UNIVERSITY TEACHERS' PERSPECTIVES ON THE IMPACT OF QUALITY ASSURANCE POLICIES IN CHINESE HIGHER EDUCATION: THREE INSTITUTIONAL CASE STUDIES

Thesis submitted for the degree of
Doctor of Philosophy in Education

Shan Huang
New College

Supervisors: Dr. Hubert Ertl, Dr. Patricie Mertova
and Professor Gordon Stanley

Trinity Term, 2016

Abstract

In the light of growing concerns regarding the quality of higher education after a period of rapid expansion, in 2003, the Chinese government launched the Undergraduate Teaching Evaluation (UTE), the first nation-wide evaluation of universities. In 2008, the UTE was replaced by the Quality Project, which signalled a change in its quality assurance approach, with a move from evaluations to the issuing of awards.

In order to investigate the impact of the two national quality assurance policies, along with the impact of two long-standing internal quality assurance mechanisms employed by universities – class observation and student evaluation of teachers – on teaching, the researcher adopted policy analysis and a case study approach. Three different universities in the same region were selected as cases. Semi-structured interviews with 56 heads of department and teachers across three departments at each university were conducted. National and university policy documents, as well as interview data, were analysed thematically in the light of concepts derived from the political sciences, namely Knoepfel et al.'s (2007; 2011) framework for policy analysis and Schneider and Ingram's (1990) classification of policy tools.

University policies and interview data revealed the patterns of impact of these two quality assurance policies. Faced with the UTE inspection, universities shifted their focus from assuring the quality of teaching to achieving good results in the exercise and therefore engaging in 'game-playing'. In order to ensure good UTE results, universities even resorted to the manipulation of data. University policies resulting from the UTE required teachers to produce standardised documents and to follow particular procedures. The majority of interviewed teachers regarded the impact of the UTE with cynicism, seeing it as a waste of time and effort, an interference with academic freedom, and believing it had the effect of undermining ethics. However, some teachers reported as positive impact of the fact that the UTE helped to keep teachers disciplined, and that it provided an impetus for ensuring teaching quality.

Interview data showed that the Quality Project awards only had an impact on the award winners, who perceived the incentives on offer to be substantial. These award winners considered the acknowledgement given by the awards and the information obtained through reflection and good examples as valuable impacts. However, the Quality Project awards did not have impact on non-award winners, i.e. the majority of teachers.

This study revealed that in order to have impact the institutional practices of class observation and the student evaluation of teachers relied on a number of conditions being in place. These practices had impact when the stakes were high or when university teachers received valid and reliable feedback. Whilst high stakes were found to lead to mixed impact, valid and reliable feedback contributed to the improvement of teaching.

This study contributes to the understanding of the impact of quality assurance policies and mechanisms on teaching in Chinese higher education, an area that has not yet been the subject of significant empirical research. Covering the two major quality assurance policies in recent years, and also the institutional mechanisms teachers face,

this research was able to capture the interdependence between these policies. No such research has previously been conducted in the context of Chinese higher education. In regard to the theoretical aspect of the research, the empirical evidence that was collected and a comprehensive review of other empirical research enabled the development of an Impact Framework. The Impact Framework identified patterns in the impact of various quality assurance policy tools and revealed their inherent strengths and weaknesses. Hence, the Impact Framework can serve in the future as an important reference for policy-makers who are seeking to design and implement effective quality assurance policy tools.

This thesis is dedicated to love and happiness.

Acknowledgements

Seven plus two years is a long time in a person's life. During this time, many things happened. Many of my family members passed away, Nai Nai (grandmother on father's side), Lao Lao (grandmother on mother's side) and Lao Ba (father). As a single child, Lao Ma (mother) is the only close family member I have now. 'When are you going to graduate' was a question that friends and family members (even I) often asked. Finally, there is an answer and fortunately my mother can witness it.

To be able to reach the end of this long journey, I received generous help from many people. My supervisors and examiners, especially Dr. Hubert Ertl, Dr. Patricie Mertova, Professor Ingrid Lunt and Dr. Janette Ryan, tolerated my clumsy language and premature analyses, read my writing again and again, and offered their insightful feedback; Professor Gordon Stanley guided my transition from a Master's student to a doctoral student. Without you, my thesis would have never reached this level. My lovely colleagues and friends, especially Dr. María Teresa Flórez Petour and Dr. Chelsea Robles, offered their critical thinking and friendship, and made me feel like home in Oxford. I am also very grateful for other members of Oxford community, especially my college advisor, Mr. David Palfreyman. All of you made this university so friendly and resourceful, which kept luring me into staying longer (and it worked). I am thankful to my contacts and interviewees in China. Without you, I would not have collected valuable data that reflected the impact of quality assurance policies on the ground. Finally, immense gratitude and love go to Nai Nai, Lao Ba, Lao Ma, Lao Lao and Wang Wang for your unwavering support and care.

Contents

Abstract **2**
Acknowledgements **5**
Contents **6**
List of Tables **10**
List of Figures **12**
Acronyms **14**
Glossary **15**
1 Introduction **20**
1.1 Contextual Rationale 22
1.1.1 The Three Critical Moments *22*
1.1.2 Evaluations of Universities *24*
1.1.3 Awards for Good Teaching *26*
1.2 Personal Rationale 28
1.3 Overview of the Thesis 30
2 Literature Review **32**
2.1 Definitions of Quality in Higher Education 32
2.1.1 Definitions in the International Literature *33*
2.1.2 Definitions in the Chinese Literature *35*
2.2 Definitions of Quality Assurance in Higher Education 38
2.2.1 Definitions in the International Literature *38*
2.2.2 Definitions in the Chinese Literature *43*
2.3 Empirical Studies on the Impact of Quality Assurance Policies 46
2.3.1 International Empirical Studies *46*
2.3.2 Chinese Empirical Studies *53*
2.4 Empirical Studies on the Impact of Awards for Good Teaching 56
2.5 Empirical Studies on the Impact of Class Observation 56
2.6 Empirical Studies on the Impact of the Student Evaluation of Teachers 59
2.7 Summary 64
3 Research Design **68**
3.1 Research Questions and Research Strategy 68
3.2 From Validity and Reliability to Alternative Criteria for Evaluating Qualitative Research 72
3.3 Methods for Data Collection 75
3.4 Case Selection 79
3.4.1 Selection of the Universities *80*
3.4.2 Selection of Departments *82*
3.4.3 Selection of Interviewees *83*
3.5 The Analytical and Conceptual Frameworks 87
3.5.1 Knoepfel et al.'s Framework for Public Policy Analysis *87*
3.5.2 Schneider and Ingram's Classification of Policy Tools *95*
3.5.3 A Proposed Understanding of Impact *97*
3.6 Data Analysis 98
3.7 Research Ethics 100
3.7.1 Negotiating Access *101*
3.7.2 Anonymity *102*

3.8 Limitations....103
3.8.1 Potential Biases....103
3.8.2 Generalisability....105

4 National Policies....108
4.1 National Institutional Rules....108
4.1.1 Central Control and Efficiency....109
4.1.2 Fragmented Power in Higher Education....112
4.2 The Political Agenda Setting of the UTE....116
4.2.1 Cause for Concerns Over the Quality of Education....116
4.2.2 Political Definition of the Public Problem....118
4.3 The Policy Programming of the UTE....122
4.4 The Political Agenda Setting of the Quality Project....125
4.4.1 Social Problems....125
4.4.2 Public Problems....127
4.5 The Policy Programming of the Quality Project....128
4.6 Summary....130

5 The Project 985 University....133
5.1 The University Context....133
5.1.1 University Institutional Rules Inherited from the National Institutional Rules....134
5.1.2 University Institutional Rules Established by the University Leadership....137
5.2 Internal Quality Assurance Mechanisms....139
5.2.1 Class Observation and the Student Evaluation of Teachers....139
5.2.2 Teacher Appraisal....142
5.3 University's Reactions to the UTE and the Quality Project Awards....144
5.3.1 Agenda Setting....144
5.3.2 Programming and Implementation....148
5.4 Perceived Impact of the UTE....153
5.4.1 Department Leaders' Perspectives....153
5.4.2 Teachers' Perspectives....155
5.4.3 Summary....159
5.5 Perceived Impact of the Quality Project Awards....162
5.5.1 Department Leaders' Perspectives....162
5.5.2 Teachers' Perspectives....165
5.5.3 Summary....168
5.6 Perceived Impact of Class Observation....171
5.6.1 Department Leaders' Perspective....171
5.6.2 Teachers' Perspectives....175
5.6.3 Summary....179
5.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal....182
5.7.1 Department Leaders' Perspective....182
5.7.2 Teachers' Perspectives....190
5.7.3 Summary....194
5.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching 198
5.8.1 Problems of Current Quality Assurance Mechanisms....199
5.8.2 What Drives Teachers to Invest in Teaching?....202
5.8.3 What Helps Teachers to Improve Teaching?....205
5.9 Chapter Conclusions....208

6 The Project 211 University....211
6.1 The University Context....211
6.1.1 University Institutional Rules Inherited from the National Institutional Rules....211
6.1.2 University Institutional Rules Established by the University Leadership....214
6.2 Internal Quality Assurance Mechanisms....216

6.2.1 Class Observation and the Student Evaluation of Teachers *216*
6.2.2 Teacher Appraisal *219*
6.3 The University's Reactions to the UTE and the Quality Project Awards 222
6.3.1 Agenda Setting *222*
6.3.2 Programming and Implementation *225*
6.4 Perceived Impact of the UTE 229
6.4.1 Department Leaders' Perspectives *230*
6.4.2 Teachers' Perspectives *232*
6.4.3 Summary *235*
6.5 Perceived Impact of the Quality Project Awards 237
6.5.1 Department Leaders' Perspectives *238*
6.5.2 Teachers' Perspectives *238*
6.5.3 Summary *241*
6.6 Perceived Impact of Class Observation 245
6.6.1 Department Leaders' Perspectives *245*
6.6.2 Teachers' Perspectives *247*
6.6.3 Summary *248*
6.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal 250
6.7.1 Department Leaders' Perspectives *250*
6.7.2 Teachers' Perspectives *251*
6.7.3 Summary *254*
6.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching 257
6.8.1 Problems with Current Quality Assurance Mechanisms *257*
6.8.2 What Drives Teachers to Invest in Teaching? *261*
6.8.3 What Helps Teachers to Improve Teaching? *263*
6.9 Chapter Conclusions 264

7 City University 267
7.1 The University Context 267
7.1.1 University Institutional Rules Inherited from the National Institutional Rules *267*
7.1.2 University Institutional Rules Established by the University Leadership *270*
7.2 Internal Quality Assurance Mechanisms 272
7.2.1 Class Observation and the Student Evaluation of Teachers *272*
7.2.2 Teacher Appraisal *275*
7.3 The University's Reactions to the UTE and the Quality Project Awards 277
7.3.1 Agenda Setting *277*
7.3.2 Programming and Implementation *280*
7.4 Perceived Impact of the UTE 285
7.4.1 Department Leaders' Perspectives *285*
7.4.2 Teachers' Perspectives *287*
7.4.3 Summary *288*
7.5 Perceived Impact of the Quality Project Awards 290
7.5.1 Department Leaders' Perspectives *291*
7.5.2 Teachers' Perspectives *292*
7.5.3 Summary *293*
7.6 Perceived Impact of Class Observation 295
7.6.1 Department Leaders' Perspectives *295*
7.6.2 Teachers' Perspectives *297*
7.6.3 Summary *302*
7.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal 304
7.7.1 Department Leaders' Perspectives *304*
7.7.2 Teachers' Perspectives *307*
7.7.3 Summary *308*

7.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching 312
7.8.1 Problems with Current Quality Assurance Mechanisms *312*
7.8.2 What Drives Teachers to Invest in Teaching? *316*
7.8.3 What Helps Teachers to Improve Teaching? *317*
7.9 Chapter Conclusions 319

8 Conclusions 321
8.1 The Impact of National Quality Assurance Exercises 322
8.1.1 Implementation Actions *324*
8.1.2 Consequences and Reactions *326*
8.2 The Impact of Awards for Good Teaching 334
8.2.1 The Impact of Awards as Incentive Tools *337*
8.2.2 The Impact of Awards as Capacity-building Tools *341*
8.3 The Impact of Class Observation 343
8.3.1 The Impact of Class Observation as an Authority Tool *346*
8.3.2 The Impact of Class Observation as a Capacity-building Tool *348*
8.4 The Impact of the Student Evaluation of Teachers 351
8.4.1 The Impact of the Student Evaluation of Teachers as an Authority Tool *353*
8.4.2 The Impact of the Student Evaluation of Teachers as a Capacity-building Tool *354*
8.5 The Impact Framework 358
8.5.1 The Impact of Quality Assurance Authority Tools *360*
8.5.2 The Impact of Quality Assurance Incentive Tools *362*
8.5.3 The Impact of Quality Assurance Capacity-building Tools *363*
8.5.4 The Impact of Quality Assurance Learning Tools *364*
8.6 Policy Implications 365
8.7 Reflections on this Project 369
8.7.1 Sensitivity to Context *370*
8.7.2 Commitment and Rigour *372*
8.7.3 Transparency and Coherence *373*
8.7.4 Impact and Importance *375*

References 377

Appendices 395
Appendix 1 A List of Reviewed Policy Documents 395
Appendix 2 The Second Interview Schedule 402
Appendix 3 The Third Interview Schedule (used in main study) 404
Appendix 4 A List of Category of Themes 411
Appendix 5 CUREC Form 412
Appendix 6 Participant Information Sheet and Consent Form 419
Appendix 7 The *2004 UTE Scheme* 422
Appendix 8 The List of the *Action Plans* of the Quality Project Awards 425

List of Tables

Table 1 Concepts underpinning the narrow sense of quality assurance and quality enhancement.....39
Table 2 Concepts underpinning accountability and improvement.....40
Table 3 Summary of the impacts seen in the literature.....65
Table 4 Profiles of interviewees in main study.....86
Table 5 The increase of both undergraduate and graduate students.....119
Table 6 A comparison between the Agenda Setting, the PAP and action plans of the UTE..124
Table 7 A comparison of the Agenda Setting and the PAP of the Quality Project.....129
Table 8 The selection criteria in the application forms of the Excellent Courses Awards, the Distinguished Teacher Awards, the Teaching Groups Awards, the Bilingual Courses Awards and Ten Thousand New Textbooks Project.....130
Table 9 A summary of the UTE and the Quality Project.....132
Table 10 The numbers of full-time students and teachers and the student–teacher ratio at the Project 985 University.....135
Table 11 Teaching evaluation forms at the Project 985 University.....140
Table 12 The Project 985 University's top priorities.....146
Table 13 The total number of the Quality Project awards that the Project 985 University received by 2010.....152
Table 14 The number of awards of the Quality Project at different departments at the Project 985 University.....164
Table 15 The numbers of undergraduate and graduate students and teachers and the student–teacher ratio at the Project 211 University.....214
Table 16 Teaching evaluation forms at the Project 211 University.....218
Table 17 The Project 211 University's top priorities.....223
Table 18 The total number of the Quality Project awards that the Project 211 University received by 2010.....229
Table 19 The number of the Quality Project awards won by different departments at the Project 211 University.....237
Table 20 The numbers of undergraduate and graduate students and teachers and student–teacher ratio at City University.....269
Table 21 Teaching evaluation forms at City University.....274
Table 22 City University's top priorities.....278
Table 23 The total number of Quality Project awards that City University had received by 2010.....285

Table 24 The number of Quality Project awards won by different departments at City University *291*
Table 25 The impact of national quality assurance exercises *323*
Table 26 The impact of teaching awards as incentive and capacity-building tools *336*
Table 27 The impact of class observation as an authority tool *344*
Table 28 The impact of class observation as a capacity-building tool *345*
Table 29 The impact of student evaluations as an authority tool *352*
Table 30 The impact of student evaluations as a capacity-building tool *352*
Table 31 Impact Framework *359*
Table 32 The strengths and weaknesses of policy tools *366*
Table 33 UTE indicators *424*
Table 34 Action Plans of the Quality Project awards *426*

List of Figures

Figure 1 The continuum model of interviews.....*76*
Figure 2 Theoretical model for the analysis of a public policy.....*91*
Figure 3 A proposed understanding of impact.....*98*
Figure 4 The impact of the UTE as perceived by head interviewees at the Project 985 University.....*160*
Figure 5 The impact of the UTE as perceived by teacher interviewees at the Project 985 University.....*160*
Figure 6 The impact of the Quality Project (QP) awards as perceived by head interviewees at the Project 985 University.....*169*
Figure 7 The impact of the Quality Project (QP) awards as perceived by teacher interviewees at the Project 985 University.....*169*
Figure 8 The impact of class observation (CO) as perceived by head interviewees at the Project 985 University.....*180*
Figure 9 The impact of class observation (CO) as perceived by teacher interviewees at the Project 985 University.....*181*
Figure 10 The impact of student evaluation (SE) as perceived by head interviewees at the Project 985 University.....*196*
Figure 11 The impact of student evaluation (SE) as perceived by teacher interviewees at the Project 985 University.....*197*
Figure 12 The impact of the UTE as perceived by head interviewees at the Project 211 University.....*235*
Figure 13 The impact of the UTE as perceived by teacher interviewees at the Project 211 University.....*236*
Figure 14 The impact of the Quality Project (QP) awards as perceived by head interviewees at the Project 211 University.....*242*
Figure 15 The impact of the Quality Project (QP) awards as perceived by teacher interviewees at the Project 211 University.....*242*
Figure 16 The impact of class observation (CO) as perceived by head interviewees at the Project 211 University.....*249*
Figure 17 The impact of the class observation (CO) as perceived by teacher interviewees at the Project 211 University.....*249*
Figure 18 The impact of student evaluation (SE) as perceived by head interviewees at the Project 211 University.....*256*

Figure 19 The impact of student evaluation (SE) as perceived by teacher interviewees at the Project 211 University *256*
Figure 20 The impact of the UTE as perceived by head interviewees at City University..... *289*
Figure 21 The impact of the UTE as perceived by teacher interviewees at City University . *290*
Figure 22 The impact of the Quality Project (PQ) awards as perceived by head interviewees at City University *294*
Figure 23 The impact of the Quality Project (PQ) awards as perceived by teacher interviewees at City University *295*
Figure 24 The impact of the class observation (CO) as perceived by head interviewees at City University *302*
Figure 25 The impact of the class observation (CO) as perceived by teacher interviewees at City University *303*
Figure 26 The impact of student evaluation (SE) as perceived by head interviewees at City University *310*
Figure 27 The impact of student evaluation (SE) as perceived by teacher interviewees at City University *311*

Acronyms

CEQ	Course Experience Questionnaire
CNY	Chinese Yuan
CO	Class observation
CPC	Communist Party of China
CUREC	Central University Research Ethics Committee
DepHum	The Humanities Department
DepSc	The Sciences Department
DepSoSc	The Social Sciences Department
EQM	External quality monitoring
ERIC	Education Resources Information Centre
GBP	British Pound Sterling
GDP	Gross domestic product
HEIs	Higher education institutions
NOKUT	Norwegian Agency for Quality Assurance in Education
NSS	National Student Survey
OECD	The Organisation for Economic Co-operation and Development
PAA	Political-administrative arrangement
PAP	Political-administrative programme
PRC	The People's Republic of China
QA	Quality Assurance
QAA	Quality Assurance Agency
QP	Quality Project
SE	Student evaluation
SEEQ	Students' Evaluations of Educational Quality
SETs	Students' Evaluations of Teaching Effectiveness
Uni211	The Project 211 University
Uni985	The Project 985 University
UniCity	City University
UTE	Undergraduate Teaching Evaluation

Glossary

A		
	Action plans	'[T]he sect of planning decisions considered as necessary for the coordinated and targeted production of administrative services (outputs)' (Knoepfel et al., 2011, 206).
	Actors	Actors include political-administrative authorities, target groups, end beneficiaries and indirectly affected third parties (Knoepfel et al., 2007; 2011).
	Agenda setting	The first stage of policy cycle, in which the first policy product - political definition of the public problem is produced (Knoepfel et al., 2007; 2011).
	Authority tools	'[S]tatements backed by the legitimate authority of government that grant permission, prohibit, or require action under designated circumstances' (Schneider and Ingram, 1990, 514).These are tools that 'rely heavily on authority and assume that lower-level agents and targets are motivated by a desire to follow orders' (Schneider and Ingram, 1997, 93)
C	Capacity-building tools	'Capacity-building tools provide information, training, education and resources to enable individuals, groups, or agencies to make decisions or carry out activities. These approaches that assume incentives are not an issue, but there may be barriers stemming from lack of information, skills, or other resources needed to make decisions or take actions that will contribute to policy goals' (Schneider and Ingram, 1990, 517).
	Consensus	Agreement between different actors. One of the resources defined by Knoepfel et al. (2007; 2011).
	Consequences	Outcomes of implementation actions or the interactions which have occurred.
D	Dedication	A self-driven will to achieve the personal best.
	Disciplinary research	Research aims at advancing knowledge in a subject area.
	Dr. Fox' effect	'The "Dr. Fox" effect is defined as the overriding influence of instructor expressiveness on SETs, and has been interpreted to mean that an enthusiastic lecturer can "seduce" students into giving favorable evaluations, even though the lecture may be devoid of meaningful content (see Marsh and Dunkin, 1997;Marsh, 1987)' (Marsh, 2007, 357)

E	End beneficiaries	'End beneficiaries are actors who benefit, more or less indirectly and according to the intended objectives of the policy concerned, from the target groups' altered behaviours' (Knoepfel et al., 2007, 54).
F	Financial resource	Money. One of the resources defined by Knoepfel et al. (2007; 2011).
G	Grading leniency	Grading leniency is a belief in a positive correlation between students' grades and teachers' results in student evaluation.
	Graduate school	A status granted by the Ministry of Education.
H	Human resources	Personnel. One of the resources defined by Knoepfel et al. (2007; 2011).
I	Impact	The term 'impact' is defined by Knoepfel et al. (2011, 230) as 'all of the - desired and undesired - changes in the behaviou of target groups that are directly attributable to the entry in force of the political-administrative programmes (PAPs), political-administrative arrangements (PAAs), action plans (APs) and the formal implementation acts (outputs) that concretise them'.
	Impetus	The driver to motivate teachers to ensure teaching quality.
	Implementation actions	Actions taken by the target groups or other actors to implement quality assurance policies. These actions are designated by the quality assurance policies (more specifically by the action plans made by the universities), so they vary policy by policy.
	Impression management	Impression management means that policies are made and actions are taken with a major purpose to leave a good impression on shareholders while the assurance of quality becomes secondary (Cheng, 2008).
	Incentive tools	Incentive tools 'rely on tangible payoffs, positive or negative, to induce compliance or encourage utilization' (Schneider and Ingram, 1990, 515). These tools assume that 'people will not be motivated to take policy-relevant action unless there is something to gain or lose' (Schneider and Ingram, 1997, 94).
	Infrastructure	One of the resouces. '…all of the tangible goods or property at the disposal of the different actors' (Knoepfel, 2011, 83).
	Interactions	Actions that enable the results of the implementation actions to reach the target groups, which complete the feedback loop for the implementation actions.

