)

(请描述你在课堂学习活动的参与情况)

(课程评价对你的英语学习方法有影响吗？如果是，怎样影响？)

(请描述你把自己在课堂评价活动的参与情况)

(课程评价对你的英语学习材料的选择有影响吗？如果是，怎样影响？)

(请描述你对现在大学英语课程评价规定的理解。)

(你怎样看待大学英语评价的目的？)

(请描述你对大学英语评价标准的理解)
9. Does assessment’s greater emphasis on the learning process affect your approaches to College English learning? If yes, how?
（大学英语评价时更重视过程会是否会影响到你的学习方法？怎样影响？）

10. How do you respond to the formative assessment practice such as self and peer-assessment in this course? Any practices different from previous ones?
（你如何应对大学英语课程的形成性评价？）

11. What kind of feedback do you usually receive from your teacher or peers?

12. Do you value feedback from teacher, peer or yourself same or differently? Why?

13. How do you act upon feedbacks from your teacher and peers?

14. How do your formative assessment responses affect your EFL learning?
（你对大学英语课程评价的应对对你的英语学习起到了怎样的作用？）

15. How do you respond to the summative assessment practice in this course? Any change from previous ones?
（你如何应对大学英语课程的终结性评价？）

16. How do your summative assessment responses affect your EFL?
（你对大学英语课程评价的应对对你的英语学习起到了怎样的作用？）

17. Do you take any external examinations (such as CET or PETS or the like)? Why or why not?
（现在学校没有硬性要求参加C E T / P E T S，你参加么？为什么？）

18. What are your views of the purposes of involving students in assessment practice?
（你怎么看学生参与课程评价的目的及意义？）

19. What are your views of the effects of involvement on your English learning?
（参与课程评价对你的英语学习有何影响？）

20. How do you think assessment can better support your English Language Learning?
Appendix D  
Classroom Observation Schedule

Location:

Time:

Session No.:

<table>
<thead>
<tr>
<th>Teacher's assessment practice</th>
<th>Assessment task and type</th>
<th>What instructions are given?</th>
<th>How instructions are given?</th>
<th>Whether students are involved?</th>
<th>How students are involved?</th>
<th>What feedback is given?</th>
<th>How feedback is given?</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Student Assessment Practice</th>
<th>Student A</th>
<th>Student B</th>
<th>Student C</th>
<th>Student D</th>
</tr>
</thead>
<tbody>
<tr>
<td>How does the student respond to the instruction given?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To what extent does the student involves him/herself in the assessment activity?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How does the student involve in the assessment activity?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How do student respond to the feedback received?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix E
### KU Teachers’ Assessment Procedures

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Tasks</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>KU-T1 (SL)</td>
<td>Quiz</td>
<td>Quizzes are used in every session to spot check if and how students do with listening materials in the textbook. 12 quizzes are scheduled in this semester.</td>
</tr>
<tr>
<td></td>
<td>Presentation</td>
<td>Presentation is used as an assessment task for speaking. Students are allowed to choose a topic of their interest and present to the class within three to four minutes. KU-T1 arranges the order and leaves time for three to four free talks each session.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Attendance is checked by counting quiz papers and transcribing quiz results onto record sheet. Two or more absences might lead to a downgrade of students’ final assessment results negatively.</td>
</tr>
<tr>
<td>KU-T2 (RWT)</td>
<td>Presentation</td>
<td>Students are divided into groups of four or five and present as a group on a given topic for 5 to 6 minutes. Each session allows one group presentation.</td>
</tr>
<tr>
<td></td>
<td>Text reading</td>
<td>Students take their turns to read text aloud in class. The purpose is to see how familiar the students are with the text. This task is abandoned later because of limited time.</td>
</tr>
<tr>
<td></td>
<td>Writing assignments</td>
<td>Writing tasks are used mostly for assignment. Several have been assigned for the semester. Topics are mainly from the institutional English proficiency test.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Checked by calling the name list sometimes.</td>
</tr>
<tr>
<td>KU-T3 (RWT)</td>
<td>Reading report</td>
<td>Students are asked to read a book each semester and write a report on it.</td>
</tr>
<tr>
<td></td>
<td>Writing task</td>
<td>Students are given a topic to write about.</td>
</tr>
<tr>
<td></td>
<td>Paragraph</td>
<td>KU-T3 chooses some paragraphs in the textbook; especially those he thinks are not easy to translate, and lets students translate.</td>
</tr>
<tr>
<td></td>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Presentation</td>
<td>Students take their turns to present on a topic of their interest.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Students are required to sign for attendance, especially when many are absent.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Mainly refers to how the students respond to teachers’ questions.</td>
</tr>
<tr>
<td>KU-T4 (RWT)</td>
<td>Presentation</td>
<td>Students do presentations on a given topic, which is usually relevant to the theme of each unit.</td>
</tr>
<tr>
<td></td>
<td>Translation</td>
<td>Students are asked to translate some paragraphs or some sentences from the text in textbooks.</td>
</tr>
<tr>
<td></td>
<td>Reading &amp; Writing task</td>
<td>Students are assigned some reading each week and asked to write reflections or summaries on the readings.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Uses classroom participation, but admits the limitations.</td>
</tr>
<tr>
<td>KU-T5 (RWT)</td>
<td>Presentation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Writing task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Checked by calling the roll, especially when not many students show up.</td>
</tr>
<tr>
<td>KU-T6 (SL)</td>
<td>Classroom participation</td>
<td>Used as a major means to assess students. Students are called once at least every session to answer questions, mainly to check their text recitation or set homework. Also used to check students’ attendance.</td>
</tr>
<tr>
<td></td>
<td>Writing task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Translation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Does not deliberately call the roll. Knows it by appointing each to answer questions.</td>
</tr>
</tbody>
</table>
Quiz for unit 9-10 (Tape 2 Side B 24:40)