	Institutional rules	Institutional rules are 'both structures and rules that are formal, explicit and generally legally formalised and as informal norms that are implicit but shared by the members of an organisation and community' (Knoepfel et al., 2007, 95).
L	Law	The legal and regulatory bases, which define the objectives and behaviours of target groups and choice of other resources. One of the resources defined by Knoepfel et al. (2007; 2011).
	Learning tools	'Learning tools are used when the basis upon which target populations might be moved to take problem-solving action is unknown or uncertain…One of the most important characteristics of learning tools is that they assume agencies and target populations can learn about behaviour, and select from the other tools those that will be effective' (Schneider and Ingram, 1990, 521).
O	Opportunity costs	Opportunity costs concern the efficient allocation of resources based on the exclusivity of the use of resources (Hay & Morris, 1991).
	Organisation	One of the resources defined by Knoepfel et al. (2011, 79) as structure formed by 'actors who fulfil the specific functions associated with the conduct of a public policy'.
P	Pedagogical research	Research aims at improving teaching and learning in any subject area in order to advance the understanding of teaching and learning.
	Passion for teaching	The enjoyment, a feel of happiness and a sense of satisfaction obtained through teaching.
	Political-administrative arrangement (PAA)	'[T]he PAA designate[s] the competent authorities and administrative services (public actors) as well as all of the other institutional rules specific to the implementation of a policy' (Knoepfel et al., 2011, 165).
	Political-administrative authorities	Policy-makers (Knoepfel et al., 2007; 2011).
	Political-administrative programme (PAP)	'The provisions of a PAP comprise the normative objectives of the solution envisaged for the resolution of the problem, the definition of target groups and their roles during the realisation of the policy (hypothesis of causality), the means made available for this purpose (intervention hypothesis) and the principles of the administrative organisation of the policy implementation' (Knoepfel, et al. 2007, 152). The PAP should have five constituent elements: objectives, evaluative elements,

		operational elements (intervention instruments), the PPA and resources and procedural elements (Knoepfel et al., 2007; 2011).
	Political definition of the public problem	The process in which social problems emerge from problematic private situations and public problems emerge from social problems (Knoepfel et al., 2007; 2011).
	Political support	Support from the majority. One of the resources defined by Knoepfel et al. (2007; 2011).
	Programming	The second stage of policy cycle, in which the second policy product - political-administrative programme (PAP) and the third policy product - political administrative-arrangement (PAA) are produced (Knoepfel et al., 2007; 2011).
	Prospective purposes of incentive tools	Prospective functions are to draw interest in teaching and motivate teachers to learn from the award winners.
	Public problems	When these social problems receive political attention but no public policy is adopted or implemented, these social problems become public problems (Knoepfel et al., 2007; 2011).
Q	Quality assurance authority tools	Quality assurance authority tools are compulsory quality assurance policies with high stakes, which target groups cannot ignore.
	Quality assurance capacity-building tools	Quality assurance capacity-building tools are quality assurance policies that help target groups to build their capacities in order to improve.
	Quality assurance incentive tools	Quality assurance incentive tools are quality assurance policies that use financial or honorary incentives to motivate target groups to change their behaviours.
	Quality assurance learning tools	Quality assurance learning tools do not dictate target groups' behaviours. When employing learning tools, political-administrative authorities, target groups and other actors assume that they do not know how to assure quality, so they have to explore and learn together.
R	Reactions	Actions taken by the target groups in response to the consequences.
	Reliability	Reliability refers to the reliability of the judges of quality assurance policies and mechanisms. This consideration entails

		asking whether the judges are able to make fair judgments about target groups' performance in regard to quality.
	Resources	Resources include legal resources, human resources, financial resources, information, organisation, consensus, time, infrastructure, political support and force (Knoepfel et al., 2007; 2011).
	Responsibility	A sense of fulfilment of duties and a will to meet others' expectations – not letting others down.
	Retrospective functions of incentive tools	Retrospective functions are to recognise pervious achievements by offering awards as honour and thereby to raise the status of award winners.
	Ritualism	Ritualism is designated to a phenomenon which policies and procedures are made and implemented as rituals to symbolise the accomplishment of quality assurance tasks, while the essence of quality is missing (Newton, 1996b).
S	Social problems	When these private problems receive public attention but are not put on the political agenda, they become social problems (Knoepfel et al., 2007; 2011).
	State Council	The Chinese centre government
T	Target groups	'The target groups are made up of people (individual or legal entities) and organisations whose behaviour is required to change' (Knoepfel et al., 2007, 53).
	Teachers	A Chinese term for academics
	Teaching Group	It is a term introduced by the Quality Project. It is designated to a group of teachers who filed an application together for the Excellent Teaching Group Award.
	Time	One of the resources defined by Knoepfel et al. (2007; 2011).
V	Validity	Validity concerns the question of whether the indicators employed by the quality assurance or mechanisms serve as the best proxies for quality.
W	Workload/ difficulty	Workload/difficulty is another popular concern among academics who believe there is a negative correlation between workload/difficulty and their results in student evaluation.

1 Introduction

This research examines the impact of quality assurance policies and mechanisms on individual teachers[1] in relation to the quality of teaching in the context of Chinese higher education. In order to address this issue, the national quality assurance policies and their influence on universities' policies were investigated before an exploration of their impact on individual teachers. In recent years, the Chinese government has focused on two major quality assurance policies: the Undergraduate Teaching Evaluation (UTE) and the Quality Project (an analysis of the two national policies is conducted in Chapter Four). The UTE is a nationwide evaluation of universities conducted by the Ministry of Education. The Quality Project is an umbrella grouping of various awards or special titles to reward good teaching practice granted by the Ministry of Education. Awards are a quality assurance mechanism employed by the Quality Project. Besides awards, the impact of other quality assurance mechanisms commonly employed by Chinese universities, i.e. class observation and the student evaluation of teachers, was also explored.

In order to research the impact of quality assurance policies and mechanisms on university policies and teachers, this study adopts a multiple case study approach and selects three different universities in the same region of China as cases (for details on the case selection, see Chapter Three). The three universities are different in their prestige levels. In mainland China (hereinafter referred to as China), universities' prestige is determined by their official affiliations to certain 'projects'. The projects are 'Project 211' and 'Project 985'. Project 211 selects around 100 universities in China and provides them with extra funding to allow them to maintain their privileges

1 In China, academics who work in higher education are referred as 'teachers', despite in most cases conducting both teaching and research.

by attracting high-achieving academics with higher salaries, more research grants and better equipment (CPC Central Committee and State Council, 1993). Project 985 selects the 39 best universities from among the universities affiliated to Project 211 and provides them with very generous funding to foster their world-class status (Ministry of Education, 1998). The first case study university is included in both Project 211 and Project 985; it is referred to as the Project 985 University in this thesis. The second university only participates in Project 211, and so is called the Project 211 University. The third university is affiliated neither to Project 211 nor Project 985; it is named City University. The universities are named this way to protect their anonymity.

Within each case study university, the existing internal quality assurance mechanisms along with the university's responses to the national quality assurance policies were examined (see Chapters Five, Six and Seven). Besides explaining the three case study universities' policies and reactions to national policies, the three data analysis chapters were largely devoted to the impact of both national and university quality assurance policies and mechanisms on teachers. Evidence of the perceived impact was collected through interviews with 56 department heads, professors, associate professors and lecturers from nine departments at the three case study universities. The empirically identified impact allows the researcher to propose an Impact Framework to articulate what the impacts are.

The rest of this chapter will explain the rationales for conducting this study. These are divided into contextual and personal rationales. The contextual rationale explains why there is a need to conduct research into the impact of quality assurance policies in

China by reviewing the development of quality assurance policies over the past few decades. The personal rationale explains the researcher's personal interest and suitability for conducting this research. This chapter concludes with an overview of the thesis.

1.1 Contextual Rationale

The quality assurance approaches adopted by the Chinese government are limited to evaluations of universities and issuing of awards for good teaching; these two tools have been repeatedly used throughout history. Before introducing the history of quality assurance policies, some context will be provided to facilitate readers' understanding. In particular, the focus will be on three critical moments that have occurred since the establishment of the People's Republic of China (PRC) in 1949, which have had dramatic influence on Chinese higher education.

1.1.1 The Three Critical Moments

The first of the three critical moments was the end of the Cultural Revolution, which took place from 1966 to 1976, and China was in chaos during that period. In terms of higher education, there was no university entry examination available to recruit students and students could not pursue their education in universities either. In 1977, however, university entry examinations were reinstated by the government and students could enrol at university again (Ministry of Education, 1977). Since then, universities have resumed normal operations.

The second critical moment was the transformation of the Chinese economy, in 1992, when the government decided to transform the Chinese economy from a planned to a market economy. Following the educational ideology of the US, the most powerful country in the world at that time, the government embraced the idea of 'general

education'. It wanted to produce well-rounded graduates to compete in the market economy. Hence, the government moved from a specialised university model, borrowed from the Soviet Union, to a western model, the comprehensive university model, by merging several specialised universities to create one comprehensive university.

The third critical moment was the launch of the policy to increase participation in higher education. From 1998 to 2008, the Chinese higher education sector experienced a rapid massification. The higher education enrolment rate jumped from 9.8% to 23.3% and student numbers in universities and colleges increased from 3,408,700 to 20,210,200. The number of universities and colleges grew from 1,022 to 2,263 (Zheng, 2010). Based on data published by the Ministry of Education, Li and Li (2010) conclude that the total number of universities reached 768, including 48 private universities, among which 202 public and 47 private universities were established between 1998 and 2009. However, such a rapid expansion of higher education caused many problems, especially a shortage of physical and human resources. According to Zeng and Zhang (2010), the government budget for teaching decreased from ¥2,921 Chinese Yuan (CNY) per student in 2000 to ¥2,352 CNY in 2003; the average campus area per student was reduced by 22 square metres from 1998 to 2003; the average student–teacher ratio increased from 12:1 in 1998 to 21:1 in 2003.

Against this background, a series of quality assurance policies were developed. The approaches adopted by these policies were evaluations of universities and issuing of

awards for rewarding good teaching practice. The history of the development of university evaluations and awards is outlined below.

1.1.2 Evaluations of Universities

The development of evaluations is addressed here and includes a distinction between the early evaluations and a more recent evaluation initiative, the UTE, in 2000s. The results of early evaluations were not publicly available, and the process of evaluations was not transparent and only covered some universities. The UTE was introduced in 2003 and was the first nationwide evaluation of universities. The evaluation processes were much more transparent and the results were published. The development of evaluations reflects the historical changes in China, which were briefly outlined in the previous section.

In 1985, eight years after the end of the Cultural Revolution, the State Education Commission,[2] which had become concerned about the quality of the recently re-established higher education system, launched an evaluation of universities which specialised in engineering to decide which of them should be offered more funding and which should be closed (State Education Commission, 1985). Starting with this evaluation, undergraduate teaching became the focus of institutional evaluations (Zhou, Zhang and Chen, 2010).

From 1992, as subject-specific and/or smaller universities were being absorbed into larger ones, the higher education system in China went through an enormous change. Around 900 higher education institutions were merged from 1992 to 2002 (Wang and Liu, 2009). Against this background, three evaluations tailored for universities with

2 The equivalent of the present Ministry of Education at that time

different prestige levels began. 192 universities that were set up after 1976 went through a 'conformity evaluation' from 1994 to 2002 with the result that universities which failed the evaluation would be closed. Sixteen universities that had a long history, i.e. had been founded by previous governments, and were highly regarded were evaluated according to an 'excellence evaluation' between 1996 and 2000. Twenty-six universities in the middle of the institutional spectrum, which were regarded as average, were selected for 'random evaluation' from 1999 to 2001. The results of the three evaluations were intended to guide government decisions on whether permission should be grounded to the universities to start new undergraduate, Master's or doctoral programmes (Zhou, Zhang and Chen, 2010).

In 2003, to address the consequences of the rapid massification of higher education, the central government launched the first round of the UTE. During this round, 589 universities were evaluated between 2003 and 2008. The majority of these were public universities that were founded before 1998. In addition, there were 68 universities founded after 1998, including three private universities (Li and Li, 2010). As it was the first time that the evaluation results were published, this evaluation had a dramatic impact on Chinese higher education (Hu, 2008).

Nine empirical studies have researched the impact of the UTE (Gao et al., 2006; Lee, Huang and Zhong, 2012; Liu, 2011; Liu, 2013; Li, Li and Qu, 2006; Qu and Li, 2008; Zhang and Xue, 2009; Zhang and Zhang, 2008a; Zhang and Zhang, 2008b), and will be reviewed in greater detail in Chapter Two. In these studies, there is a distinction between the findings of research initiated by academics and those commissioned by the Ministry of Education. On the one hand, Li, Li and Qu (2006) and Qu and Li's

(2008) studies were commissioned by the Ministry of Education and claimed a generally positive impact of the UTE on teaching. On the other hand, research initiated by academics was dominated by perceptions of negative impact.

Although there have been nine empirical studies on the impact of the UTE, the number of empirical studies on the impact of general evaluation is still very small, a situation which, the author believes, is disproportional to the importance of evaluations in quality assurance. By conducting this study, the researcher aims not only to contribute another piece of empirical research but also to offer an explanation for the discrepancies between the existing empirical studies. Furthermore, none of the existing empirical studies have focused on the impact of the UTE on teaching. As the Ministry of Education (2001) claims that the ultimate goal of the UTE is to improve the quality of teaching, the impact on teaching quality ought to be the focus of research on quality assurance. Hence, this study intends to address this deficiency.

1.1.3 Awards for Good Teaching

In 1989, the State Council and the Ministry of Education granted the first National Higher Education Teaching Achievement Awards. Subsequently, the awards were granted in a four-year cycle in 1993, 1997, 2001, 2005 and 2009. In 2003, the year that the UTE was launched, the Ministry of Education introduced two more awards for teaching: the Higher Education Distinguished Teacher Awards (Ministry of Education, 2003a) and the Excellent Course Awards (Ministry of Education, 2003b). The two awards shared one common purpose: to encourage professors to teach undergraduates. The Excellent Course Award was the only award that gave winners financial rewards to support them to build websites and put their courses online. All winners were required to share their courses online, because another main purpose of

the Excellent Course Awards was to promote the design of courses and the sharing of the resulting outcomes.

In 2007, the Quality Project was formally launched. The Distinguished Teacher Awards and the Excellent Course Awards were included in and reinforced by the Quality Project. Furthermore, the Quality Project established more awards, such as the Excellent Teaching Group Awards and the Excellent Textbook Awards (Ministry of Education, 2007). The 'teaching group' was a new term created by the Quality Project. To form a teaching group, group members had to be teachers who teach the same course at the same department. In theory, group members had to collaborate in designing and delivering a course. In practice, group members could be teachers who filed one application together for the Excellent Teaching Group Award. Although the Excellent Textbook Awards were new awards, in practice they overlapped with the existing awards granted by the Ten Thousand New Textbooks Project to reward high quality textbooks identified by the Ministry of Education.

There is a lack of empirical research on the impact of the various awards. The only empirical study that the researcher has found was conducted by Zhou (2009). She tested the accessibility of courses that received the Excellent Course Awards in the field of education and showed that a third of the online courses' video materials were not accessible at the time of her study. When such a large number of the award-winning courses were unable to be accessed online, fulfilling the purpose of sharing teaching resources would be problematic.

Most publications in the existing literature are discursive articles based on analyses of award criteria. The most common problem identified by this literature is the validity of award criteria. After analysing the alterations in the criteria of Teaching Achievement Awards, Lan (2006) concluded that the criteria were biased towards academics who also took on management roles or even towards administrative staff who did not teach. Similarly, Wang (2010) found that, instead of teaching, research performance became the decisive factor in receiving the Distinguished Teacher Awards. Research-related indicators accounted for 45% of the total weight, while teaching-related indicators accounted for only 30%.

Xie (2010) argued that the Excellent Teaching Group Awards were unable to fulfil their goals because the incentives were not strong enough to attract high-achieving professors to remain dedicated to the teaching groups and, as external drivers, the awards were thus unable to establish meaningful collaborations between teachers.

Questions then arise as to what the empirical data suggest. Do teachers think that the awards are able to promote the sharing of teaching materials, to reward teaching or to promote collaborations? By collecting empirical evidence, this research aims to investigate what impact these awards had on the ground.

1.2 Personal Rationale

Besides the contextual rationale, there are also personal motivations for undertaking this study. I experienced the UTE twice at two different universities, once as a student and once as a teacher. My experience of undergoing quality assurance exercises and, more precisely, a dissatisfaction with the existing policy, motivated me to conduct research to explore what changes could be made to enable quality assurance systems

to better contribute to raising the quality of education. Both experiences featured ritualism and impression management.

I first encountered the UTE when I was an undergraduate student. We were handed booklets with answers to potential questions that might be asked by evaluators. A series of meetings was held to remind us of the importance of the UTE and to warn us about the consequences of misbehaving during the inspection. The whole campus was refurbished and decorated to impress the UTE evaluators. For example, the front gate of the main campus was redesigned and reconstructed. The external walls of buildings were repainted or retiled. New trees were planted along the main avenue. Even small details were looked at. For instance, during the UTE inspection, hand-wash liquids, which were luxuries at that time, were provided in the toilets. However, after the UTE inspection, the special efforts made for the UTE stopped with the disappearance of the hand-wash.

My second encounter with the UTE happened after I started lecturing in another university. Similar preparations were made. All the campuses were spruced up and decorated and a series of meetings again held to stress the importance of the UTE. At my department, all teachers were asked to work at weekends, checking the examination papers for the previous three years. We had to check whether the total scores were calculated correctly. If there were mistakes, we had to correct them. We had to fill in additional forms and re-bind all examination papers. In the process, we found many unexpected mistakes in calculating the total scores. However, once the UTE was finished, neither the university nor the department took measures to prevent similar mistakes from happening again.

When looking at the newly refurbished campuses, I was not pleased. Instead, I started to ask questions. Was this the only difference that the quality assurance policy was able to make? Why were these quality assurance policies unable to make a substantial contribution to the quality of education? After I had read many commentaries written by teachers at other universities to reflect on the practices of their universities, I found that many frontline teachers shared similar negative attitudes and doubts about the UTE. On the other hands, many government officials on the other hand often restated the position that the UTE made various contributions and facilitated the improvement of quality. This sharp contrast made me uneasy. It provoked my interest in quality assurance and subsequently a strong desire to investigate the real impact of quality assurance policy in the Chinese context.

In order to find a solution to improve the Chinese quality assurance policies by learning from international experience, I devoted my first Master's degree to research on the quality assurance schemes in China, England and the US (Huang, 2007). Building on my previous research and personal experiences, the current doctoral thesis investigates the impact of these policies on the ground. By collecting empirical evidence, I hope that this study can make a contribution to further our understanding of what impact different quality assurance policies can have and why. Moreover, the aim of this study is to improve future quality assurance policies by learning from experience.

1.3 Overview of the Thesis

This thesis is divided into eight chapters. Following this chapter is the Literature Review. It discusses definitions of quality and quality assurance and reviews

empirical studies on the impact of quality assurance policies and mechanisms on teaching practice. Chapter Three: Research Design includes research questions, analytical and conceptual frameworks and the methodology, followed by a discussion of research ethics and limitations. Chapter Four presents the national quality assurance policies. The two national quality assurance policies focused on are the UTE and the Quality Project. Chapters Five to Seven articulate the impact of national policies on the three case study universities and their teachers. The final chapter concludes by comparing and contrasting the impact of quality assurance policies and mechanisms at the three case study universities and putting forward the Impact Framework. The thesis closes with a reflection on the project.

2 Literature Review

This chapter reviews the existing empirical research on the impact of quality assurance policies and mechanisms. However, it is necessary to clarify two key concepts: the definitions of quality and quality assurance in the international and Chinese literature. Defining these two concepts is more than just a matter of terminology. The ways in which quality and quality assurance are defined will have important practical implications. How quality is defined can determine the criteria against which quality should be assessed; how quality assurance is defined can determine what purposes the quality assurance policy should fulfil. These concepts are not straightforward. Different understandings coexist through the perspectives of different actors, which subsequently account for the discrepancies between the intended outcomes (see Chapter Four) and the impact on the ground (Chapters Five to Seven).

This chapter consists of three parts. The first two parts focus on the definitions of quality and quality assurance. The third part reviews empirical studies on the impact of quality assurance policies and mechanisms on staff in order to identify the contribution that this study could make.

2.1 Definitions of Quality in Higher Education

The definitions of quality discussed in the international literature are presented first, followed by definitions found in the Chinese literature.

2.1.1 Definitions in the International Literature

Quality is a contested concept (Barnett, 1990), and as Cheng (2008, 45) concludes, 'it means different things to different people in different circumstances'. The most cited definition of quality in higher education is the one offered by Harvey and Green (1993). The strength of Harvey and Green's definition is that it provides an umbrella that covers all the contested concepts. According to them, 'quality can be viewed as exception, as perfection, as fitness for purpose, as value for money and as transformative' (Harvey and Green, 1993, 9). Each of the notions encompasses several sub-notions, which will be elaborated below.

The notion of 'quality as exceptional' comprises three sub-concepts: the traditional notion of quality, exceeding high standards, and checking standards. The traditional notion of quality regards quality as self-evident, so there is no need or way to assess quality. Quality is associated with exclusive distinctiveness. When quality is viewed as exceeding high standards, these standards are used to assess quality. High standards are achieved through excelling in input and output, i.e. excellent students, staff and resources, which then result in excellent outcomes. When high standards are lowered to minimum standards, quality becomes attainable. Quality can be checked by an external agency against an absolute benchmark, i.e. as conformation to standards. This allows the comparison between a range of products or services.

When quality is viewed as 'perfection and consistency', quality is achieved by conforming to specifications. Unlike standards, specifications are measurable. Beyond input and output, the focus is extended to every step of a process in order to assure 'zero defects'.

‘Quality as fitness for purpose’ judges the quality of universities on whether universities fulfil their purposes. These purposes can be specified by their customers’ requirements. Harvey and Green (1993) point out the two problems underpinning ‘fitness for purpose’. First, all the stakeholders (i.e. students, government, employers and academics within higher education) can be regarded as customers. Due to the difference in their interests, different customers can have different or even conflicting requirements (the difference between academics’ and students’ perceptions of quality can be illustrated by Cheng’s (2011) study). Second, even given the opportunity, not all customers, particularly students, are capable of articulating what they expect from their university (Elton, 1992b). Therefore, Harvey and Green (1993) contend that in practice the ‘fitness for purpose’ is benchmarked against the objectives or missions identified by universities themselves – not by customers’ requirements.

‘Quality as value for money’ emphasises efficiency or effectiveness. This definition is popular among governments, who would like to have as many graduates as possible while investing as little as possible. Under this circumstance, the inconsistent education standards across different institutions are justifiable. Because the tuition fees charged by different universities vary, so the education offered by different education providers can vary also.

‘Quality as transformation’ defines quality against the institution’s influence on students. Quality is achieved by enhancing and empowering students. Enhancing students can be captured by the value-added – how much value education adds to students. Empowering students involves students in decision-making; it allows

students to monitor if the minimum standards are achieved, gives students control over their learning, and develops their critical thinking.

Different definitions suggest different approaches to the assessment of quality. When quality is defined as the traditional notion of quality, there is simply no need to assess quality because it is self-evident. When quality is defined as exceeding high standards and checking standards, quality is benchmarked against standards. When quality is defined as perfection and consistency, quality is checked against specifications. When quality is defined as fitness for purpose, quality is evaluated by the fulfilment of the purposes. When quality is defined as value for money, quality is measured by efficiency, i.e. the inputs against the outputs. When quality is defined as transformation, quality can be assessed by the value-added and other measures of personal development.

2.1.2 Definitions in the Chinese Literature

The definitions of quality discussed in the Chinese literature fall broadly into the categories put forward by Harvey and Green (1993). Sharing a similar view with Harvey and Green's (1993) traditional notion of quality, An (1997), Tian (2005) and Yang (2010) believe that that quality is such an intangible concept which cannot be captured by words.

In fact, the most popular definition of quality among Chinese academics is 'quality as fitness for purpose'. Ha (2009), Han (2003) and Yang (2010) advocate evaluating the quality of higher education institutions by the fulfilment of purposes. In terms of who determines the purposes, Han (2003), Yan (2011) and Yang (2010) propose that this

is students; Gu (1998), Qi and Chen (1997), Yan (2011) and Yang (2010) suggest universities; Han (2003), Huang (2002) and Qi (2002) indicate that this is the market, including students and employers; Qi and Chen (1997) place the government at the foreground; and Huang (2002), Qi (2002) and Qi and Chen (1997) include all these actors and call it 'society'.

Many Chinese academics associate 'quality as perfection and consistency' with 'quality as value for money'. They regard education as a product (An, 1997; Ha, 2009; Han, 2003; Yan, 2011) or a service (Ha, 2009; Liu, 2002; Yan, 2011) and argue that it is necessary to ensure its effectiveness or efficiency (An, 1997; Yan, 2011). Hence, specifications should be constructed to assess quality (An, 1997; Han, 2003).

The definition of 'quality as transformation' is shared by some Chinese academics. In agreement with the notion of 'enhancement', a sub-concept of 'transformation' defined by Harvey and Green (1993), Qi and Chen (1997), Qian and An (2006) and Yan (2011) advocate the well-rounded development of students. In addition, the notion of 'empowerment' – the other sub-concept of 'transformation' – is mentioned but with a narrower perspective. Only Huang (2002) argues that students should be given the freedom to tailor education provision to meet their needs, but none of the Chinese scholars mentioned student involvement in decision-making and monitoring of education standards. Furthermore, Harvey and Green (1993) and Chinese academics' definitions differ in the scope of the transformation. Beyond the transformation of students, Qian and An (2006) and Yan (2011) add the transformation of the whole society.

There are some definitions of quality in the Chinese literature that are not categorised by Harvey and Green (1993). Yan (2011) and Yang (2010) define quality as distinguishing characteristics, contending that quality means the unique characteristics that make a university different from its counterparts. One definition of quality is limited to the quality of research (Huang, 2002; Qi, 2002; Yan, 2011), while another definition confines quality to the quality of teaching (Yan, 2011). Han (2003) and Yan (2011) advocate a holistic approach to quality, maintaining that quality in higher education should cover all aspects of education.

After reviewing the definitions of quality discussed in the Chinese literature, two features call for further attention. Firstly, there is an emphasis on education as a product or a service. This can be attributed to the second critical moment in recent Chinese history – the transformation of the Chinese economy (see Chapter One). When the Chinese economy was transformed from the planned to the market economy, market-orientated concepts, including market-orientated definitions of quality, gained popularity.

Secondly, many Chinese academics advocate ‘quality as fitness for purpose’, which benchmarks quality against multiple standards rather than a consistent standard. This could result from the fact that the central government often uses the same set of criteria to judge all universities in China, such as the UTE and the Quality Project awards (see Chapter Four). Using a unified standard disadvantages less prestigious universities. Thus, there is a call for multiple standards (Lee, Huang and Zhong, 2012; Yang, 2010).

2.2 Definitions of Quality Assurance in Higher Education

The definitions of quality assurance discussed in the international literature are presented first, followed by the definitions to be found in the Chinese literature.

2.2.1 Definitions in the International Literature

In the western literature, the term 'quality assurance' is used in both a narrow and a broader sense. The narrow sense of quality assurance is linked to accountability. When quality assurance is discussed in the broader sense, it includes both accountability and enhancement of quality.

Regarding quality assurance in the narrow sense, Teelken and Lomas (2009) draw a distinction between quality assurance and quality enhancement and believe that the two concepts encompass two completely different ideas:

> Emphasis on assurance is an indicator of government agencies' desire to ensure that higher education institutions (HEIs) are accountable for their public funding and their services are fit for purpose. Indeed, it has been argued that the growth of managerialism and the audit culture are signs of managerial distrust of academic staff (Morley, 2002; Newton, 2002; Shore and Wright, 1999). On the other hand, emphasis on enhancement suggests more of a transformative approach to quality which allows lecturers a greater level of autonomy with less interference from colleagues within and government agencies outside the university.
>
> (Teelken and Lomas, 2009, 260)

According to Teelken and Lomas (2009), the concepts underpinning quality assurance are audit, compliance, accountability, managerialism, fitness for purpose and value for money, whereas the concepts underpinning quality enhancement are transformation and autonomy.

In terms of enhancement, Elton (1992a) defines enhancement as empowerment, enthusiasm, expertise and excellence. Table 1 summarises the contesting concepts

underpinning quality assurance and quality enhancement, as defined by Teelken and Lomas (2009) and Elton (1992a).

Quality Assurance	Quality Enhancement
Audit	Transformation
Compliance	Autonomy
Accountability	Empowerment
Managerialism	Enthusiasm
Fitness for purpose	Expertise
Value for money	Excellence

Table 1 Concepts underpinning the narrow sense of quality assurance and quality enhancement

However, some authors use the term 'quality assurance' in a broader sense, which includes quality enhancement. Ratcliff (2003, 118) maintains that 'by extension, the term "quality assurance" refers to the policies, systems, and processes designed to ensure the maintenance and enhancement of quality within a programme or institution'. Also, Vroeijenstijn (1995, xviii) defines quality assurance as the 'systematic, structured and continuous attention to quality in terms of quality maintenance and quality improvement'. Hodson and Thomas (2003, 376) comment that 'this definition captures the dual nature of quality assurance'.

When the term 'quality assurance' is used in its narrow sense, tension is caused by the opposing factors of quality assurance and quality enhancement. When the term 'quality assurance' is used in its broader sense, tension is caused by the contest between accountability and improvement. Recognising the tension between accountability and improvement, Harvey and Knight (1996) distinguish them as follows:

> Accountability is about value for money and fitness for purpose and sees quality in those terms… Continuous improvement in teaching and learning is about enhancement of the student experience, empowering students as life-long learners.
>
> (Harvey and Knight, 1996, 97-100)

They also elaborate on the concepts associated with accountability and improvement. They associate managerialism, effectiveness and efficiency with accountability, while associating collegialism, autonomy, teaching and transformative learning with improvement. These concepts are listed in Table 2.

Accountability	Improvement
Value for money	Empowerment
Fitness for purpose	Collegialism
Managerialism	Autonomy
Effectiveness	Teaching
Efficiency	Transformative learning

Table 2 Concepts underpinning accountability and improvement

Comparing Table 1 and Table 2 shows that the concepts underpinning accountability are similar to those associated with the narrow sense of quality assurance. Academics also use the narrow sense of quality assurance interchangeably with accountability (Brown, 2004; Ratcliff, 2003), and in the same way they interchange enhancement and improvement.

While some practitioners believe quality assurance initiatives hamper the development of quality enhancement, others think quality enhancement can develop alongside quality assurance (Filippakou and Tapper, 2008; Filippakou, 2009). Harvey and Knight (1996) suggest that what has started happening across the world is the convergence of accountability and improvement in the external quality monitoring (EQM) system, an external quality assurance mechanism. Brown (2004) proposes solutions to integrate enhancement into accountability.

In fact, the narrow and the broader senses of quality assurance reflect two different perspectives and two underlying assumptions. When quality assurance is discussed in

a narrow sense, quality assurance and quality enhancement are two separate concepts. The underlying assumption is that quality assurance and quality enhancement cannot facilitate each other's development, as there is an intrinsic tension between them. On the other hand, when quality assurance is used in a broader sense, enhancement, by nature, is a part of quality assurance. The underlying assumption is that quality assurance initiatives should promote both accountability and enhancement and that this is achievable.

However, even when it is agreed that accountability and enhancement can develop alongside each other, there is an argument as to which one comes first. Harvey and Knight (1996, 100) oppose the idea that 'requiring accountability… will lead to review of practices, which in turn will result in improvement' for the following three reasons:

> First, it is likely that faced with a monitoring system that demands accountability, academics will comply with requirements as in such a way as to minimize disruption to their existing academic practices.
> Second…accountability systems, in short, are unlikely to lead to a process of *continuous* quality improvement.
> Third, accountability approaches tend to demotivate staff who are already involved in innovation and quality initiatives.
>
> (Harvey and Knight, 1996, 100)

Consequently, they believe that improvement comes first and that it will lead to the fulfilment of accountability.

The tension between quality assurance and quality enhancement or between accountability and improvement often results in an emphasis on quality assurance (in the narrow sense) or accountability and an underdevelopment of enhancement in policy and practice. The first obstacle to the balanced development of quality assurance (in the narrow sense) or accountability and quality enhancement is the

absence of quality enhancement factors at policy level. Although agreeing with Vroeijenstijn (1995) that quality assurance has a dual nature of quality maintenance and quality improvement, Hodson and Thomas (2003, 376) note that '… pressures at the end of the second millennium tended to focus attention on those elements of the definition that emphasised fitness of purpose rather than continuous enhancement'. Similarly, Filippakou and Tapper (2008, 92) suggest that 'quality enhancement was always an element of the quality assurance methodology, but the recent change in emphasis is marked'. However, more and more external quality assurance bodies have noticed the absence of enhancement and have started to include enhancement in their policies. For example, in the early 2000s the Quality Assurance Agency (QAA) in the UK claimed its policy has started moving towards enhancement (McGhee, 2003).

However, even when enhancement has been taken into account at policy level, Filippakou and Tapper (2008) question whether the institutionalisation of quality enhancement at the policy level can promote quality enhancement in practice. When external bodies reinforce quality enhancement, the external bodies may try to put into effect a prescriptive quality enhancement model, which may defeat the whole purpose of quality enhancement as a process that is, by definition, non-mandatory and that should be shaped by the actual participants involved in the teaching and learning process. Research also shows that the QAA's policy has not promoted enhancement but rather has reinforced accountability. In his research at two business schools in modern UK universities, Hoecht (2006) found that the informal quality management systems of the early 1990s, which were based on trust, have developed into highly prescriptive control mechanisms. Thus, even when enhancement has been taken into

account in terms of a relevant policy, it is questionable whether enhancement can be promoted in practice.

Whether or not quality assurance is discussed in the broader or the narrow sense, academics in western countries keep calling for enhancement. However, the following review will show that, in China, the concept of enhancement is not only missing in government policies but also in academic writing.

2.2.2 Definitions in the Chinese Literature

The understanding of quality assurance in higher education in the Chinese literature is still heavily influenced by the early business-orientated theory in the western literature, with the effect that quality assurance is largely focused on accountability. In addition, quality assurance in China only focuses on teaching, while excluding research. This is reflected in the interchangeable use of 'teaching quality assurance' and 'quality assurance'.

The Chinese versions of 'quality enhancement' or 'quality improvement' have not been used much in the Chinese context. The two concepts mainly appear in articles that introduce experience from other countries, mainly the UK. When introducing the enhancement-led institutional review in Scotland, Wang (2009) translates quality enhancement into 'Zhi Liang Zeng Qiang[3]' or 'Zhi Liang Qiang Hua[4]'. When talking about the QAA's practice, Wang (2008) translates quality improvement into 'Zhi Liang Ti Gao[5]' or 'Zhi Liang Ti Sheng[6]'.

3 'Zhi Liang' means quality; 'Zeng Qiang' means strengthen, enhance, boost or heighten.

4 'Zhi Liang' means quality; 'Qiang Hua' means strengthen, intensify or consolidate.

5 'Zhi Liang' means quality; 'Ti Gao' means raise, heighten, enhance, increase, improve, raising, lifting, up-grading, enhancement or hoist(ing).

Yuan (2010) uses another Chinese term –'Zhi Liang Gai Jin'– to describe quality improvement. However, Yuan regards it as a negative term in one article (Yuan, 2010) but as positive in another article (Yuan, 2011). In the first article, Yuan (2010) develops a systematic comparison between quality assurance and quality improvement, in which he associates quality improvement with higher pressure, more external interference, higher demand and punishment. He feels that asking for quality improvement places universities under more pressure than when asking for quality assurance to take place, and it is suggested that this is because he mistakes quality improvement for quality assurance. However, Yuan's (2011) work focuses on the literal meaning of 'quality improvement'– i.e. on the improvement of the quality of education – and here the term is regarded positively.

The fact that the same author uses different Chinese terms interchangeably in the same article or the same author uses the same Chinese term to mean different concepts illustrates the lack of a stable term, which in turn reflects a shortage of literature and a lack of understanding. This may be attributed to the attachment to the origin of quality assurance.

The concept of quality assurance originated from business in industrialised countries. Quality assurance in higher education in China followed this tradition. It is associated with predetermined standards or goals and customer satisfaction. For instance, Li (2006) argues that quality assurance gives customers and other related parties confidence in that it shows that the processes and the products have met certain

---

6 'Zhi Liang' means quality; 'Ti Sheng' means promote, advance, hoist, elevate or lifting.

standards. Similarly, Chen et al. (2004) define quality assurance as the manufacturer ensuring that its products or services consistently meet predetermined goals in order to satisfy its customers. These uncritically borrowed definitions of quality assurance assume that education functions in the same way as manufacturing.

Moreover, when discussing quality assurance, many Chinese academics focus on accountability and show a supportive attitude towards this focus. Xi and Huang (2010) regard quality assurance as systematic evaluation and quality control.[7] Jia (2003) defines quality assurance as monitoring or inspection and suggests that government should create an external quality assurance system to which universities are accountable. Additionally, it is posited that universities should establish an internal quality assurance system that they hold themselves accountable to (Du and Zeng, 2007).

While Harvey and Newton (2007) criticise the fact that accountability-led EQM does not take teaching into account, quality assurance often stresses teaching in China. Teaching takes primary importance in quality assurance in China (Guan, Cai and Gong, 2011). Some academics think the purpose of quality assurance is only to ensure the quality of teaching (Jia, 2003; Xi and Huang, 2010). On the other hand, other academics extend the scope of 'teaching quality assurance' to the whole concept of quality assurance (Gu, 2007; Li, 2010; Wang and Zhang, 2008). These academics expand 'teaching quality assurance' from teaching to research, facilities, personnel, finance, support and services, graduate employment rates, performance-related pay schemes for teachers and the predominance of a quality culture. The interchangeable

7 Quality control is a system of routine technical activities, to measure and control the quality of the inventory as it is being developed.

use of quality assurance and teaching quality assurance is also evident in government policy; after all, the central government named the first-ever nationwide evaluation of universities the UTE in order to show the focus on teaching. However, the UTE does not only include the evaluation of teaching, but also of research, resources, personnel and management (Ministry of Education, 2004b). Thus, in the Chinese literature some academics limit the general term 'quality assurance' to just teaching quality assurance while other academics use the term 'teaching quality assurance' to cover the whole concept of quality assurance. However, the emphasis on teaching does not necessarily mean that there is an emphasis on enhancement in the Chinese context.

Facing the imbalanced emphasis on accountability, many Chinese academics call for attention to be placed on teacher reward schemes, unlike their western counterparts who demand autonomy. Many Chinese academics demand that quality assurance should involve a reward scheme for teachers, who should be rewarded according to their achievements (Gu, 2007; Sun and Zhang, 2008; Xi and Huang, 2010).

2.3 Empirical Studies on the Impact of Quality Assurance Policies

The international empirical studies on the impact of quality assurance will be presented first, followed by Chinese empirical studies.

2.3.1 International Empirical Studies

Reviewing research on quality assurance over the previous 15 years, Harvey and Williams (2010) observe that the impact of quality assurance is an under-researched area. Among the existing literature in this area, there are both empirical research articles and discursive articles or reviews. Given that this research is empirical, only empirical research articles will be reviewed here. Furthermore, the existing empirical research on the impact of quality assurance can be categorised by actors and areas of

impact. The actors can be students, academics, academic managers and administrative staff. The areas of impact can include impact on student learning, teaching, research, programmes, institutional policy and management, etc.

For example, there is research on the impact of quality assurance on student learning (Cardoso, Santiago and Sarrico, 2012; Gift and Bell-Hutchinson, 2007; Horsburgh, 1999; Tam, 2007; Volkwein et al., 2006; Webster et al., 2009; Zhang et al., 2011); on research (McNay, 1999; Westerheijden, 1997); on how academics perceive the definition of quality (Barandiaran-Galdós et al., 2012; Bauer, 1996; Bauer and Henkel, 1997); on how academics perceive academic values and autonomy (Bauer, 1996; Bauer and Henkel, 1997); on programmes (Frederiks, Westerheijden and Weusthof, 1994; Gerbic and Kranenburg, 2003; Silva, Reich and Gallegos, 1997); and on institutional policy and management (Brunetto and Farr-Wharton, 2005; Shah, 2012; Vilgats and Heidmets, 2011).