You will hear a conversation between a man and a woman. The man has a health problem and the woman is trying to give him some advice. Listening to the recording and complete the chart below with the information you hear.

<table>
<thead>
<tr>
<th>The man’s symptoms (His looks, the feeling etc.)</th>
<th>The woman’s advice (what the man should do, what the woman should do)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1.</td>
</tr>
<tr>
<td>2.</td>
<td>2.</td>
</tr>
<tr>
<td>3.</td>
<td>3.</td>
</tr>
<tr>
<td>4.</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
</tr>
</tbody>
</table>

Note:
1. Maybe criteria communicated but not specifications for standards.
2. Demand of accountability indicates the strong influence of examination tradition and the administration, whereas the focus shifting to learning shows a tendency of change.
Appendix G
KU Student Quiz Paper Sample

1. 500 dollars
2. a hot and hot water
3. 5th
4. one bed room is very large and another is very small
5. two
6. one
7. her husband 47 and herself.
8. Friday afternoon
9. 4:30
10. 362

A
### Appendix H

**KU-T1's Record Sheet**

<table>
<thead>
<tr>
<th>表格</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>F</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Appendices 309
## Appendix I

### NKU Teachers’ Assessment Procedures

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Procedures</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKU-T2</td>
<td>Assignments</td>
<td>NKU-T2 assigned writing, translation or other tasks and looked at mainly the number of submitted assignment.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Mainly look at how students respond to teacher’s questions in class and how they participate in classroom activities. NKU-T2 did not note down all the details but took her impression of students’ performance into account when making records.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Never call the roll, but would make records if the student was appointed to answer questions in class.</td>
</tr>
<tr>
<td>NKU-T1</td>
<td>Presentation</td>
<td>Referred to as daily report. NKU-T1 learned it from other teachers. Students took turns to give a presentation on any topic they interested in within three to five minutes. Two for every session.</td>
</tr>
<tr>
<td></td>
<td>Translation</td>
<td>Ask students to translate key and difficult sentences in the text.</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>Not yet, but would have it in next semester.</td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>Students were asked to look up words of similar meanings in dictionary, but seldom.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Mainly look at students’ performance at the presentation task and their responses to teachers’ questions in class.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Call the roll sometimes.</td>
</tr>
<tr>
<td>NKU-T3</td>
<td>Presentation</td>
<td>Referred to as morning report. NKU-T3 learned it from other teachers. It was aimed at providing some oral practice opportunity and purposeful training. Students were free to choose whatever topic they interested in. Students’ participation was used as part of classroom participation.</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>Assign writing tasks, read and give feedback.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>The same with NKU-T1.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Call the roll sometimes.</td>
</tr>
<tr>
<td>NKU-T4</td>
<td>Presentation</td>
<td>Referred to as morning report.</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>Same with NKU-T3.</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Look at students’ responses to teachers’ questions and observe students’ attitude to see if they were happy or attentive.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Same with NKU-S3</td>
</tr>
<tr>
<td>NKU-T5</td>
<td>Presentation</td>
<td>Intended as an opportunity for students to talk. Topics appointed weekly. If students did not feel like to present, they were asked to hand in the hard copy.</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>Look at students’ responses to teachers’ questions in class</td>
</tr>
<tr>
<td></td>
<td>Classroom participation</td>
<td>Look at students’ responses to teachers’ questions in class</td>
</tr>
<tr>
<td>NKU-T6</td>
<td>Material gathering</td>
<td>Students were asked to gather good sentences or phrases for each unit from the textbook or somewhere else.</td>
</tr>
<tr>
<td></td>
<td>Group discussion</td>
<td>Students were grouped into four. A group leader and a topic were appointed for each group. The group leader coordinated group discussion after class and presented it to class when selected.</td>
</tr>
<tr>
<td></td>
<td>Listening</td>
<td>Lead students to do spot dictation and summarise main idea</td>
</tr>
<tr>
<td>NKU-T7</td>
<td>Presentation</td>
<td>Was not used this semester, but used in last semester. Students chose a topic and presented what they prepared to class.</td>
</tr>
<tr>
<td></td>
<td>Writing</td>
<td>Same with others.</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
<td>Same with others.</td>
</tr>
</tbody>
</table>
# Appendix J
## NKU-T2’s Record Sheet