However, the focus of the current research lies in only two actors, which are academics and academic managers, and one area of impact, which is the impact on teaching practice. This literature review therefore only includes empirical research articles exploring academics' and academic managers' views on the impact of quality assurance policies concerning the quality of teaching. The methodologies described in those studies are dominated by surveys and interviews. Survey studies are reviewed first, followed by a review of interview studies.

In regard to whether quality assurance policies have made a contribution to the improvement of quality or not, Newton (1999b) and Saarinen's (1995) studies

suggested that the majority of their respondents answered no. The respondents of Newton's (1999b) study were 72 academics and academic managers at a post-1992 university (NewColl) in Wales (the response rate was 66%). Newton (1999b) reported that only 28% of staff with administrative responsibilities and 2% of staff without administrative responsibilities believed that quality assurance contributed to real improvement in quality. Moreover, only 27% of staff with administrative responsibilities and only 9% staff without administrative responsibilities considered quality assurance as having contributed to their personal improvement. The respondents of Saarinen's (1995) study were 151 departments among the Humanities and Sciences Departments at all universities in Finland and all departments at University of Oulu and University of Jyväskylä (the response rate was 63%). Saarinen (1995, 230) found only 'a quarter of the respondents who mentioned impacts referred in some way to the development of teaching'.

Langfeldt et al. (2008) revealed divided opinions over the effects of quality policies on the quality of teaching/education among different groups. Among the 335 respondents, there were institutional leaders, administrative staff, academics and students, who had been visited by the Norwegian Agency for Quality Assurance in Education (NOKUT) panels (the response rate was 64%). The percentages of institutional leaders and administrative staff who believed in clearly positive effects were around 11% to 23% higher than those leaders and administrative staff who perceived no effect or those who thought that the effects were partly positive and partly negative. Among academics, the opinions were distributed more evenly between clearly positive effects (23%), partly positive and partly negative effects (26%) and no effect (28%). The percentages of respondents who perceived clearly

negative effects were very low, 1% for institutional leaders, 9% for administrative staff and 2% for academics. Langfeldt et al.'s findings suggested firstly that there was a division of opinion between teaching and administrative staff and secondly that the quality exercises conducted by NOKUT did not have clearly negative effects on the quality of teaching/education but did have some negative effects or no effect.

What conclusions can be drawn from Langfeldt et al. (2008), Saarinen (1995) and Newton's (1999b) research? If 51% is used as a threshold (admittedly, the decision on the threshold is an arbitrary choice), it can be concluded that these quality assurance exercises failed because fewer than half of the respondents considered that they contributed to the improvement of quality.

A number of interview studies have been conducted to explore whether academic managers and academics consider quality assurance policy and practice as having promoted the improvement of teaching and why. They are case studies of the impact of (a) quality assurance exercise(s) by managers or/and academics in the UK (Brennan, Frederiks and Shah, 1997; Cheng, 2008; Cheng, 2010; Cheng, 2011; Hoecht, 2006; Laughton, 2003; Newton, 1999b; Newton, 1999a; Newton, 2000; Williamson, 2005), in Sweden (Askling, 1997), in Finland (Haapakorpi, 2011), in Italy and the Netherlands (Minelli et al., 2006), in 16 countries in the Organisation for Economic Co-operation and Development (OECD) (Brennan and Shah, 2000) and in Hong Kong (Jones and Darshi De Saram, 2005).

As suggested by Brennan and Shah (2000) and Lee, Huang and Zhong (2012), impact can be categorised into the national, institutional, departmental, and individual levels.

The researcher finds that the impact at national and institutional levels was perceived positively by both managers and academics but that a negative impact was more frequently perceived at the individual level.

The perceived impact at national level was in ensuring the consistency of academic standards across different universities (Cheng, 2008) and in promoting transparency within higher education and information sharing among higher education's stakeholders (Cheng, 2008; Haapakorpi, 2011; Minelli et al., 2006).

The perceived impact at institutional level was revitalising the institutional and departmental focus on the quality of education (Williamson, 2005) and ensuring a minimum standard of programmes (Cheng, 2008).

There were only two positive impacts at individual level mentioned in Minelli et al. (2006) and Cheng's (2008) research. Managers at the University of Amsterdam reported new teaching methodologies and improvement actions (Minelli et al., 2006). The QAA institutional audit led to an increasing awareness of good teaching among academics at the pre-1992 university (Cheng, 2008). However, with the exception of these two positive impacts, all of the other impacts at the individual level were perceived as negative, including in regard to impression management, bureaucracy, the increased burden of administrative work, and issues with validity[8] and reliability[9].

8 Validity concerns the question of whether the indicators employed by the quality assurance or mechanisms serve as the best proxies for quality.

9 Reliability refers to the reliability of the judges of quality assurance policies and mechanisms. This consideration entails asking whether the judges are able to make fair judgments about target groups' performance in regard to quality.

In fact, one of the most frequently mentioned negative impacts of quality assurance policies was ritualism or impression management, which was referred to by interviewees of the studies conducted by Brennan, Fredriks and Shah (1997), Cheng (2008), Laughton (2003) and Newton (1999b). Ritualism refers to a phenomenon in which policies and procedures are made and implemented as rituals to symbolise the accomplishment of quality assurance tasks, while the essence of quality is missing (Newton, 1999b). Impression management means that policies are made and actions are taken with a major purpose to leave a good impression on stakeholders while the assurance of quality becomes secondary (Cheng, 2008). The extreme case of ritualism and impression management was misbehaviour. Interviewees reported that teachers decreased students' workload in order to increase productivity (Minelli et al., 2006). The second frequently mentioned negative impact was bureaucracy (Cheng, 2008; Hoecht, 2006). The extreme case of bureaucracy was that quality systems and procedures became too complex and time-consuming (Brennan, Frederiks and Shah, 1997; Jones and Darshi De Saram, 2005). The third negative impact was the increased burden of administrative work (Brennan, Frederiks and Shah, 1997; Cheng, 2008; Haapakorpi, 2011; Laughton, 2003; Minelli et al., 2006; Newton, 1999b). The fourth negative impact was a threat to autonomy. Audit was regarded as a fault-finding exercise (Cheng, 2008). Some academics feel their autonomy was undermined by the demand for transparency (Askling, 1997). For example, peer observation was regarded as an interference in teaching (Cheng, 2008). Consequently, Haapakorpi (2011) observed insignificant benefits at the 'grass-roots' level.

As suggested by Brennan and Shah (2000), the impact of quality assurance policies was fundamentally affected by people's perception of the validity and reliability of

the quality assurance instruments. Hence, many findings regarding the impact of quality assurance policy and practice related to managers' or academics' perceptions of the reliability and the validity of quality assurance instruments.

These concerns over the validity of quality assurance instruments centred on the discrepancy between what was assured and what should be assured. Newton (1999b) found that quality was not assured by a total quality approach; instead, it was reduced to discipline and technology. Jones and Darshi De Saram (2005) also reported a similar finding in that quality procedures were prescribed in too much detail, which blurred the whole picture of quality, and that there was an over-reliance on quantitative performance indicators. Moreover, 'the physical/administrative environment and services provided by various support units fall short of what is required and often are not aligned with the real requirements for facilitating good quality teaching and learning environments' (Jones and Darshi De Saram, 2005, 51). Hence, Jones and Darshi De Saram (2005) observed that 'quality systems are often peripheral to the real substantive issues involved in the assurance and enhancement of quality in teaching and learning' (Jones and Darshi De Saram, 2005, 51); 'the administrative demands of overall systems can actually subvert what should be the essential foci on quality in teaching and learning' (Jones and Darshi De Saram, 2005, 51). In order to tackle this problem, Jones and Darshi De Saram (2005) called for an overall communication and agreement on procedures and standards.

Laughton (2003) questioned the reliability of quality assurance instruments. The issue of reliability was mainly covered by Brennan, Fredriks and Shah's (1997) and Cheng's (2008) studies in England. Data collected through interviews with managers

and academics revealed concerns over the reliability of the QAA's institutional audit, external examining system, the National Student Survey (NSS) and the student evaluation of teaching. In each of the quality assurance mechanisms, there was always an actor that failed. The concerns of academics were as follows: some QAA or Higher Education Funding Council for England assessors were unprofessional; the quality of external examiners varied; some external examiners knew little of the reviewed department; it was difficult for external examiners to compare standards across the institutions; the sample of students who participated in the NSS was very small when it was broken down to the department level; and students did not take the student evaluation of teaching seriously.

2.3.2 Chinese Empirical Studies

Discursive articles dominate the Chinese literature and there is a limited number of empirical studies on the impact of quality assurance policy and practice. There are six survey studies on the impact of the UTE (Gao et al., 2006; Li, Li and Qu, 2006; Qu and Li, 2008; Zhang and Xue, 2009; Zhang and Zhang, 2008a; Zhang and Zhang, 2008b) on the quality of teaching from the perspective of leaders and teachers. In terms of the Quality Project, as the scale and the influence of the Quality Project awards are much smaller, the researcher has not found any empirical studies on the impact of the Quality Project on teaching practice.

The perceived positive impacts brought by the UTE were mainly concentrated on impacts at institutional level, which involved improvements in standardising teaching practice (Li, Li and Qu, 2006), establishing internal quality assurance mechanisms (Li, Li and Qu, 2006), improving teaching facilities (Li, Li and Qu, 2006; Qu and Li, 2008), strategic planning (Qu and Li, 2008) and staffing (Zhang and Xue, 2009). The

importance of teaching was reinstated at both institutional and individual level (Qu and Li, 2008). On the other hand, the perceived impacts at individual level were mixed. Zhang and Xue (2009) reported no impact on professional ethics and teaching methods. However, Qu and Li (2008) reported an impact in terms of facilitating the improvement of teaching quality.

There is a contrast in the findings of research commissioned by the Ministry of Education and that initiated by teachers. Qu and Li's (2008) research was commissioned by the Ministry of Education. They asked respondents whether the UTE contributed to the improvement of the quality of teaching and claimed that around 70% of respondents thought so. This suggests that most of the respondents thought the UTE had improved the quality of teaching. Although none of the other researchers asked this question directly, answers to related questions asked by Zhang and Xue (2009) suggest a different perspective. According to their research, 55.5% of respondents disagreed with the statement that the UTE contributed to an improvement in teaching methods and 64% disagreed that the UTE improved teachers' professional ethics. Such answers seem to indicate that the UTE was not perceived to an improvement in the quality of teaching.

Standardisation was a topic touched on by Li, Li and Qu (2006) and Zhang and Xue (2009). Li, Li and Qu (2006) state that 95% of their respondents reported a significant contribution to the standardisation of teaching practice. Zhang and Xue (2009) found that 75.3% of respondents agreed with the statement that the UTE made teaching plans more standardised.

Zhang and Zhang (2008b) and Qu and Li (2008) were interested in perceptions of the necessity of launching the UTE. They reported that 80% of respondents (Zhang and Zhang, 2008b) and 94% of winners of Distinguished Teacher Awards (Qu and Li, 2008) believed that the UTE was very necessary or necessary.

Besides survey studies, there are two interview studies investigating administrative staff and teachers' perceptions of the impact of the UTE and the reasons behind their perceptions (Lee, Huang and Zhong, 2012; Liu, 2011; Liu, 2013). Lee, Huang and Zhong (2012) and Liu (2011; 2013) conducted interview studies to investigate managers' and teachers' perceptions of the impact of the UTE and their reasoning. The perceived positive impacts were again mainly at the national and institutional levels. At national level, the UTE promoted greater local government investment in higher education (Lee, Huang and Zhong, 2012). At an institutional level, universities' strategic planning and internal quality assurance mechanisms were enhanced (Lee, Huang and Zhong, 2012; Liu, 2011), teaching facilities were improved, and staff recruitment policies were revised (Liu, 2011).

Lee, Huang and Zhong (2012) and Liu's (2011; 2013) studies also revealed negative impacts at the individual level. As touched on above, the first problem was ritualism or impression management. The extreme case of impression management was using excessive means to impress evaluators or falsifying documents in order to meet the requirements. A considerable amount of money, time and effort were wasted and the UTE became a hindrance to daily teaching (Lee, Huang and Zhong, 2012). Liu (2011) concluded that there was no influence on teaching content and teaching methods at her three case universities.

2.4 Empirical Studies on the Impact of Awards for Good Teaching

The researcher found only four pieces of empirical research on the impact of awards for good teaching (no empirical research on the impact of Quality Project awards was found). Dinham and Scott (2002) and Vaden-Kiernan, Michie and Frechtling (1994) confirmed that awards offered acknowledgement to award winners. However, Jacobsen (1989) and Tollefson and Tracy (1983) did not find significant differences in teaching effectiveness between award winners and non-award winners.

2.5 Empirical Studies on the Impact of Class Observation

The researcher found only discursive articles on the impact of class observation on staff in the Chinese literature. Hence, only empirical studies found in the international literature are included. As illustrated by Keig and Waggoner's investigation (1994), the majority of academics expressed a belief that class observation would contribute to the improvement of teaching in theory. Keig and Waggoner (1994) asked faculty members at 26 baccalaureate degree-granting independent colleges and universities whether or not they thought peer involvement in instructional evaluation would benefit the quality of instruction. Among the 372 respondents (response rate: 49.5%), 77% of them believed so.

However, whether class observation is able to fulfil this purpose in practice remains debatable. Some empirical research demonstrated that class observation was able to achieve it. Having examined the recent literature on peer review of teaching, Thomas et al. (2014) summarised its positive impacts identified by empirical research, which are: offering confirmation (Hanson, 1993) or a different perspective (Bell and Mladenovic, 2008; Hanson, 1993), promoting collegiality, autonomy or self-

regulation (Al Qahtani et al., 2011), and identifying strengths and weakness (Thomas et al., 2014).

Thomas et al. (2014) also summarised threats to class observation identified by empirical research: fear (Taylor, Atwood and Hutchings, 2000; Lomas and Nicholls, 2005), uncertainty (Taylor, Atwood and Hutchings, 2000), lack of time (Kell and Annetts, 2009), interference with academic freedom (Keig and Waggoner, 1994). These factors can prevent implementation of and meaningful participation in class observation.

Britt (1982) pointed out that the most common objection to class observation relates to academic freedom. Goldschmid (1978, 226) argued that 'the classroom remains a private domain and in the name of academic freedom, public scrutiny, even in the form of peer assessment, is taboo'. This negative perception can cause resistance and dissatisfaction. However, Britt's (1982) study demonstrated a different picture. She surveyed 215 professors, associate professors and assistant professors at three public colleges or universities (response rate: 55.4%). Surprisingly, 78% of the respondents disagreed with the statement that academic freedom was jeopardised by colleague observation of classroom teaching; 76% of respondents disagreed that colleague evaluation using classroom visits constituted an infringement of the private domain.

The positive and negative impacts can co-exist in the same institution due to variations in staff commitment and reflective stands or in the same higher education system due to variations in implementation approaches adopted by different institutions.

Three studies revealed variations in staff commitment and reflective stands. Shortland (2004) studied a trio who actively engaged in a peer observation scheme. The trio identified the process as valuable and incorporated good practice into their teaching. However, the trio reported 'some staff members completing observation documentation simply to comply with the managerial requirement to do so, having not actually undertaken the observation process' (Shortland, 2004, 226). Hammersley-Fletcher and Orsmond (2005) investigated the impact of peer observation at two schools at a university. They found that, while many faculty members experienced fear, anxiety and difficulty in giving or receiving criticism, other faculty members took this opportunity to reflect their teaching techniques and materials. Cockburn (2005) shared the same opinion. He contended that, when academics took a non-reflective stand they perceived class observation as bureaucratic scrutiny, which was artificial and incredible and reduced quality to box-ticking exercise. On the other hand, when academics took a reflective stand, they would find class observation served as a reflective mirror, which helped to boost their professional identity, to improve their teaching skills, and to enable them to give feedback to their colleagues.

How class observation was conducted at different institutions could result in different impacts. Bell (2001), Bell and Copper (2013), Blackwell and McLean (1996) and Lomas and Nicholls (2005) reported positive impacts at the institutions where they conducted their studies. Bell and Copper (2013) and Lomas and Nicholls (2005) summarised the conditions which enable class observation to contribute to the improvement of teaching quality as: voluntary participation, a formative process, collegial collaborations, and support from the leadership. Their underlying argument

was: if these conditions were met, class observation would have positive impact, whereas ‘inappropriate implementation may undermine developmental potential and promote cynicism’ (Blackwell and McLean, 1996, 163).

Blackmore (2005) and Douglas and Douglas (2006) pointed out the importance of institutional culture. They found that ‘at the case study institution people generally did not like to feedback any negative criticism as they did not want to undermine colleagues’ confidence’ (Blackmore, 2005, 228; Douglas and Douglas, 2006, 9), because of the lack of culture of constructive criticism. The lack of this culture prevented faculty members from offering useful feedback and thereby improving the quality of teaching.

Although Bell (2001) reported that the Teaching Development Programme based on class observation at the University of Wollongong promoted the development of collegiality, whether the scale of change was substantial enough to change institutional culture was questionable. In Keig and Waggoner’s (1994) study, the majority of respondents did not think that ‘the collegial climate of the college or university would be improved by faculty participation in formative evaluation’ (Keig and Waggoner, 1994, 100).

2.6 Empirical Studies on the Impact of the Student Evaluation of Teachers

The researcher did not find any empirical research on the impact of class observation on staff in the Chinese literature, therefore this review only includes empirical studies in the international literature. Student evaluation is a well-studied area in international educational research (Marsh, 1987, 260). However, studies on ‘feedback reaction’, i.e.

the effects of student evaluation of staff, are a relative underdeveloped area (Moore and Kuol, 2005). Academics' views on student evaluation vary 'from support and acceptance to complete opposition and rejection' (Nasser and Fresko, 2002, 188).

Many empirical studies found that the majority of academics trusted that 'ratings they received from students were an accurate reflection of their teaching performance as perceived by students' (Baxter, 1991; Moore and Kuol, 2005; Ryan, Anderson and Birchler, 1980) and instructor self-evaluations showed significant agreement with students' evaluations (Howard, Conway and Maxwell, 1985; Marsh, 1982; Marsh, Overall and Kesler, 1979; Webb and Nolan, 1955).

However, individual academics' views on student evaluation are more complicated than the conclusions only drawn from the majority's opinions. Nasser and Fresko (2002) revealed that many academics believed that 'there exist[ed] a relationship between the quality of instruction and student evaluations… Meanwhile, many academics also believe[d] that other factors affect[ed] student ratings. For example, many of them believe[d] that instructors who demand[ed] a lot from their students, or courses which [were] difficult, [would] receive lower ratings' (Nasser and Fresko, 2002, 192-193). The 'other factors' that are discussed frequently in the literature are grading leniency, workload/difficulty and 'Dr. Fox' effect. They affect the perceived reliability of student evaluation.

Grading leniency is a belief in a positive correlation between students' grades and teachers' results of student evaluation. Against this common belief, Marsh and Dunkin (1997) reported that class-average grades correlated .20 with overall teacher

ratings in [Students' Evaluations of Educational Quality] (SEEQ) research, and this finding was consistent with the extensive review of this relation reported by Feldman (1976; 1997). Marsh and Dunkin (1992) concluded that the best single estimate of the relation between the overall teacher rating and expected grades was probably the .2 value reported by Centra and Creech (1976) based on 9,194 class-average responses from a diversity of different universities, courses, settings, and situations. However, 'Centra (2003), in subsequent research based on a much larger, diverse sample of 55,549 classes, found a slightly lower correlation of only .11' (Marsh, 2007, 352).

Workload/difficulty is another popular concern among academics who believe there is a negative correlation between workload/difficulty and the results of student evaluation. In fact, research found that workload/difficulty was positively correlated with students' evaluations (Marsh and Dunkin, 1997; Marsh and Roche, 2000; Marsh, 2001). Furthermore, Marsh (2001) revealed that students distinguished between good workload, which posed an appropriate challenge and on which time spent was valuable, and bad workload, which was too difficult or too easy and a waste of time. 'Good workload was positively related to [Students' Evaluations of Teaching Effectiveness] (SETs) and learning, but bad workload … had negative relations' (Marsh, 2007, 350).

Another frequently mentioned bias is the 'Dr. Fox' effect. 'The "Dr. Fox" effect is defined as the overriding influence of instructor expressiveness on SETs, and has been interpreted to mean that an enthusiastic lecturer can "seduce" students into giving favourable evaluations, even though the lecture may be devoid of meaningful content (see Marsh and Dunkin, 1997; Marsh, 1987)' (Marsh, 2007, 357). Although

most research in experimental settings supported the Dr. Fox effect (Abrami, Leventhal and Perry, 1982), Marsh and Ware (1982) demonstrated that 'in conditions most like university classroom, where students were given incentives to do well on the achievement test, the Dr. Fox effect was not supported' (Marsh, 2007, 357).

Besides reliability, the perceived validity of student evaluation also affects academics' perceptions. The perceived validity can be determined by whether academics' definition of quality teaching agrees with students' or not. Fisher, Alder and Avasalu (1998, 153) found academics and students 'agree[d on] the criteria [of lecturing], but portray[ed] different pictures of the ways in which the criteria combine[d] to produce the understanding of what is a good lecture'. Oliveira, Oliveira and Costa (2012, 49) revealed 'that, although teachers and students refer[red] to the same dimensions that influence quality, they [had] different perspectives about their importance'. Cheng (2011, 3) discovered that 'most of the [Teaching Excellence] Award winners associated the concept of quality with transformative learning. However, students, while recognising the concept, defined quality in more instrumental terms'. These studies show that students and academics can agree on the elements that constitute quality teaching, but can disagree on the combination and weightings of elements, and the purpose of quality teaching.

In addition, the validity of student evaluation can be affected by how the results of student evaluation are reported to academics. One common problem that violates the validity of student evaluations at the three case study universities is the use of total scores (see Sections 5.2.1, 6.2.1 and 7.2.1 for details). Marsh (1984; 1987; 2007) contends that the quality of teaching was a multidimensional concept. Therefore, he

opposes the use of total scores, because: ‘First, there are many possible indicators of effective teaching and many possible uses for student ratings; the component that is “most valid” will depend on the criteria being considered (Marsh and Overall, 1980). Second, reviews of different validity criteria show that specific components of student ratings are more highly correlated with individual validity criteria than an overall or total rating (e.g. student learning (Cohen, 1981); instructor self-evaluations (Marsh, 1982;Marsh, Overall and Kesler, 1979); the effect of feedback for the improvement of teaching (Cohen, 1980). Third, the influence of a variety of background characteristics suggested by some as “biases” to student ratings is more difficult to interpret with total ratings than with specific components (Marsh, 1980; Marsh, 1983). Fourth, the usefulness of student ratings, particularly as diagnostic feedback to faculty, is enhanced by the presentation of separate components. Finally, even if it were agreed that student ratings should be summarized by a single score for a particular purpose, the weighting of different factors should be a function of logical and empirical analyses of the multiple factors for the particular purpose…’ (Marsh, 1984, 710).

When studying the reactions taken by academics, Ryan, Anderson and Birchler (1980, 322) found that ‘the most frequently reported changes were for activities that reduced coursework demands on students either directly by lowering the difficulty level of course content and exams’. This shows that academics were concerned about grading leniency and workload/difficulty. However, Ryan, Anderson and Birchler (1980) also discovered that 50% of respondents did not report any change. Similarly, Spencer and Flyr (1992) reported that the majority of respondents did not make changes. On the other hand, Baxter (1991) and Moses (1988) stated that the majority of their

respondents reported that they made changes to their teaching method, teaching style and teaching materials. Based on their empirical research, Franklin and Theall (1989, 1) contended that 'those with a negative attitude were likely to use student ratings less often and less well. Positive attitudes were associated with higher knowledge scores'.

Different research reported different effects of student evaluation on staff morale and job satisfaction. Ryan, Anderson and Birchler (1980) and Murray and Newby (1982) reported a dominance of negative effects, whereas Baxter (1991), Jacobs (1987) and Moses (1988) found improved job satisfaction and morale among the majority of staff. Based on the research on academics' reactions to student evaluation reviewed above, it can be concluded that it is uncertain how academics would react to student evaluation and why.

2.7 Summary

Quality is a multifaceted concept. It can be viewed as pertaining to the exceptional, to perfection or consistency, to fitness for purpose, to value for money or to transformation. What purposes the quality assurance policy should fulfil is also understood differently. In the international literature, there is a debate on whether quality assurance should only address accountability or whether it should also encompass improvement or, in other words, enhancement and what is achievable. A review of the Chinese literature reveals the absence of the concept of improvement or enhancement. Another problem found in the Chinese literature is that there is a dominant focus on the quality of teaching.

By reviewing interview studies in the international and Chinese literature on the impact of quality assurance policies, the researcher finds that the impact revealed by

these studies either relates to the actions taken by teachers in regard to policies or to the consequences or the outcomes of the actions. Therefore, the impact of the literature reviewed above can be summarised as follows:

Actions	No actions
New teaching methodologies and improvement actions (Jones and Darshi De Saram, 2005)	Lack of resources (Jones and Darshi De Saram, 2005)
Bureaucracy (Cheng, 2008; Hoecht, 2006; Jones and Darshi De Saram, 2005)	
Impression management (Brennan, Frederiks and Shah, 1997; Cheng, 2008; Laughton, 2003; Lee, Huang and Zhong, 2012; Newton, 1999b)	
Consequences	**No consequences**
An increasing awareness of good teaching (Cheng, 2008)	Problem of validity (Brennan, Frederiks and Shah, 1997; Cheng, 2008)
Increased administrative burden (Brennan, Frederiks and Shah, 1997; Jones and Darshi De Saram, 2005)	Problem of reliability (Brennan, Fredriks and Shah, 1997; Cheng, 2008)
	No influence on teaching content and teaching methods (Liu, 2011)

Table 3 Summary of the impacts seen in the literature

The categories of impact illustrated by Table 3 inspired the researcher to develop a conceptual framework to analyse impact, which will be elaborated in Chapter Three.

As mentioned before, these impact studies mainly used either surveys or interviews. A weakness of interview studies is their lack of statistical generalisability, while a strength that distinguishes survey studies from interview studies is their statistical generalisability. The level of generalisability is determined by the sampling frame, the sample size, the sampling strategy and the response rate (Bryman, 2008). However, the generalisability of the studies reviewed above varies considerably. For instance, Zhang and Zhang's (2008a) study is very weak in its generalisability. Their sample population was the administrative staff at 54 universities and colleges in China – but

only 73 valid questionnaires were returned to them. It is questionable if the opinions of these 73 staff can be taken to represent all administrative staff at the 54 universities and colleges. However, even when the sample size is big, the non-response rate could still endanger generalisability. For example, the sample population of Newton's (1999b) study is the 147 staff at the NewColl and the sample size is 109 staff – however, only 72 staff returned their questionnaires and it is unclear if the information that could have been provided by the other 75 staff who did not respond would alter the findings. Hence, quite often the generalisability of survey studies is also problematic.

The strength of the interview approach is the ability to probe, to ask for more details, to clarify information and to investigate causes. This addresses some of the weakness that survey studies suffer from. However, only one of the existing interview studies (Cheng, 2008) focuses on the impact of quality assurance policies on teachers and their day-to-day teaching and no research has been conducted with a particular focus on investigating the impact of quality assurance policies in China. More importantly, none of the existing studies articulates what the impact is. This research intends to address these deficiencies.

Impact studies of three quality assurance mechanisms are reviewed above. The number of studies on the impact of awards is limited. They find that awards are able to offer acknowledgement to award winners. However, the impact on non-award winners has not yet been investigated. This research addresses this gap. The literature on class observation reviewed above mainly looks at the impact of one type of class observation, which is peer class observation. Within peer observation, it draws a line

between summative and formative class observation and argues that the formative observation, which is non-judgemental and only used for developmental purposes, can promote improvement and vice versa. However, class observations implemented at the three case universities are not peer observation, as they are conducted by leaders, university or department inspectors, and are summative evaluations. This research explores whether these summative class observations are able to promote improvement or not. Research on student evaluation shows that the majority of academics have confidence in student evaluation. However, overall confidence among the majority of academics does not mean that academics are not concerned about its reliability and validity. The reliability issues include grading leniency, workload/difficulty and the Dr. Fox effect; the validity issues contain definitions of quality teaching and use of total scores. Moreover, research demonstrates a mixed picture in teachers' reactions. Student evaluations at the three case study universities illustrate the typical student evaluation used in Chinese universities, which is not rigorously tested for its validity and reliability, but is associated with appraisal for promotion (for details see Chapters Five, Six and Seven). This research is able to examine how teachers perceive and react to student evaluation when it is flawed but with high stakes.

3 Research Design

This chapter starts by articulating the research questions. Then, various criteria for evaluating qualitative research will be presented in order to outline the most appropriate ones for this project. The research methodology utilised by this project will then be elaborated, including data collection, the selection of cases, the analytical and the conceptual frameworks and data analysis. This chapter concludes with considerations of research ethics and limitations.

3.1 Research Questions and Research Strategy

As mentioned in previous chapters, this research is interested in the impact of quality assurance policies on teachers regarding teaching practice in the context of Chinese higher education. In order to investigate these issues, the following questions were addressed:

- What are the key policies that have shaped the quality assurance agenda in Chinese higher education in recent years?
- How are these quality assurance policies implemented at the case study universities?
- How do teachers perceive the impact of these quality assurance policies on their teaching practice?
- What are the unaddressed tensions and factors affecting the assurance of the quality of teaching?

The first research question examines the development of quality assurance policies in China. Key quality assurance policies are outlined in order to reveal how quality should be assured. The second question investigates how universities implement

national quality assurance policies; in other words, how individual universities respond to the national policies. Answers to the second question provide a deeper understanding of what teachers are theoretically expected to do in order to implement university policies. The third question explores how teachers respond to these policies in practice and how they perceive the influence of these policies. University policies and teachers' responses are compared across the three case universities. The fourth question allows the researcher and readers to step back and look at the factors that contribute to the quality of teaching beyond government initiatives. Answers to this question will provide insights into how to improve policy-making in the future.

This research adopts a case study approach because the nature of case studies meets the needs of this research. The leading exponents of case studies in the social sciences are Yin and Stake. Yin defines a case study as an empirical inquiry that

> …investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident.
>
> (Yin, 2003, 13)

Stake describes the case study as:

> … the study of the particularity and complexity of a single case, coming to understand its activity within important circumstances.
>
> (Stake, 1995, xi)

These quotes from Yin and Stake highlight what distinguishes case studies from other qualitative studies: in essence, case studies adopt a holistic approach to understanding a phenomenon or a case within its context or circumstances.

The case study approach's awareness of contexts matches the research questions. The first research question explores the context created by the national quality assurance policies; the second question explores the context created by university quality

assurance policies. These contexts provide the foundation for the analysis of teachers' opinions, addressed by the third and fourth questions.

Yin and Stake view the purpose of case studies differently. From a positivist perspective, Yin (2003) maintains that their purpose is to develop or test theory; on the other hand, from the interpretivist perspective, Stake (2005) states that the purpose is to learn from the particular cases. According to Stake (2005) through case studies researchers convey the experience of those who are studied into experiential knowledge, meaning case studies are strong in transferring this knowledge to readers to advance or modify their existing understanding.

In terms of fulfilling the purpose of case studies, this research is able to fulfil both purposes advocated by Stake (2005) and Yin (2003). To the researcher's knowledge, this research is the first of its kind, using empirical evidence to examine the impact of quality assurance policies in relation to the quality of teaching and exploring factors contributing to the improvement of the quality of teaching in China. The perceptions and experiences of teachers who were interviewed in this study will be passed on to readers, who could be other teachers, policy-makers and the general public. This study will provide insight into the impact and problems in the existing quality assurance policies in China and hopefully will help policy-makers, leaders and practitioners reflect on their opinions and practices. Hence, this research will fulfil the purpose suggested by Stake. Furthermore, the researcher intends to develop and demonstrate a systematic approach to the examination of quality assurance policy in higher education by applying existing political science theories. In addition, building upon impacts revealed by existing research and empirical evidence collected by this

study, the researcher proposes a framework to articulate impact. By contributing to the development of theory, this research will also fulfil the purpose promoted by Yin.

Within the ambit of the case study, different types of case studies can be discerned. Stake (2005) maintains that different interests in learning (intrinsic and instrumental interests) lead to different types of case studies. He distinguishes among intrinsic, instrumental and multiple case studies:

> I call a study an intrinsic case study if the study is undertaken because, first and last, one wants better understanding of this particular case…
> I use the term instrumental case study if a particular case is examined mainly to provide insight into an issue or to redraw a generalization. The case is of secondary interest, it plays a supportive role, and it facilities our understanding of something else…
> When there is even less interest in one particular case, a number of cases may be studied jointly in order to investigate a phenomenon, population, or general condition. I call this multiple case study or collective case study. It is instrumental study extended to several cases.
>
> (Stake, 2005, 445-446)

As this study is interested in the implementation of national quality assurance policies, this purpose defines the research as an instrumental case study. The adoption of either a single instrumental case study or multiple case study approach depends on which design maximises the opportunity to learn. If one university was chosen, this research could explore how this university shapes its policies so as to respond to the national policy and teachers' perceptions of the national and university policies. If different universities were chosen, this research could explore how different universities respond to the same national policies and if there are differences between the perceptions of teachers at different universities. The latter would be able to provide data to distinguish the impact made by national policies from that wrought by university policies. Hence, the multiple case study approach maximises the opportunity to learn.

Admittedly, the multiple case study design decreases the number of embedded cases, namely departments and individuals, in each university. Nevertheless, Stake (2005) argues that, because the purpose of a case study is to learn from particular cases, variety weighs more than representativeness. Having more case study universities means more variety.

3.2 From Validity and Reliability to Alternative Criteria for Evaluating Qualitative Research

There are a number of criteria for evaluating the quality of case studies or qualitative studies in general. These criteria are introduced below and subsequently their weaknesses are discussed.

To evaluate the quality of case studies, Yin (2003) follows the positivist tradition and insists that case studies should be tested by construct validity, internal validity, external validity and reliability. However, the positivist features underpinning the four tests are in conflict with the nature of qualitative study. From the interpretivist paradigm, Stake's main concern is that 'we have ethical obligations to minimize misrepresentation and misunderstanding' (Stake, 1995, 109). To validate research, Stake (1995) calls for triangulation, including data source triangulation, investigator triangulation, theory triangulation and methodological triangulation. Triangulation is necessary but is nonetheless insufficient: it is just a means to assure rigour and coherence. There are more aspects deserving consideration.

Other popular criteria for judging qualitative research are trustworthiness (Lincoln and Guba, 1985) and authenticity (Guba and Lincoln, 1989). Trustworthiness

comprises credibility, transferability, dependability and conformability. However, even Lincoln and Guba are uneasy about trustworthiness. They explain this as follows:

> They [criteria on trustworthiness] have their roots and origins in positivist assumptions, and while adjustments have been made for the different assumptions of the naturalist paradigm, there remains a feeling of constraint, a feeling of continuing to play "in the friendly confines" of the opposition's home court. In addition to their positivist ring, they share a second characteristic that leaves an uncomfortable feeling: they are primarily methodological criteria.
>
> (Guba and Lincoln, 1989, 245)

Consequently, they propose authenticity criteria, 'which spring directly from constructivism's own basic assumptions' (Guba and Lincoln, 1989, 245). Authenticity criteria include fairness, ontological authenticity, educative authenticity, catalytic authenticity and tactical authenticity (Guba and Lincoln, 1989). In fact, when Guba and Lincoln propose them, the original purpose is to judge the adequacy of evaluation. Because of this, authenticity emphasises research impact while paying less attention to research design, data collection and analysis. Therefore, the researcher thinks that these criteria are not adequate to judge the quality of qualitative research in general.

Yardley (2000) proposes alternative criteria for evaluating qualitative research – the characteristics of good qualitative research. She writes:

> Sensitivity to context
> *Theoretical; relevant literature; empirical data; sociocultural setting; participants' perspectives; ethical issues.*
>
> Commitment and rigour
> *In-depth engagement with topic; methodological competence/skill; thorough data collection; depth/breadth of analysis.*
>
> Transparency and coherence
> *Clarity and power of description/argument; transparent methods and data presentation; fit between theory and method; reflexivity.*

> Impact and importance
> *Theoretical (enriching understanding); socio-cultural; practical (for community, policy makers, health workers).*
>
> (Yardley, 2000, 219)

The strength of this framework is that the thinking underpinning these criteria is rooted in qualitative research and covers all aspects of research, from research design, data collection and data analysis to the impact of the research.

Another framework to assess the quality of qualitative research is put forward by Spencer et al. (2003), who state that 'It was developed with particular reference to evaluations concerned with the development and implementation of social policy, programmes and practice' (Spencer et al., 2003, 2). They identify 18 appraisal questions to assist an assessment of the quality of research. Under each question, they also provide a series of quality indicators, which point to 'the kinds of information needed to judge whether or not the quality feature concerned has been secured' (Spencer et al., 2003, 8). The appraisal questions cover key processes involved in qualitative research, including design, sampling, data collection, analysis, findings, reporting, reflexivity and neutrality, and ethics and auditability. These appraisal questions and quality indicators are developed under the guidance of four central principles, which are:

> **contributory** in advancing wider knowledge or understanding about policy, practice, theory or a particular substantive field;
>
> **defensible in design** by providing a research strategy that can address the evaluative questions posed;
>
> **rigorous in conduct** through the systematic and transparent collection, analysis and interpretation of qualitative data;
>
> **credible in claim** through offering well-founded and plausible arguments about the significance of the evidence generated.
>
> (Spencer et al., 2003, 7)

These four central principles promoted by Spencer et al. (2003) partly overlap with the criteria proposed by Yardley (2000). 'Contributory' parallels 'impact and importance'; 'defensible in design' and 'rigorous in conduct' parallel 'commitment and rigour'; and 'credible in claim' parallels 'coherence'. However, some of the ideas put forward by Yardley are not mentioned by Spencer et al., which are 'sensitivity to context' and 'transparency'. Hence, this research uses Yardley's characteristics of good qualitative research as the general guidance throughout the whole research. At the end of this thesis, the author will use Yardley's criteria as tools to reflect on the way this project was conducted.

3.3 Methods for Data Collection

To answer the first research question, national policy documents were collected from the public domain. Data for the second research question were university and departmental policies. Some university policy documents were available in the public domain and some were collected during the fieldwork. Appendix 1 lists all the national and university policy documents reviewed for this study. Departmental policies were not available in the public domain. A feasible way to collect departmental policies was by talking to departmental policy-makers, i.e. the department heads. Consequently, interviews were conducted to inquire into departmental policies. In addition, how these policies were implemented was also triangulated by teachers' accounts. Answers to the third and the fourth research questions had to be built upon teachers' accounts and perceptions. The purpose of acquiring teachers' accounts in order to understand how teachers perceive the influence of these policies and to explore what factors contributed to the improvement of teaching quality led the data collection method to be interviews.

Kvale (1996, 5-6) defines the research interview as 'an interview whose purpose is to obtain descriptions of the life world of the interviewee with respect to interpreting the meaning of the described phenomena'. Tuckman describes the functions of interviews in the following terms:

> By providing access to what is "inside a person's head", [it] makes it possible to measure what a person knows (knowledge or information), what a person likes or dislikes (values and preferences), and what a person thinks (attitudes and beliefs).
>
> (Tuckman, 1972, 197)

The nature of interviews suits the purpose of this research, as the third and the fourth research questions are intended to explore the 'life world' of the teachers and what is hapenning inside teachers' heads, including in regard to their knowledge, information, values, preferences, attitudes and beliefs.

There are many different types of interviews. Minichiello et al. (1990) provides a continuum to conceptualise different interviews (see Figure 1).

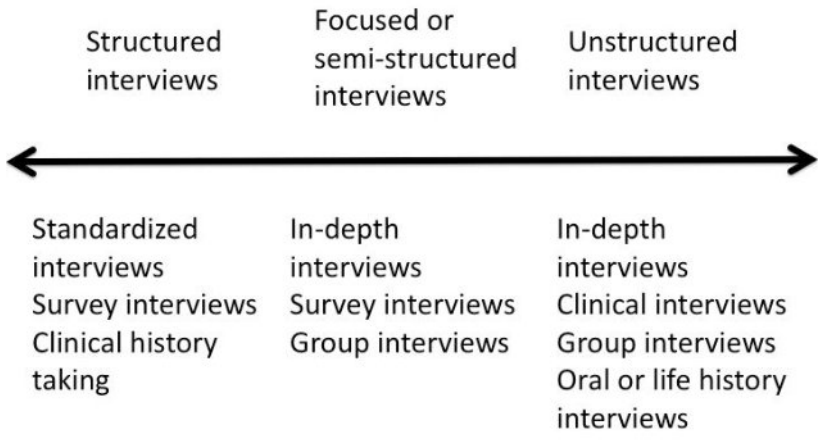


Figure 1 The continuum model of interviews

(Adapted from Minichiello et al., 1990, 89)

Robson (2002) summarises the features of different interviews:

> *Fully structured interview* Has predetermined questions with fixed wording, usually in a pre-set order. The use of mainly open-response questions is the only essential difference from an interview-based survey questionnaire.
>
> *Semi-structured interview* Has predetermined questions, but the order can be modified based upon the interviewer's perception of what seems most appropriate. Question wording can be changed and explanations given; particular questions which seem inappropriate with a particular interviewee can be omitted, or additional ones included.
>
> *Unstructured interviews* The interviewer has a general area of interest and concern, but lets the conversation develop within this area. It can be completely informal.
>
> (Robson, 2002, 270)

When choosing the appropriate type of interview, Bryman (2008) describes several scenarios:

> Researchers who are concerned that the use of even the most rudimentary interview guide will not allow genuine access to the world views of members of a social setting or of people sharing common attributes are likely to favour an unstructured interview.
>
> If the researcher is beginning the investigation with a fairly clear focus, rather than a very general notion of wanting to do research on a topic, it is likely that the interviews will be semi-structured ones, so that the more specific issues can be addressed.
>
> If more than one person is to carry out the fieldwork, in order to assure a modicum of comparability of interviewing style, it is likely that semi-structured interviewing will be preferred...
>
> If you are doing multiple-case study research, you are likely to find that you will need some structure in order to assure cross-case comparability. Certainly, all my qualitative research on different kinds of organization has entailed semi-structured interviewing, and it is not a coincidence that this is because most of it has been multiple-case research...
>
> (Bryman, 2008, 439-440)

As pointed out by Bryman, the semi-structured interview is able to assure data are comparable across cases. It also allows the researcher to follow the flow of each individual interview, which enables interviewees to open up their conversation and let

data emerge. Hence, the semi-structured interview was selected as the data collection method for this research. However, as shown in Figure 1, the degree of structure of semi-structured interviews varies.