2008级学生大学英语分级教学班级学生名单

<table>
<thead>
<tr>
<th>序号</th>
<th>学号</th>
<th>姓名</th>
<th>专业</th>
<th>学院</th>
<th>备注</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>2008181095</td>
<td></td>
<td>法学</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>2008181102</td>
<td></td>
<td>法学</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>2008181108</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>2008181112</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>2008181117</td>
<td></td>
<td>法学√√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>2008181122</td>
<td></td>
<td>法学√√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>07</td>
<td>2008181136</td>
<td></td>
<td>法学√√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>08</td>
<td>2008181137</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>09</td>
<td>2008181148</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2008181155</td>
<td></td>
<td>法学√√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>2008181157</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2008181158</td>
<td></td>
<td>法学√</td>
<td>法学院</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>2008331001</td>
<td></td>
<td>环境科学×</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2008331003</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2008331004</td>
<td></td>
<td>环境科学√√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2008331007</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2008331009</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2008331017</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2008331020</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2008331023</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2008331029</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>2008331035</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>2008331040</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2008331043</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>2008331044</td>
<td></td>
<td>环境科学√</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>2008331045</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>2008331047</td>
<td></td>
<td>环境科学</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>2008332001</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>2008332003</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>2008332004</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2008332007</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>2008332009</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>2008332015</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>2008332018</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>2008332019</td>
<td></td>
<td>资源环境与城乡规划管理</td>
<td>环境与资源学院</td>
<td></td>
</tr>
</tbody>
</table>
Appendix K
NKU Student Writing Assignment Sample

Are We Really Safe In Modern Society

Looking back at the changes having taken place in recent years, one of the most obvious is the amount of locks. Faced with the change, we can't help thinking deeply: Are we really safe in modern society?

Nowadays, people become more concerned about security. More and more security devices have been invented and applied to practice. Undoubtedly, with all kinds of locks installed, we can protect ourselves against crimes to some degree. However, we have locked ourselves in as well. When an emergency occurs, it may be difficult for us to escape. What's more, it's reported that the crime rate is increasing sharply. Sometimes, installing security systems is only a comfort to our hearts but doesn't ensure the safety of our bodies and possessions.

In my opinion, it's the distrust between people that makes us feel insecure. If we to some extent believe others more, maybe we'll feel more relaxed and secure.

At the same time, we take measures to defend crimes, we'd better build a harmonious atmosphere full of trust.

Excellent!
Apr. 5
Appendix L
Ethics approval from QUT Research Ethic Unit

Dear Ms Guixian Chen,

A UHREC should clearly communicate its decisions about a research proposal to the researcher and the final decision to approve or reject a proposal should be communicated to the researcher in writing. This Approval Certificate serves as your written notice that the proposal has met the requirements of the National Statement on Research Involving Human Participation and has been approved on that basis. You are therefore authorized to commence activities as outlined in your proposal application, subject to any specific and standard conditions detailed in this document.

Within this Approval Certificate are:

* Project Details
* Participant Details
* Conditions of Approval (Specific and Standard)

Researchers should report to the UHREC, via the Research Ethics Officer, events that might affect continued ethical acceptability of the project, including, but not limited to:

(a) serious or unexpected adverse effects on participants; and
(b) proposed significant changes in the conduct, the participant profile or the risks of the proposed research.

Further information regarding your ongoing obligations regarding human based research can be found via the Research Ethics website http://www.research.qut.edu.au/ethics/ or by contacting the Research Ethics Coordinator on 07 3138 2340 or ethicscontact@qut.edu.au

If any details within this Approval Certificate are incorrect please advise Research Ethics within 10 days of receipt of this certificate.