The design of interview schedules was shaped by a pilot study conducted in 2011 and by the inputs of the examiners of the earlier Transfer of Status draft of the present work. In total, three interview schedules were produced. The first interview schedule was created and used for the pilot study. The second interview schedule was produced in the form of a questionnaire, based on the lessons learned from the pilot study. Although it was made with the intention of collecting data in the main study, it was not used due to the production of the third interview schedule. The third interview schedule was designed by taking account of the suggestions made by the examiners of the Transfer of Status.

Before the pilot study, the researcher was not aware of which existing quality assurance mechanisms were in place at particular universities. Hence, the semi-structured interview schedule was loose in its structure. According to the first interview schedule, the researcher provided interviewees with five elements – the teacher, students, the university, the government and the society – and asked interviewees to draw a diagram to indicate the relationships between these elements in regards to the quality of teaching. Then, the researcher probed regarding quality assurance mechanisms and how they influence interviewees' work. Ten interviews were conducted at two universities in China. Based on the data collected from the pilot study, the researcher produced the second interview schedule – a one-page questionnaire (see Appendix 2), which listed various quality assurance mechanisms.

Interviewees would need to rate each quality assurance mechanism according to their perceived influence. The rationale behind the use of questionnaires stems from the fact that they were frequently used in China to collect data; moreover, many interviewees during the pilot study suggested that they expected the researcher to give them a questionnaire. The intention behind the use of the questionnaire was not to collect quantitative data; instead, the questionnaire was intended to open up a conversation. The questionnaire was tried with two professors in China.

The use of a questionnaire may look firmly structured but follow-up questions can mean interviews move more into the territory of an unstructured interview. Due to this reason, the examiners at the Transfer of Status stage suggested that the researcher produce a full interview schedule, specifying the major questions. After several revisions, the third interview schedule, which was used in the main study, was produced and agreed upon by the examiners. The third interview schedule has two versions: one for teachers and the other for the department heads (see Appendix 3). The third interview schedule was subsequently translated into Chinese and used for data collection.

3.4 Case Selection

> Perhaps the most unusual aspect of case study in the social sciences and human services is the selection of case to study.
>
> (Stake, 2005, 450)

As mentioned above, to maximise the opportunity to learn the researcher decided to adopt a multiple case study approach, this is, a case study that comprises multiple cases. Yin (2003) argues that a case can be an individual, an event or an entity. Influenced by Smith (1978), Stake defines case as '"a bounded system", drawing attention to it as an object rather than a process' (Stake, 1995, 2). He explains:

> The crux of the definition is having some conception of the unity or totality of a system with some kind of outlines or boundaries.
>
> (Stake, 1988, 255)

In this study, universities are defined as cases, because they are bounded systems. In China, universities have their own campuses, which are surrounded by walls, so they have clear geographical boundaries. In addition, universities also have their contextual boundaries. Every university has its own history and culture. More importantly, universities have their own organisational boundaries. They make their policies in response to national policies. Every aspect of universities is influenced by university policies, including staff and student intake, allocation of resources, facilities or teachers' perceptions. This is the fundamental reason why this study chooses universities as the cases.

3.4.1 Selection of the Universities

The intention of this research was to select different kinds of universities to see whether there are different responses to the same national policies. However, there are various ways that universities differ and universities can be differentiated by location or by classification.

The pilot study confirmed that different geographic and economic regions had an influence on teachers' perceptions. It was conducted in two different cities in China. University A is in a capital city of a less economically developed province in the west of China, while University B is in a more economically developed city in the east of China. In the 2011 ranking of universities compiled by the China University Alumni Association, the two universities' ranks were close to each other. At University A many interviewees were unmotivated because of low salaries, while at University B interviewees were content with their salaries and thought their efforts were

acknowledged by the university. In terms of classification, Yu's (2010) research shows that universities at different levels in the classification of Chinese universities enjoy different reputations and resources, thus attracting different kinds of students.

The researcher intended to select universities that differ in a number of ways, but due to the scope of this research only one factor that differentiates universities was included in this research. If both regions and levels in the university classification were included in the main study, it would have been difficult to identify which features contribute to the differences. Therefore, the researcher had to make a choice between selecting different universities from the same region or similar universities from different regions. Choosing universities from the same region would enable the exploration of the differences between universities positioned at different levels; choosing universities from different regions would enable an exploration of the differences created by imbalanced economic development. Because quality assurance is initiated by government policy and the university classification is also a product of government policy, it is sensible to keep the scope of this research within a policy scenario, rather than extending it to encompass imbalanced economic development between the east and the west of China.

The case study universities were selected based on the conventional classification of universities in China – affiliation to Project 985 and Project 211. As was stated in the Introduction, the first case university is referred to as the 'the Project 985 University' (Uni985) in this research, and is affiliated with both Project 985 and Project 211. The second case university is called the 'the Project 211 University' (Uni211), and

participates only in Project 211. The third university is named the 'City University' (UniCity), and is excluded from both Project 985 and Project 211.

In China, some universities specialise in sciences, some in teacher education, some in agriculture and some in arts. Others are comprehensive universities. To restrict the differences between universities to one factor, the three case universities were selected from the same speciality category. However, the speciality category cannot be revealed here in order to protect the anonymity of the interviewees, given that there are only a few regions in China host to universities that belong to this category but differ in classification.

The region in which the selected universities are is an economically well-developed region in China. This fact shapes the macro context of this study. This macro context has particular implications, which in turn means that findings generated from this study may not be applicable to universities located in economically less developed regions. However, it gives this study a new perspective, enabling the researcher to investigate how universities and teachers in an affluent environment respond to national policies.

3.4.2 Selection of Departments

Yin (2003) categorises multiple case studies into multiple holistic and embedded case studies. At university, teachers are clustered into departments. This feature determines that this research is a multiple embedded case study. Thus, departments are the embedded units of analysis.

Following the same rationale of assuring diversity, this study selected three departments at each university: the humanities, the sciences and the social sciences. The pilot study provided another criterion to select departments. When selecting departments, different sources of income are considered because they can influence teachers' choices when prioritising competing tasks. In a department that easily attracts research grants, the major task competing with teaching is likely to be research; in a department in which academics are offered many part-time opportunities outside the university, the major competing task is likely to be a part-time job; in a department that does not attract research grants or part-time job offers, the temptation to prioritise other competing tasks is likely to be weaker. This design is regarded a strength of this research because it enables the researcher to assess whether the sources of income affect teachers' views.

Applying the two criteria, three departments were selected for this study. The department that was selected in the humanities is referred to as the 'Humanities Department' (DepHum). It is the department that does not easily attract grants or offers of part-time jobs. The department selected in the sciences is referred to as the 'Sciences Department' (DepSc). It is the department that is offered generous funding. The department selected from the social sciences is referred to as the 'Social Sciences Department' (DepSoSc). It is the department where academics are often offered part-time opportunities outside the university. The same departments were selected across the three universities.

3.4.3 Selection of Interviewees

The main study initially intended to recruit one professor, one associate professor and two lecturers, with an aim to diversify interviewees with different academic titles, as

well as two department heads in each department. The department heads in fact are teachers who conduct teaching, but they have some extra management responsibility. In each department, there is one department head and several deputy heads. The department head oversees the working of the department and each deputy head is responsible for a particular area, such as research, graduate education or undergraduate education. In this research, the term ‘department head’ refers to both the head and the deputy head.

In the process of gaining access, the department head and the deputy head responsible for undergraduate education were contacted first. If they refused to participate in this study, other heads were contacted. When recruiting interviewees with the same academic title, the researcher attempted to recruit a diverse group. Selection criteria included different research records, records of winning teaching- or research- related awards and experiences of being involved in government quality assurance initiatives. Initially, the researcher contacted 54 interviewees at nine departments at three case study universities. However, in practice the recruited interviewees differed from the plan. The profile of recruited interviewees is listed in Table 4.

University	Department	ID	Title	Gender	Years of Teaching Experience
The Project 985 University	The Humanities Department	Head A	Head/Professor	Male	12
		Head B	Head/Professor	Male	20
		Lecturer A	Lecturer	Female	5
		Associate Professor A	Associate Professor	Male	8
		Associate Professor B	Associate Professor	Male	2
		Professor A	Professor	Female	28
	The Sciences Department	Head C	Head/Professor	Male	19
		Head D	Head/Professor	Male	15
		Lecturer B	Lecturer	Male	3
		Lecturer C	Lecturer	Male	2
		Associate Professor C	Associate Professor	Male	6
		Professor B	Professor	Male	30
	The Social Sciences Department	Head E	Head/Professor	Male	12
		Head F	Head/Professor	Female	23
		Lecturer D	Lecturer	Male	1
		Lecturer E	Lecturer	Female	26
		Associate Professor D	Associate Professor	Male	9
		Professor C	Professor	Male	25
The Project 211 University	The Humanities Department	Head G	Head/Professor	Male	39
		Head H	Head/Associate Professor	Male	9
		Lecturer F	Lecturer	Female	20
		Lecturer G	Lecturer	Female	21
		Associate Professor E	Associate Professor	Male	6
		Professor D	Professor	Male	10
	The Sciences Department	Head I	Head/Professor	Male	26
		Head J	Head/Professor	Male	4
		Lecturer H	Lecturer	Female	7
		Associate Professor F	Associate Professor	Female	8
		Associate Professor G	Associate Professor	Female	11
		Professor E	Professor	Female	26
	The Social Sciences Department	Head K	Head/Associate Professor	Male	17

		Lecturer I	Lecturer	Female	3
		Lecturer J	Lecturer	Male	26
		Associate Professor H	Associate Professor	Female	21
		Associate Professor I	Associate Professor	Female	29
		Professor F	Professor	Female	16
		Professor G	Professor	Male	26
City University	The Humanities Department	Head L	Head/Professor	Female	27
		Head M	Head/Professor	Male	24
		Lecturer K	Lecturer	Male	21
		Associate Professor J	Associate Professor	Female	13
		Associate Professor K	Associate Professor	Male	3
		Professor H	Professor	Male	24
		Professor I	Professor	Male	24
	The Sciences Department	Head N	Head/Professor	Male	20
		Head O	Head/Professor	Male	20
		Lecturer L	Lecturer	Female	15
		Lecturer M	Lecturer	Male	3
		Associate Professor L	Associate Professor	Female	23
		Professor J	Professor	Male	20
	The Social Sciences Department	Head P	Head/Professor	Male	8
		Head Q	Head/Professor	Male	18
		Lecturer N	Lecturer	Female	10
		Lecturer O	Lecturer	Male	15
		Professor K	Professor	Male	20
		Professor L	Professor	Male	27

Table 4 Profiles of interviewees in main study

Table 4 shows that, at every department, two department heads were successfully recruited except at the Social Sciences Department at the Project 211 University. The researcher invited other heads at this department but they all refused to participate in the research. At some departments, the pool of lecturers was very small. After the attempt to recruit the lecturers failed, the researcher recruited associate professors instead. At some departments the number of recruited interviewees exceeded the

original plan. Consequently, the profile of recruited interviewees is different from the original plan but still meets the selection criteria in general. In total, 56 interviews were conducted.

3.5 The Analytical and Conceptual Frameworks

In terms of the analytical framework, the analysis of national and university policies is guided by Knoepfel et al.'s (2007; 2011) framework for public policy analysis. Knoepfel et al.'s framework enables the researcher to present a whole picture of policy-making and implementation and thereby to identify the problems in policy-making and to highlight the changes that occurred during the implementation. The conceptual framework is built upon Schneider and Ingram's (1990) classification of policy tools. Inspired by Schneider and Ingram's classification, the researcher intends to outline a pattern of the impact of different policy tools based on the empirical evidence collected in this study. This proposed conceptual framework would be able to lift the understanding of the impact of quality assurance policy on individuals to a theoretic level.

3.5.1 Knoepfel et al.'s Framework for Public Policy Analysis

Knoepfel et al.'s (2007; 2011) framework is published in their book *Public Policy Analysis*. Dunstan (2008, 518) praises their work in the following terms: 'what this book does exceptionally well, whilst remaining theoretically neutral, is present an analytically and conceptually sophisticated account of public policy'. The sophistication of Knoepfel et al.'s framework firstly lies in conceptualising the dynamic between *actors, resources* and *institutional rules*. Knoepfel et al.'s framework identifies *actors* in the policy arena and provides an exhaustive list of all of the possible *resources* that can be processed and used by the various *actors*. Moreover, Knoepful et al. propose the concept of *institutional rules*, in other words,

the context shaping the policy-making and implementation process. According to the dynamic between *actors, resources* and *institutional rules,* it can be concluded that the *resources* owned by various *actors* determine the *institutional rules* and subsequently the roles played by *actors* in the policy-making and implementation.

According to Knoepfel et al. (2007; 2011), *actors* consist of *political-administrative authorities*, *target groups*, *end beneficiaries* and *indirectly affected third parties*. The *political-administrative authorities* are the policy-makers who are in the public service sphere. 'The *target groups* are made up of people (individual or legal entities) and organisations whose behaviour is required to change' (Knoepfel et al., 2007, 53). '*End beneficiaries* are actors who benefit, more or less indirectly and according to the intended objectives of the policy concerned, from the target groups' altered behaviours' (Knoepfel et al., 2007, 53). 'Finally, *the third-party groups* comprise all people (individuals or legal entities) and the organisations representing their interests who, without being directly targeted by the policy, see their individual and/or collective situation altered permanently' (Knoepfel et al., 2007, 53).

In this research, the *political-administrative authorities* are the Communist Party of China (CPC) and the State Council (the central government). The CPC is the only ruling party in China and it controls all legislative, administrative and judicial departments. Its administrative institution is the State Council and the Ministry of Education is a department of the State Council. The three institutions in fact represent the same interest and they are regarded as the same entity in this research. When discussing university policies, universities are also *political-administrative authorities.* The *target groups* can be universities and/or teachers. For example, the

target groups of the UTE are universities and teachers. The *target groups* of the Quality Project awards are teachers. The *end beneficiaries* of the two policy initiatives are students. The *third-party groups* can be companies that are contracted for refurbishment of buildings or construction of websites. However, the actors this research is concerned with are *political-administrative authorities* and *target groups*, i.e. the CPC or the State Council, universities and teachers.

According to Knoepfel et al. (2007; 2011), *resources* can be *legal resources, human resources, financial resources, information, organisation, consensus, time, infrastructure, political support* and *force*. Different *actors* possess various *resources*. What is striking in Chinese higher education is that the CPC or the State Council owns most of the *resources* apart from *consensus* and it uses them to create the national *institutional rules*. Universities have limited *legal, human, financial resources* and *organisation*. Teachers have few *resources* at their disposal. Various *resources* processed by different *actors* affect the *institutional rules* and consequently the whole policy cycle.

Institutional rules are 'both structures and rules that are formal, explicit and generally legally formalised and as informal norms that are implicit but shared by the members of an organisation and community' (Knoepfel et al., 2007, 53). Because both the State Council and universities are *political-administrative authorities*, the researcher proposes to distinguish national *institutional rules* and university *institutional rules,* which are made by the State Council and universities respectively. The dominant feature of national *institutional rules* discovered in this research is the strong control of the CPC and the State Council over higher education. University *institutional rules*

comprise rules made by the national *institutional rules* and university leaders. University *institutional rules* can differ at different universities. These *institutional rules* contribute to the contexts of the universities and potentially account for policies and practices at universities.

The model proposed by Knoepfel et al. (2007; 2011) treats policy as a policy cycle, starting with *agenda setting* and moving to *programming, implementation* and then *evaluation* (see Figure 2) for the purpose of analysis. At each stage, there are policy products being created under the influence of *institutional rules, actors* and *resources*.

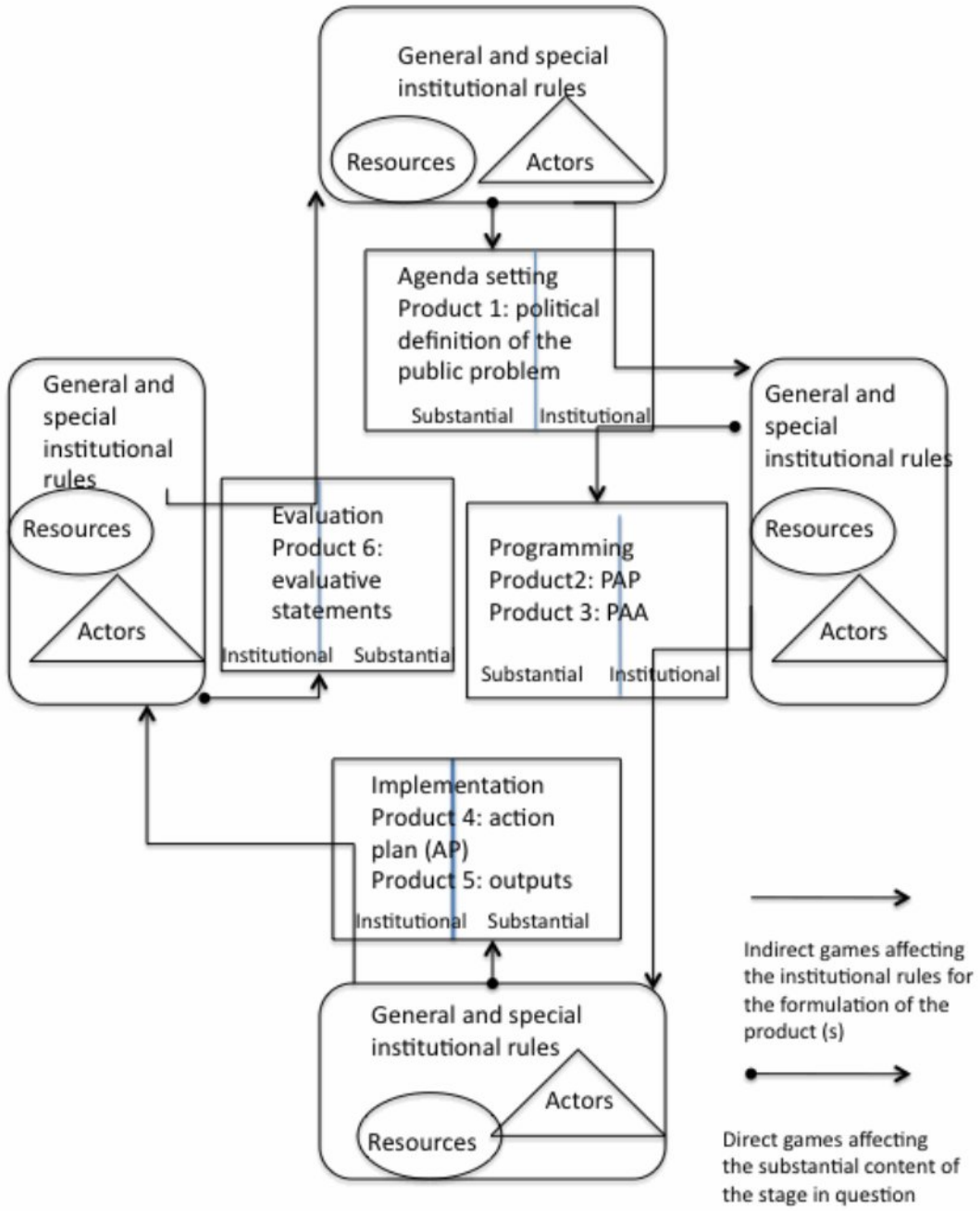


Figure 2 Theoretical model for the analysis of a public policy

(Extracted from Knoepfel et al., 2011, 126)

At the stage of *agenda setting*, the first policy product – *political definition of the public problem* – is created. There are many problems that *actors* face. However, not all of these problems can enter the policy arena. Problems with no social recognition

are called *problematic private situations*. When these *private problems* receive public attention but are not put on the political agenda, they become *social problems*. When these *social problems* receive political attention but no public policy is adopted or implemented, these *social problems* become *public problems*. When these *public problems* are addressed, public policy emerges. Consequently, the *political definition of the public problem* is achieved. In this process, some problems are left out and some problems are selected and addressed. This mechanism is called the *filtering mechanism*.

At the stage of *programming*, there are two products: the *political-administrative programme (PAP)* and the *political-administrative arrangement (PAA)*. 'The provisions of a *PAP* comprise the normative objectives of the solution envisaged for the resolution of the problem, the definition of *target groups* and their roles during the realisation of the policy (hypothesis of causality), the means made available for this purpose (intervention hypothesis) and the principles of the administrative organisation of the policy implementation' (Knoepfel et al., 2007, 53). The *PAP* should have five constituent elements: *objectives, evaluative elements, operational elements (intervention instruments), the PPA* and *resources* and *procedural elements*.

At the *implementation* stage, *action plans* and *implementation acts* should be produced. *Action plans* are more detailed and guide the implementation of the policy. *Implementation acts* are a set of decisions made by the public actors so as to affect the *target groups*.