Research Ethics Officer (on behalf of the Chairperson, UHREC)

<table>
<thead>
<tr>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>13/2/2009</td>
</tr>
</tbody>
</table>

Project Details

Category of Approval: Human non-HREC
Approved Until: 6/02/2012
Approval Number: 090000092
Project Title: A study of assessment policy changes in two Chinese universities: perspectives on English as a foreign language (EFL) learning
Project Chief Investigator: Ms Guixian Chen

Other Project Staff/Students: Prof Val Klenowski, Dr Margaret Kettle, Dr Lynette May

Experiment Summary: Investigate the interactions between the assessment policy change and its implementation, and the impact on English language learning within two Chinese universities (one key and the other non-key).

Participant Details

Participants:
Deans of the College English Departments, teachers and classes

Location/s of the Work:
China

RU Report No. E801 Version 2
Appendix M
Information sheet for teacher interview participants

PARTICIPANT INFORMATION for QUT RESEARCH PROJECT

Assessment Policy Change in Relation to English Language Teaching and Learning in
China: A study of Perspectives from two universities
(Teacher individual Interview and classroom observation)

Research Team Contacts

Name: Chen Qiuxian
Position: PhD student in Centre of Learning Innovation, Faculty of Education, Queensland University of Technology (Australia)
Phone: 0011 61 (0)411692279 (In Australia), 0011 86 13934645989 (in China)
Email: qiuxian.chen@student.qut.edu.au

Description
This project is being undertaken as part of a PhD project by Chen Qiuxian. The project is funded by the China Scholarship Council (CSC) and Queensland University of Technology (QUT). The funding body will not have access to the data obtained during the project. The purpose of this project is to investigate the impact of assessment policy change as experienced by EFL teachers and students in the Chinese higher education context. It takes a case study approach and involves two cases. The project has received ethical clearance from the QUT Human Research Ethics Committee [No. 0900000092]. This information sheet describes the project. Please read it carefully before deciding whether to participate.

Participation
The research team requests your assistance because you and your class have been purposefully selected to achieve the intended purposes of this research. I have obtained approval from the school authority about conducting this research in your school. Your participation in this project is voluntary. If you do agree to participate, you can withdraw from participation at any time during the interview or observation without any penalty. Your decision to participate will in no way impact upon your current or future relationship with the school.
Your participation will involve an interview and observation of your class. The interview is concerned with your views of English language assessment and your classroom assessment practices. It will take approximately 30-35 minutes. It will be audio-recorded for the purpose of later transcription. The researcher plans to observe four periods of your classes. This will involve video-recording for the purpose of information retrieval only.

Expected benefits
This study will not benefit you directly but the anticipated results, which is about how the changes to assessment can better sustain English language learning, are intended to provide insights regarding assessment practice in English language classrooms. You will be able to have access to the findings after the completion of the PhD thesis.

Risks
There are no risks beyond normal day-to-day living associated with your participation in the interview or observation. Your identity will be protected by the use of pseudonym or code. As to the classroom observation, I will observe as a non-participant and not intervene in any way. I will come to the classroom early, set up the video recording machine before the class begins and recording from the back row to avoid distraction. The videoed image will not be publicized in any form and will be used for this research.
only.

**Confidentiality**

All comments and responses will be treated confidentially. The names of individual persons are not required in any of the responses. Neither your name nor your university’s name will be disclosed in the research report or in any publications. Your consent will be sought if they are be used for any other purposes. Transcribed data will be stored securely during the research project in a locked storage facility. E-data will be password protected, then destroyed five years later.

**Consent to Participate**

We would like to ask you to sign a written consent form (enclosed) to confirm your agreement to participate.

**Questions / further information about the project**

Please contact the researcher named above if you have any questions or if you require further information about the project.

**Concerns / complaints regarding the conduct of the project**

QUT is committed to researcher integrity and the ethical conduct of research projects. However, if you do have any concerns or complaints about the ethical conduct of the project you may contact the QUT Research Ethics Officer on 3138 2340 or ethicscontact@qut.edu.au. The Research Ethics Officer is not connected with the research project and can facilitate a resolution to your concern in an impartial manner.
Appendix N
 Consent form for teacher interview participants

CONSENT FORM for QUT RESEARCH PROJECT
(Teacher interview and classroom observation)

Statement of consent
By signing below, you are indicating that you:

• have read and understood the information document regarding this project
• have had any questions answered to your satisfaction
• understand that if you have any additional questions you can contact the researcher
• understand that you are free to withdraw at any time, without comment or penalty
• understand that you can contact the Research Ethics Officer on 3138 2340 or ethicscontact@qut.edu.au if you have concerns about the ethical conduct of the project
• agree to participate in the interview and classroom observation
• understand that the interview will be audio-recorded and the observation will be video recorded for later information retrieval

Name
________________________________________________________

Signature
________________________________________________________

Date
___________ / ___________ / ___________.