At the *evaluation* stage, evaluation statements should be produced to assess *impact, outcome, relevance, effectiveness* and *efficiency*. *Impact* is the influence on the *target groups*; *outcome* is the influence on the *final beneficiaries*. *Effectiveness* describes whether the impacts contribute to the outcomes; *efficiency* means whether the *outcomes* are achieved with minimal *resources*. *Relevance* means whether *outcomes* contribute to solving the *social problem* defined at the beginning.

The sophistication of Knoepfel et al.'s framework secondly lies in its scope. It places the analysis of one policy against the backdrop of the big picture. It reminds policy analysts of the influence of parallel agendas. The analysis of a particular policy should not be limited to what is included but also must consider what is left out. As described above, *agenda setting* is the process of filtering. However, the *filtering mechanism* is not confined to *agenda setting*; rather Knoepfel et al. (2007; 2011) argue that it exists at each stage of the policy process. At the stage of *programming*, some *public problems* will be filtered out and others will be addressed by the *PAP* and the *PAA*. The *filtering mechanism* also works at the stage of *implementation*, filtering some issues in the *PAP* out. The rest of the issues are tackled by the *action plan*.

As Knoepfel et al.'s (2007; 2011) model is not tailored for analysing quality assurance policies and the impact on Chinese higher education, it has some inadequacies and redundancies when applied to this task. As a result, the researcher has amended Knoepfel et al.'s framework to meet the needs of this study. In fact, amendments of their model are welcome by the authors. They encourage researchers 'to revisit the arguments presented, develop them further, complement them with other theoretical approaches and apply them in an actual analysis situation' (Knoepfel et al., 2007, 53).

The first amendment of Knoepfel et al.'s (2007; 2011) framework is caused by the multiple affiliations in the case at hand. Universities are both *target groups* and *political-administrative authorities*. Universities are the *target groups* who implement government policies but they also make their own policies in terms of targeting teachers in response to government policies. Hence, unlike the model proposed by Knoepfel et al., which has only one *agenda setting* and *programming*, there are two *agenda settings* and *programmings.* One is the *agenda setting* and *programming* made by the State Council and the other is made by universities.

The second amendment is a simplification of Knoepfel et al.'s (2007; 2011) model. Their framework is very complex but this research will not use all its elements. Instead, the current research will keep the key definitions (*actors, resources and institutional rules*) but will simplify the policy cycle. As mentioned earlier, Knoepfel et al. divide the policy cycle into *agenda setting*, *programming, implementation and evaluation.* What they mean by *implementation* is the process by which *political-administrative authorities* make *action plans* and *implementation acts*. In other words, it is the stage of making detailed decisions and policies to implement the *PAPs*. In the researcher's opinion, the *implementation* as it is defined by Knoepfel et al. (2007; 2011) is still a part of policy-making. The researcher contends that the term 'implementation' should be understood in reference to its literal meaning, i.e. what *target groups* do to implement the policy in order to have a term to describe the 'implementation action' as it pertains to policy.

Therefore, the simplified policy cycle consists of two parts: policy-making and implementation. The two terms retain their literal meanings. Policy-making is the stage of making policy; implementation is the stage of implementing policy. The policy-making is divided into *agenda setting* and *programming.* The *agenda setting* remains the same. The *programming* used by this thesis is a combination of Knoepfel et al.'s *programming* and their *implementation*, the policy products of which include the *PAPs,* the *PAAs* and the *action plans*. Knoepfel et al.'s (2007; 2011) *evaluation* is absent in this research because it is often omitted by in national policies.

3.5.2 Schneider and Ingram's Classification of Policy Tools

Although Knoepfel et al. (2007; 2011) make reference to *intervention instruments*, they do not elaborate on this. According to a widely cited classification of policy tools in the political sciences proposed by Schneider and Ingram (1990), policy tools can be categorised into *authority, incentive, capacity-building, symbolic and hortatory or learning tools*.

Authority tools are 'simply statements backed by the legitimate authority of government that grant permission, prohibit, or require action under designated circumstances' (Schneider and Ingram, 1990, 514). These are tools that 'rely heavily on authority and assume that lower-level agents and targets are motivated by a desire to follow orders'(Schneider and Ingram, 1997, 93).

Incentive tools 'rely on tangible payoffs, positive or negative, to induce compliance or encourage utilization' (Schneider and Ingram, 1990, 514). These tools assume that

'people will not be motivated to take policy-relevant action unless there is something to gain or lose' (Schneider and Ingram, 1997, 93).

'*Capacity-building tools* provide information, training, education and resources to enable individuals, groups, or agencies to make decisions or carry out activities. These approaches that assume incentives are not an issue, but there may be barriers stemming from lack of information, skills, or other resources needed to make decisions or take actions that will contribute to policy goals' (Schneider and Ingram, 1990, 514).

'*Symbolic and hortatory tools* assume that people are motivated from within and decide whether or not to take policy-related actions on the basis of their beliefs and values' (Schneider and Ingram, 1990, 514). Furthermore, 'These tools aim either to change the values of agents or targets or to convince them that the behaviour needed to achieve policy objectives is consistent with values already held' (Schneider and Ingram, 1997, 93).

'*Learning tools* are used when the basis upon which target populations might be moved to take problem-solving action is unknown or uncertain…One of the most important characteristics of learning tools is that they assume agencies and target populations can learn about behaviour, and select from the other tools those that will be effective' (Schneider and Ingram, 1990, 514).

Although one policy can have some overlapping elements of different policy tools, this thesis identifies policy by its most distinctive feature. The most distinctive feature

should be the necessary and sufficient condition for a policy's implementation. It can be identified by eliminating other features. For instance, the most distinctive feature of *authority tools* is their compulsoriness. Policy implementers have to implement the policy despite the lack of incentives and of the prospect of building their capacities. They implement the policy because the *political-administrative authorities* use their authority to force them to do so. The most distinctive feature of *incentive tools* is the motivation provided by incentives. A policy can be identified as an *incentive tool* when the policy is not compulsory and policy implementers will not react to it unless they are incentivised by either financial or reputational rewards or by both. In terms of the quality assurance policies at hand, the UTE is identified as an *authority tool* because it is compulsory. The Quality Project awards are regarded as an *incentive tool* because they serve to motivate teachers through either financial or reputational rewards.

3.5.3 A Proposed Understanding of Impact

Inspired by Schneider and Ingram's (1990) classification and the literature reviewed in Chapter Two, the researcher proposes four types of impact: *implementation actions, interactions, consequences* and *reactions*, as illustrated in Figure 3. They will be used to guide the analysis of interview data. *Implementation actions* are actions taken by the *target groups* or other actors to implement quality assurance policies. *Interactions* are actions that enable the results of the *implementation actions* to reach the *target groups*, which complete the feedback loop for the *implementation actions*. *Consequences* are the outcomes of *implementation actions* or the *interactions* which have occurred. *Reactions* are actions taken by the *target groups* in response to the *consequences*.

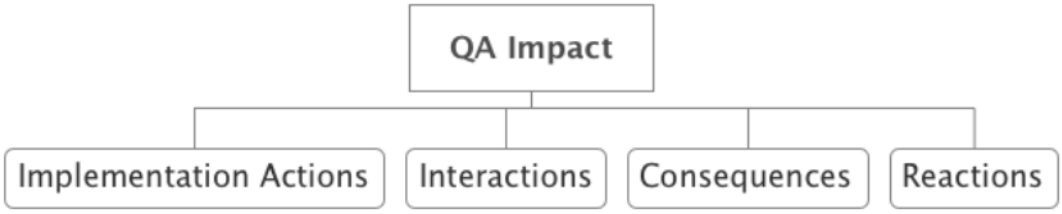


Figure 3 A proposed understanding of impact

The relationship between the four types of impact remains to be investigated. Several questions arise: if the *target groups* did not execute the *implementation actions*, would there be *consequences*? if there were no interactions, would there be *consequences*? were there always *consequences*? if not, what would be the reasons? is it possible to find general patterns of *implementation actions, interactions, consequences* and *reactions* of different policy tools? These questions will be explored based on the data collected by this research.

3.6 Data Analysis

In this research, both policy documents and interviews were analysed by thematic analysis. Thematic analysis can be defined as 'analysis based on the identification of themes in qualitative material' (Seale, 2004, 509). It can also be regarded as:

1. A way of seeing
2. A way of making sense out of seemingly unrelated material
3. A way of analyzing qualitative information
4. A way of systematically observing a person, an interaction, a group, a situation, an organization, or a culture
5. A way of converting qualitative information into quantitative data

(Boyatzis, 1998, 4-5)

Boyatzis (1998) identifies three coding approaches: theory driven, prior-research driven and data driven. These differ 'in the degree to which the thematic analysis starts with a theory or the raw information' (Boyatzis, 1998, 29).

In this research, all three coding approaches proved useful. The coding approach for policy documents was theory driven, following Knoepfel et al.'s (2007; 2011) framework. National and university policy documents were categorised into documents related to *agenda setting* and documents related to *programming*, such as the *PAP* and *action plans*.

Some interview coding approaches were prior-research driven. Before the interview questions were designed, the initial literature review showed that the two major national quality assurance policies are the UTE and the Quality Project, so each policy became a theme.

Moreover, in order to answer the research questions, the interview questions were structured by the following themes, which were pre-research driven:

- Teachers' definitions of high quality teaching;
- What helped teachers to improve their teaching;
- What motivated teachers to improve their teaching;
- What were the quality assurance mechanisms that they thought had some influence on their teaching;
- The influence of the UTE;
- The influence of the Quality Project;
- Suggestions regarding current quality assurance mechanisms.

A data-driven coding approach was also utilised in the coding process. The internal quality assurance mechanisms repeatedly mentioned by interviewees were class

observation, the student evaluation of teachers and teacher appraisal, so they became important themes as well.

All 56 interviews were transcribed in full in the original language – Chinese. Thus, the analysis was conducted in Chinese in order to preserve the original meanings of interview data. When accounts from teachers and heads were quoted, these accounts were then translated into English. Data from policy documents and all 56 interview transcripts were then grouped by themes and further analysed using the method suggested by Ryan and Bernard (2003). They recommend looking for repetition, indigenous typologies or categories, metaphors and analogies, transitions, similarities and differences, linguistic connectors, missing data and theory-related material. Comparing similarities and differences was the most commonly used method. The categories of themes are listed in Appendix 4. As illustrated in the following chapters, these themes not only contribute to answering the research questions but also guide the structuring of these chapters.

3.7 Research Ethics

The research complied with the ethics policy of the University of Oxford. Before the pilot study, an application was submitted to the Social Sciences Division Interdisciplinary Research Ethics Committees, as part of the Central University Research Ethics Committee (CUREC). After the researcher demonstrated how the ethical requirements and procedures would be fulfilled, permission was granted. Subsequently, the pilot study and the fieldwork were carried out according to the university's ethical policy. The CUREC forms are included in Appendix 5. The most relevant ethical issues – in this case, access and anonymity – are discussed below.

3.7.1 Negotiating Access

Gaining access was crucial to this study. Access had to be negotiated with the 'gatekeeper' of each individual department, in particular the department head. The first attempt to contact the heads of nine departments was via email, which was because the researcher was based in the UK. Unfortunately, this failed as none of the nine department heads replied. This was not unexpected, given the fact that personal contact plays a very important role in the Chinese cultural context. People are naturally suspicious when a stranger approaches them for an interview, let alone by email. Hence, the researcher decided to travel to China to negotiate access in September 2012. There were two options to establish contacts. The approach in tune with Chinese culture would be to approach the department heads through several contacts. The abrupt approach with a high risk of failure would be to contact the department heads directly in person or by telephone.

The researcher decided to use the more traditional Chinese approach first. After going through four or five contacts, the researcher reached the department heads. If one department refused to grant the researcher access, this would have compromised the research intentions, because the study had to be based on the very same departments across three universities. With honesty and openness, the researcher introduced herself and her research and presented the information sheet, consent form and proposed interview schedule (see Appendix 6 for the information sheet and consent form). In the end, nine department heads agreed to the interview. The department heads' agreement to interview meant that the researcher gained access to all nine departments.

For each interview with teachers, the researcher contacted every potential interviewee to get permission to conduct the interview. Not all teachers' names and contact details were listed on the departmental or university websites. Some potential interviewees' names and contact details were obtained from the website directly, while others were obtained from the department heads. An email to introduce the researcher and the research with the information sheet, consent form and the interview schedule was sent to each potential interviewee. When potential interviewees' contact details were not available, the researcher waited outside the individual's classroom to introduce herself and the project. The printed information sheet, consent form and interview schedule were provided in order to help the potential interviewee make an informed decision. Due to time constraints, contacts were first made with those whose contact details were available. Classroom visits were used as the last resort. The criteria to recruit diverse interviewees were also applied throughout the selection process.

Besides interview data, the researcher planned to collect university policy documents that are not available in the public domain. The researcher requested access to three universities' archives but all three requests were denied because she was affiliated with a foreign university. However, when the researcher resorted to university libraries, many librarians gave her access to university yearbooks that were not available in the catalogue. Hence, the university yearbooks became another major source of university policy detail and provided much background information about these universities.

3.7.2 Anonymity

The researcher endeavoured to protect interviewees' anonymity. None of the names of interviewees, the departments or of the universities were revealed at any stage. As

illustrated in Table 4, interviewees' names were replaced by letters; departments were referred to by their generic names and the universities were referred to by their classification. The location of the universities was not revealed. When interviewees mentioned other teachers' names, the names of a department or the names of universities, these names were anonymised in the same way. When citing interviewees' accounts, any information that had the potential to reveal any identity was replaced with more generic wording.

3.8 Limitations

The major limitations of this study included the potential biases and generalisations, which are discussed separately in the following sections.

3.8.1 Potential Biases

Despite the researcher's best efforts to represent diverse opinions by recruiting a broad range of interviewees, the fact that she could only interview individuals who accepted her invitation for interview introduced certain biases. Accepting the invitation suggests that the interviewee is interested in the research topic and recognises the importance of the quality of undergraduate education. Furthermore, interviewees who accept the request might be more likely to be confident teachers. However, it does not mean that they are necessarily 'better' teachers than others. Some interviewees admitted that they had received negative comments from students or were not good at communicating with students. It was difficult to draw any conclusions about interviewees' teaching given the multifaceted nature of teaching quality and the fact that the researcher did not conduct classroom observations.

On the other hand, by not accepting the invitation for interview, those who were not interested in undergraduate education or who were not confident in teaching excluded

themselves from this study. In addition, paradoxically, some teachers, such as Teacher Q, who were enthusiastic about undergraduate teaching declined the researcher's invitation maybe because they were deeply frustrated by a dominance of performance-based culture that emphasises research.

The inclusion of a certain group of interviewees suggests that the perceptions extracted from interview data were inclined to reflect the opinions of teachers who were interested in teaching undergraduates or had confidence in their teaching. As a result, this research is more relevant to addressing the issues concerning this group of teachers. In fact, this group of teachers can be the main *target group* of government policy. If government policies cannot influence a group of teachers who are interested in undergraduate teaching, then such policies are unlikely to influence anybody else. Thus, investigating how government policy influences this particular group of teachers represents a sensible place to start when it comes to improving quality assurance policies.

The researcher is also aware of the potential biases introduced by the researcher's own personality. To follow the ethical procedure, the researcher needed to ask every interviewee to sign a formal written consent form. Most interviewees were unfamiliar with this practice. In the pilot study, some interviewees were quite cautious about it because they perceived the signature as legally binding and thus as holding them responsible for their speech and actions. In the main study, the requirement for a signature did cause concern among certain potential interviewees who then refused to participate in the research. Nevertheless, for those who accepted the interview invitation, the ethical procedure demanded by the University of Oxford reinforced

their confidence in the researcher. The researcher's gender (being a female) and her identity as a student plus her frankness and honesty reassured the interviewees that participating in this research would not cause any harm to come to them. The fact that interviewees' accounts, including the accounts of the department heads, contain criticisms of current policy underlined the openness and frankness of interviewees.

3.8.2 Generalisability

Where case studies are concerned, the issue of generalisation often comes to the fore. Advocates of the case study approach admit that case studies are weak in producing certain types of generalisations, i.e. 'statistical generalisations' (Yin, 2003), 'predictive generalisations' (Stenhouse, 1978) or 'propositional generalisations' (Stake, 1995). Yin (2003) describes the practice of making an inference about a population based on empirical data collected about a sample as a statistical generalisations. Because generalisations tend to take the form of assertions, Stake (1995) names this type of generalisation 'propositional generalisations'. Stenhouse (1978) emphasises the role that generalisations play in making predictions, so he calls them 'predictive generalisations'. Stake (1995, 85) points out that 'single cases are not as strong a base for generalizing to a population of cases as other research designs'. In short, the purpose of case studies is neither to make generalisable assertions nor to make predictions.

Instead, these advocates point out that case studies are designed to achieve other kinds of generalisations. Yin (2003, 10) argues that the goal of case studies is 'to expand and generalize theories (analytic generalization) and not to enumerate frequencies (statistical generalization)'. Stenhouse (1978) distinguishes retrospective generalisations, which case studies aim to achieve, from predictive generalisations:

> Retrospective generalisations are attempts to map the range of experience rather than to perceive within that range the operation of laws in the scientific sense.
>
> (Stenhouse, 1978, 22)

> While predictive generalisations claim to supersede the need for individual judgement, retrospective generalisations seek to strengthen individual judgement where it cannot be superseded.
>
> (Stenhouse, 1978, 22)

Furthermore, Stake (1995) proposes the concept of naturalistic generalisations:

> Naturalistic generalizations are conclusions arrived at through personal engagement in life's affairs or by vicarious experience so well constructed that the person feels as if it happened to themselves.
>
> (Stake, 1995, 85)

Stake (1995) also argues that, to optimise readers' opportunity to learn, researchers should provide readers with 'thick description'. As illustrated above, case studies intend to achieve analytical, retrospective or naturalistic generalisations and the purposes of conducting case studies are to develop theory, document experiences, strengthen judgement, advance understanding, or promote learning.

Similar to other case studies, this research is inadequate in terms of achieving 'statistical generalisation' (Yin, 2003), 'predictive generalisation' (Stenhouse, 1978) or 'propositional generalisation' (Stake, 1995). Due to the constraints of time and resources, it was impossible for the researcher to achieve a statistically representative sample. Despite the researcher's effort to maximise the diversity of interviewees, it is necessary to remind readers that the findings of this research are based on the perceptions of 56 interviewees. When the researcher uses the term 'teachers', it should only be taken to refer to the regular teachers who were interviewed. This applies to the use of the term 'heads' too. The claims regarding the case study universities and departments are also drawn from the available data, so the researcher acknowledges that findings could be different if different data were collected.

Moreover, whether the findings of this research can be generalisable to different universities remains an open question.

Nevertheless, this research will be able to achieve retrospective and naturalistic generalisations. To achieve them, the researcher endeavours to keep the research as transparent as possible without endangering anonymity, to preserve the completeness of interviewees' accounts, and to present the diversity of interviewees' opinions. These efforts should enable readers to gain insight into interviewees' experiences and to form their own judgements about the impacts of quality assurance policies and factors that contribute to enhancing the quality of teaching.

This research alone is unable to achieve analytic generalisation as defined by Yin (2003). To achieve the analytic generalisation, the findings of this research have to be supported by other studies. This work will be attempted in Chapter Eight with the aim of achieving the analytic generalisation.

4 National Policies

The national *institutional rules* are introduced first in order to provide a background to facilitate readers' understanding. The analyses of the two major national quality assurance policies will be followed by the simplified policy cycle, consisting of the *agenda setting* and the *programming*. The analysis of the UTE will be presented first, followed by that of the Quality Project.

4.1 National Institutional Rules

The term '*national institutional rules'* denotes legal norms, formal and informal procedures, social norms and cultural values established by the ruling party and its government agencies to limit other actors' access to resources, to guide other actors' behaviours and to define political and social legitimacy. This term is adopted from Knoepfel et al.'s (2007; 2011) framework for policy analysis rather than a term used by the ruling party or its government agencies in policy documents. *National institutional rules* can be understood as the context of national policies.

The *national institutional rules* of the governance of higher education in China are characterised by centralised control and fragmented regulative power. The centralised control of higher education gives the central government the power to make educational policies without consulting other actors. By using this right, the central government institutionalises its definition of quality (as efficiency) in policies. The fragmented regulative power makes the quality assurance policies issued by the Ministry of Education only focus on the quality of undergraduate teaching.

4.1.1 Central Control and Efficiency

Higher education in China is controlled centrally. The power relationship in higher education policy-making can be illustrated as a pyramid (Yan, 2006). The Communist Party of China (CPC) is situated at the top, followed by the legislative department, education administrative departments, and higher education institutions. The CPC is the governing party of China and oversees the legislative and administrative departments, which regulate higher education institutions. The National People's Congress is the highest legislative body and the State Council, i.e. the central government, is the highest administrative authority. Hence, the ultimate policy-maker is the CPC.

After the Cultural Revolution, the CPC initiated a series of reforms. After announcing the transformation of the Chinese economy from a planned economy to a market economy (CPC Central Committee, 1985), educational reforms were carried out, including reducing funding and sharing the financial burden with local governments. To compensate for the decreased funding to higher education, the State Council resorted to charging tuition fees. Higher education was free from 1949 to 1988. In 1989, the State Education Commission (the then Ministry of Education), the State Price Bureau and the Ministry of Finance issued the *Regulations for Collecting Tuition, Sundry and Boarding Fees in Ordinary Universities* (State Education Commission, 1989). The newly issued Higher Education Law (PRC President, 1998) also demanded that all students in higher education pay tuition fees according to regulations. As a result, the State Council was able to decrease the education budget dramatically. Based on data extracted from *Chinese Higher Education Statistics Year Books*, Yang and Zhao (2012) concluded that the contribution of the education budget

to the total funding of higher education dropped from 63% in 1999 to 43% in 2006 and the portion contributed by students including tuition and other fees increased from 17% in 1999 to 29% in 2006.