Thank you for your participation!
Appendix O

Information sheet for student participants

PARTICIPANT INFORMATION for QUT RESEARCH PROJECT

Assessment Policy Change in Relation to English Language Teaching and Learning in China: A study of Perspectives from Two Universities
(Student interview and observation)

Research Team Contacts

Name: Chen Qiuxian
Position: PhD student in Centre of Learning Innovation, Faculty of Education, Queensland University of Technology (Australia)
Phone: 0011 61 (0)411692279 (In Australia), 0011 86 13934645989 (in China)
Email: qiuxian.chen@student.qut.edu.au

Description

This project is being undertaken as part of a PhD research project by Chen Qiuxian. The purpose of this project is to investigate the impact of assessment policy change as experienced by EFL teachers and students in the Chinese higher education context. The findings aim to shed light on how assessment can support English as a Foreign Language (EFL) learning in the Chinese higher educational context. It takes a case study approach and involves two cases. The project has received ethical clearance from the QUT Human Research Ethics Committee [No.0900000092]. This information sheet describes the project. Please read it carefully before deciding whether to participate.

Participation

The researcher requests your assistance because you have been purposefully selected as one of the interviewees designed to achieve the optimal research purposes. The researcher has obtained approval from the school authority and your teacher about conducting this research in your class. Your participation in this project is voluntary. If you do agree to participate, you can withdraw at any time without penalty. Your decision to participate or withdrawal will in no way impact upon your current or future relationship with your school and your teacher. Your participation will involve a face to face interview and being observed in EFL classroom. The interview is mainly about your views of and responses to the assessment and the assessment practices in the English Language Learning classroom, and will take approximately 30-40 minutes. It will be audio-recorded for the purpose of later transcription. The classroom observations are about your involvement in assessment activities and your responses in relationship to the teacher’s assessment practice. It will be video-recorded for four sessions (60 minutes for each). All the data will be used for the purpose of this study only.

Expected benefits

It will not benefit you directly but the anticipated results of this study, which is about how the changes to assessment can better sustain English language learning, are intended to provide insights regarding assessment practice in English language classrooms. The insights from the students’ perspective can be of value for students’ English language learning and assessment experience in the Chinese higher educational context. You will be able to have access to them after the completion of the PhD thesis.

Risks

There are no risks beyond normal day-to-day living associated with your participation in the interview. Your identity will be protected in any publication of results of this study by the use of pseudonym (false name). The recorded data will not be publicised in any form.

Confidentiality
All comments and responses will be treated confidentially. The names of individual persons are not required in any of the responses. Neither your name nor your university’s name will be disclosed in the research report or in any publications. Transcribed data will be stored securely during the research project in a locked storage facility. E-data will be protected with a password and stored securely for five years and then destroyed.

**Consent to Participate**

We would like to ask you to sign a written consent form (enclosed) to confirm your agreement to participate in the interview.

**Questions / further information about the project**

Please contact the researcher named above if you have any questions to be answered or if you require further information about the project.

**Concerns / complaints regarding the conduct of the project**

QUT is committed to researcher integrity and the ethical conduct of research projects. However, if you do have any concerns or complaints about the ethical conduct of the project you may contact the QUT Research Ethics Officer on 3138 2340 or ethicscontact@qut.edu.au. The Research Ethics Officer is not connected with the research project and can facilitate a resolution to your concern in an impartial manner.
Appendices 319

Appendix P
Consent form for student interview participants

CONSENT FORM for QUT RESEARCH PROJECT

(Student interview and observation)

Statement of consent
By signing below, you are indicating that you:

• have read and understood the information document regarding this project
• have had any questions answered to your satisfaction
• understand that if you have any additional questions you can contact the researcher
• understand that you are free to withdraw at any time, without comment or penalty
• understand that you can contact the Research Ethics Officer on 3138 2340 or ethicscontact@qut.edu.au if you have concerns about the ethical conduct of the project
• agree to participate in the interview
• understand that the interview will be audio-recorded and observation will be video-recorded.

Name

__________________________________________________________

Signature

__________________________________________________________

Date

_ / _ / _

Thank you for your participation!