Reducing funding and charging tuition fees reflected the State Council's intention to improve the efficiency of the central government's funding. By doing so, the number of students enrolled in higher education increased despite the lower investment. Also, students could be turned into an educated workforce to supply the economy. Charging tuition fees is indicative of the shift from the communist ideology to a market-oriented ideology. Education became a commodity rather than a public service.

Another reform concerned sharing the financial burden with local governments. In China, universities were regulated by the central or local governments. The regulators had to provide funding for regulated universities. *An Action Plan for Invigorating Education Oriented to the 21st Century* reiterated the central government's intention to transfer the ownership of universities to local governments in order to reduce the expenditure on higher education coming from central government (Ministry of Education, 1998). By extracting data from *Chinese Higher Education Statistics Year Books,* Xu and Fang (2012) found that, in 1997, there were 1,020 regular higher education institutions[10] and 345 (34%) were governed by the central government; however, by 2001 this number had dropped to 111 (10%).

10 In Chinese higher education, there are some institutions mainly offering vocational and continuing education programmes and others mainly offering full-time undergraduate and graduate programmes. The latter type is referred to as regular higher education institutions by the government.

By transferring responsibilities to local governments, the State Council was able to ease its financial burden. The intention behind the move to share the financial burden with local governments was to improve efficiency. After transferring the ownership to local governments, the State Council regulated only elite universities. By doing so, the State Council was able to achieve a better return on its investment.

The quest for efficiency is amplified in the definition of quality. Quality has never been defined formally in policy documents; instead, a definition is only provided in speeches delivered by the ministers of education. Zhou Ji, the minister of education from 2003 to 2009, defined quality as follows:

> What is high quality education? I think high quality education is high employment rate. That the society and employers are satisfied with the graduates means high quality education.
>
> (Zhou, 2004, 19)

Zhou's definition is very market-oriented. According to his definition, quality is not judged by teachers, universities or students; it is judged by employers and measured by the employment rate. The purpose of education is only to provide a workforce. In fact, when the quality of higher education is equal to the employment rate, quality is therefore defined by efficiency.

Under such centralised control, the policy-making is problematic, especially regulatory policies. Regulatory policies are made and enforced by the State Council, while legislative policies come from the National People's Congress (Yan, 2006). In fact, most of the national policies mentioned in this research are regulatory policies. Reviewing the education policies over the past 50 years, Yan (2006) points out the following problems in regard to regulatory policies. Unlike as is the case with legislative policies, the State Council has total control over policy-making. The State

Council's policy-making process lacks transparency because the council does not need to consult other actors. Few people are aware of the policy-making and very few people can debate or question the policy. Without consultation many policies are hastily made and implemented.

While the state maintains a tight control over higher education, universities are left with limited autonomy. Before 1998 the State Council controlled almost every aspect of university affairs. After 1998, the Higher Education Law has given universities certain level of autonomy (PRC President, 1998). The majority of the issues on which universities have autonomy relate to teaching, including the rights to decide what subjects and programmes to teach, how lesson plans should be prepared, how textbooks should be compiled or selected, and how teaching-related activities should be organised. However, as discussed in later sections, these aspects often become the targets of the quality assurance mechanisms. Hence, quality assurance mechanisms are tools for the State Council to steer from a distance when it does not have a direct control.

4.1.2 Fragmented Power in Higher Education

Although higher education is controlled centrally by the State Council, regulative power is shared by different departments within the State Council and these different departments launch different quality assurance mechanisms. An overview of these quality assurance mechanisms will be provided in this section. As a result, the Ministry of Education is only responsible for the quality assurance of undergraduate teaching.

According to the Higher Education Law, 'the State Council leads and manages education in China. Local governments assist the State Council in managing the higher education institutions in the regions and those higher education institutions directly controlled by the State Council' (PRC President, 1998, 2). The highest education administrative authority is the Ministry of Education (PRC President, 1998, 2). Nevertheless, the power to regulate higher education is distributed between different departments within the State Council and these departments have designed and implemented various quality assurance mechanisms since the 1980s.

The quality of research is assured by the National Natural Science Foundation, the Ministry of Science and Technology, and the National Planning Office of Philosophy and Social Science through the grant bidding mechanisms. To apply for funding from these bodies, teachers have to submit applications and only a limited number of applications receive funding. The quality of research is expected to be assured through competition. This also reflects the State Council's pursuit of efficiency.

The quality of graduate education is assured by the Academic Degrees Committee. The quality assurance mechanisms used by the committee include degree-awarding power, the Postgraduate Degree Evaluation, the selection of National Key Disciplines and the Disciplinary Ranking. According to the *Degree Regulations of the People's Republic of China*, the Academic Degrees Committee decides on whether programmes receive degree-awarding power (State Council, 1981). A degree awarding power is granted based on whether the programme offers adequate courses, delivered by high-achieving researchers, has adequate resources and has a sound student assessment system (State Council Academic Degrees Committee, 1981). The

Postgraduate Degree Evaluation for Master's and doctoral programmes has been conducted since 2005. A programme's degree-awarding power will be removed if it fails the evaluation. The evaluation criteria include the amount of funding received over the previous five years. After the first evaluations between 2005 and 2006, seven doctoral and 33 Master's programmes were disqualified.

The selection of the National Key Disciplines was announced in the *CPC Central Committee Decision on the Reform of the Education System* in 1984 (CPC Central Committee, 1984). It is intended to allocate funding based on programmes' research performance in order to maximise the efficiency of funding. The awarding of the title of National Key Discipline is based on an evaluation of research performance. After viewing the applications submitted by departments and conducting site visits, evaluators select the best and award them the title of National Key Discipline. Beyond the honour the winners of National Key Disciplines are given financial rewards as well. The first, second and third selection of the National Key Disciplines were conducted in 1986, 2001 and 2006, in which 416, 964 and 1,180 subjects were selected as National Key Disciplines.

Besides the selection of National Key Disciplines, another mechanism to assure the quality of graduate education is the Disciplinary Ranking. The Disciplinary Ranking has been compiled since 2002. It ranks programmes by researchers, research outputs, graduate student outputs and reputations. To conclude, the quality monitoring mechanisms for graduate education mainly evaluate programmes based on research.

Due to the division of responsibilities, the Ministry of Education is only responsible for the quality of undergraduate education with a focus on teaching. The most influential quality assurance policies are the UTE and the Quality Project (for the history of other quality assurance policies launched by the Ministry of Education, see Chapter One). The two policies will be explained in full detail later in this chapter.

The term 'quality assurance' refers to evaluations in policy documents issued by the Ministry of Education. The idea of quality assurance was first mentioned in 2001 in *Some Opinions on the Improvement of the Quality of Teaching in Higher Education*, which recommended establishing a sound quality monitoring and assurance system to assure the quality of teaching (Ministry of Education, 2001). However, the exact term 'quality assurance system' did not appear, until the *2003-2007 Education Action Plan* was issued. It states:

> To improve the quality assurance system, an evaluation and consultant agency and a five-year cycle higher education teaching quality evaluation mechanism have to be established.
>
> (Ministry of Education, 2004a, 8)

The quality assurance system referred to in the above document is the UTE. This quote shows the limitation of the quality assurance policies issued by the Ministry of Education. The scope of quality assurance is confined to the quality of teaching. This limitation of the Ministry of Education's quality assurance policies is caused by the fragmented regulative power mentioned earlier.

In summary, the quality of higher education is divided into quality of research, quality of graduate education and quality of undergraduate teaching, and is assured by different departments within the State Council. Due to the division of power, the Ministry of Education is only responsible for assuring the quality of undergraduate

teaching. This division may look reasonable at first glance. However, the analysis of interview data in the following chapters will reveal a key problem with this division, which is that teachers are struggling to maintain a balance between the competing targets created by different quality assurance policies issued by different departments.

4.2 The Political Agenda Setting of the UTE

According to Knoepfel et al. (2007; 2011), public policies are intended to solve *public problems*. However, not all problems can become *public problems*. When individuals encounter some problems, Knoepfel et al. call them *problematic private situations*. When there is a social recognition of these problems, they become *social problems*. When *social problems* receive political attention, they become *public problems*.

The identification of *social* or *public problems* determines the *target groups* and establishes the *causal hypothesis*, i.e. who and what contribute to poor quality teaching. This section will draw a comparison between the *social problems* identified by the public and the *public problems* selected by the government so as to reveal the discrepancies between the two. The cause for concern over the quality of higher education will be introduced first, followed by the *political definition of public problems*.

4.2.1 Cause for Concerns Over the Quality of Education

In 1998, the Ministry of Education issued *An Action Plan for Invigorating Education for the 21st Century* (Ministry of Education, 1998), which announced the government's decision to increase participation in higher education. The target set by the *1998 Action Plan* was as follows:

> To improve the quality and the efficiency of education…by 2000 the number of both three-year and four-year undergraduates should reach 6,600,000… The higher education enrolment rate reaches 11%. The student–teacher ratio

> should increase to 12:1… In 2010… the higher education enrolment rate reaches 15%...
>
> (Ministry of Education, 1998, 6)

As shown in the quote, the Ministry of Education required universities to increase the student–teacher ratio, although a low student–teacher ratio was often regarded as an indicator of high quality. This reflects the State Council's pursuit of efficiency.

The target set for 2000 was very ambitious. In 1998, the number of both three-year and four-year undergraduates was 3,408,764 (Ministry of Education, 2005) and the higher education gross enrolment rate was 9.8% (Ministry of Education, 2008). This meant that within two years the number of undergraduates had to be almost doubled. The State Council believed that 15% level was critical, because it is the threshold of mass education as defined by Trow (1973). Once gross enrolment reached 15%, the State Council could claim that Chinese higher education had been transformed from an elite higher education system to a mass higher education system. Achieving mass higher education was regarded as a milestone on the way to becoming a developed country. Although the plan was to reach the threshold by 2010, in practice the threshold was achieved in 2002 (Zhou, 2004, 18). Reaching the threshold of mass higher education in 2002 coincided with the end of Jiang Zeming's[11] term in office. Bringing the gross enrolment rate from 9.8% to 15% in five years, just one term of office, could be seen as being politically motivated, because this government wanted to take the credit for transforming Chinese higher education from an elite education to a mass education system.

11 Jiang Zeming was the president of China from 1993 to 2003.

Beyond the political reasons, the decision to increase access to higher education was made due to economic considerations. Kang (2000) argued that mass higher education would produce a better-educated workforce to boost the knowledge economy. In addition, Kang maintained that another crucial reason was to stimulate the domestic economy that was affected by the 1997 Asian financial crisis by directing more household spending to education. Development Research Centre of the Ministry of Education estimates that increasing access to higher education would generate an extra 7 billion CNY of gross domestic product (GDP) in 1999 (Anonymous, 1999).

To conclude, the policy of increasing participation in higher education was politically and economically motivated at the cost of the quality of education. Unsurprisingly then, concerns over the quality of higher education started to mount.

4.2.2 Political Definition of the Public Problem

4.2.2.1 Social Problems

After launching the policy of increasing access to higher education, problems relating to the quality of education were increasingly reported in the media. Due to great media exposure, these problems drew public attention and became *social problems* especially after 2000. There was a concern about academic standards because admission requirements were lowered to allow more students to enter higher education (Di, 2000; Gao, 2000). Gao (2000) highlighted on the artificially high graduation rate, which would endanger academic standards. Parents and students also worried about the financial burden added by raising tuition fees, and doubted if universities could recruit students on the basis of affordability rather than merit (Gao, 2000).

Another concern was a lack of *resources*, particularly *infrastructure* and *human resources*. Interviewing university managers, Di (2000) revealed the problems of inadequate infrastructure and a lack of teachers. Even when there were enough teachers, they were overwhelmed by the overcrowded classrooms and could hardly maintain the quality of teaching (Gao, 2000). With universities struggling in terms of their basic resources, high quality education became a luxury.

The policy of increasing access to higher education also affected postgraduate education. The increase in the number of both undergraduate and graduates is illustrated in Table 5. He (2003) voiced public concern over the quality of postgraduate education. Sun (2003) reported the problems of a lack of supervision and funding caused by a significant increase in student numbers.

Year	Number of undergraduates	Increase in undergraduate participation rate	Number of graduate students	Increase in graduate students' participation rate
1998	3,408,764	7%	22,532	13%
1999	4,085,874	20%	34,628	17%
2000	5,560,900	36%	67,726	29%
2001	7,190,658	29%	92,017	31%
2002	8,620,989	20%	107,724	27%
2003	11,085,642	29%	150,280	30%

Table 5 The increase of both undergraduate and graduate students

Data extracted from the official website of the Ministry of Education

Overall, the public were concerned about both undergraduate and graduate education. The *social problems* involved lower academic standards, a shortage of resources and therefore heavier workloads for teachers. These problems were caused by the State Council due to political and economic considerations. Hence, the behaviours of the State Council should have changed and State Council should have been the *target group*. The *end beneficiaries* should have been universities, teachers and students.

4.2.2.2 Public Problems

From November 2000 to June 2001, the Ministry of Education launched an investigation (hereinafter referred to as the *2000-2001 Investigation*) into the impact of the policy of increasing access to higher education and identified seven problems (Song, 2002). These were:

Problem One: Outdated programmes

Problem Two: Lack of *Human Resources*: teachers

Problem Three: Lack of *Human Resources*: management personnel

Problem Four: Lack of *Consensus* on the importance of teaching

Problem Five: Lack of *Infrastructure*: teaching facilities

Problem: Six: Outdated teaching content

Problem Seven: Outdated management

The *social problems* – i.e. the lack of *human resources* and *infrastructure* – were acknowledged by the Ministry of Education. The Ministry of Education added the lack of another *resource* – *consensus* – as it believed there was a lack of *consensus* on the importance of teaching. Moreover, the Ministry of Education added several problems related to teaching, which are problems of programme design, syllabi, and teaching content. The two *social problems – graduate education and lower academic standards* – were ignored by the Ministry of Education. *Graduate education* was overlooked due to the fragmented power structure mentioned earlier in this chapter, with the Ministry of Education not being responsible for graduate education. The issue of *lower academic standards* were neglected because of the pursuit of efficiency.

Lower recruitment standards were caused by government policies. The number of students enrolled by the universities was not decided by the universities themselves but instead by the Ministry of Education. In order to increase participation in higher education, universities were required to recruit more students. Universities recruited students based on their scores in the national university entrance examination. With more students accessing higher education, the admission scores required were not as high as before. Hence, the recruitment standards became lower.

However, lower recruitment standards did not necessarily lead to lower academic standards. If universities let students fail when they could not meet universities' standards, academic standards would be maintained. However, when universities tried to maintain an artificially high retention and graduation rate, academic standards would drop. High retention and graduation rates were the norm in Chinese higher education. In the stage of elite education, because higher education was free and universities undertook the task of training an elite workforce for allocated positions, the retention and graduation rates were maintained at a very high level in order to ensure the efficiency of the investment (Yang and Li, 2011).

High retention and graduation rates also endangered academic standards. At the stage of mass higher education, student enrolments were not as selective as before. When universities strove to keep the retention and graduation rates as high as before, they were unable to maintain the same academic standards as before. The reason why universities did not reduce the retention and graduation rates was because the State Council encouraged them to keep the rates high in order to ensure efficiency. This was evidenced in the UTE evaluation scheme (see Table 6).

4.3 The Policy Programming of the UTE

The researcher defines *programming* as the stage of making policies to solve the *public problems* identified in *agenda setting*. According to this definition, the *political-administrative programmes (PAPs)* and *action plans* are produced at this stage. 'The *PAP* defines the legal bases for the objectives, intervention instruments and operational arrangements of the public action' (Knoepfel et al., 2007, 53). *Action Plans* are 'the set of planning decisions considered as necessary for the coordinated and targeted production of administrative services (outputs)' (Knoepfel et al., 2007, 53). In the case of the UTE, the *PAP* is *Some Opinions on the Improvement of the Quality of Teaching in Higher Education* (Ministry of Education, 2001) and the *action plan* is the *2004 Undergraduate Teaching Evaluation Scheme*[12] (Ministry of Education, 2004c). The former is referred to as the *2001 Opinions* and the latter is referred to as the *2004 UTE Scheme* in what follows.

The *2001 Opinions* indicated the State Council's intention to use evaluation – *an authority tool* – as the intervention instrument. The *2001 Opinions* made 12 recommendations (see the second column of Table 6) to tackle the seven *public problems* identified by the *2000-2001 Investigation* (see the first column of Table 6). *The 2004 UTE Scheme* made a detailed *action plan* (see the third column of Table 6) to implement these recommendations. A comparison between the *political definition of the public problem* (*the 2000-2001 Investigation*), the *PAP* (*the 2001 Opinions*) and the *action plan* (*the 2004 UTE Scheme*) is drawn in Table 6. For the full *2004 UTE Scheme*, see Appendix 7.

12 In 2002 a UTE scheme was released. It was derived from the evaluation schemes of previous evaluations. When the UTE started in 2003, this evaluation scheme was used. After one year, some minor revisions were made and a revised UTE scheme was formally published in 2004. Some revisions of certain criteria were made for medical universities in 2004 and for key universities, sports and arts universities in 2006.

Public problems	***PAP***	***Action Plan***
Problem One: Outdated programmes	Recommendation Eleven: Universities should provide experiments or placements to develop students' innovation or practical abilities.	Programmes ®
Problem Two: Lack of *Human Resources*: teachers	Recommendation Six: The central government, local governments and universities should provide training for teachers. Universities should recruit and train early-and mid-career teachers to meet the demand for high quality teaching.	*Resources - human resources*: student–teacher ratio
		Resources - human resources: teachers' qualifications (T) ®
Problem Three: Lack of *Human Resources*: management personnel		*Resources - human resources*: the number of management personnel
		Resources - human resources: qualifications of management personnel
Problem Four: Lack of *Consensus* on the importance of teaching	Recommendation Three: Universities should encourage professors to teach core undergraduate courses.	*Resources-consensus*: Percentages of professors who taught undergraduate courses and professional ethics (T)
	Recommendation Five: Universities should promote professional ethics	
	Recommendation One: Government education departments and university leadership should understand the importance of teaching.	*Resources-consensus: university leadership*
	Recommendation Four: Universities should make teaching one of the important criteria for employment and appointment.	
Problem Five: Lack of *Infrastructure*: teaching facilities	*Recommendation Two: Local government should increase the investment in teaching.*	*Resources-infrastructure: buildings, laboratories, libraries, intranet and sports facilities; Resources- financial resources: 'four kinds of funds'*
Problem Six: Outdated teaching content	Recommendation Nine: Universities should encourage the use of new, imported or award-winning textbooks or compiling high quality textbooks.	Teaching Content: syllabi and textbooks (T)
	Recommendation Eight: Universities should promote English teaching.	*Teaching Method: English (T)*
	Recommendation Ten: Teachers should use modern technology to improve teaching.	*Teaching Method: technology (T)*

		Teaching Quality (T) ®
Problem Seven: Outdated management	Recommendation Twelve: Universities should establish sound quality monitoring and assurance systems.	Management personnel's publications ® and an internal quality assurance mechanism (T)
	Recommendation Seven: Universities should promote a learning environment and motivate students to learn.	*Learning Environment*
		Strategic Planning
		Characteristics
		Learning Outcomes
		Employment Rate
		Reputation ®

Table 6 A comparison between the *Agenda Setting*, the *PAP* and *action plans* of the UTE

('T' indicates that the target group is teachers; ® symbolises that the indicator is fulfilled by the demonstration of research performance; Italics highlight the inconsistencies)

As illustrated in Table 6, the *PAP* and the *action plan* of the UTE addressed all of the *public problems*, but there are inconsistencies between the *public problems*, the *PAP* and the *action plans*. Some factors were removed and others were added.

There were inconsistencies in the *target groups* identified by the *PAP* and the *action plan*. The *target groups* identified by the *PAP* included the central government, local governments, universities and teachers. Given that *the 2004 UTE Scheme* was designed to evaluate universities, the central and local governments were removed from the *target groups* in the *action plan* (see changes regarding Recommendations One and Two). Moreover, the *action plan* added a new *target group* – students, whose learning outcomes and employment were used as indicators. Hence, the *target groups* of the UTE became universities, teachers and students.

Although the UTE was an evaluation of teaching, *the 2004 UTE Scheme* often used research-related indicators to evaluate teaching (see the indictors marked with a ® in Table 6). This could be attributed to the fact that it was difficult to find quantifiable indicators to measure teaching, unlike research which can be measured by publications and grants. This could also be the reason why Recommendation Four in the *PAP* was ignored in the *action plan*. Furthermore, using research-related indicators to evaluate teaching reflected the logic followed by the Ministry of Education, which is that good research leads to good teaching and good teaching is evidenced by good research. Tan (2005), Wang and Zheng (2008) and Zhou and Zhu (2004) expressed a concern that the abundance of research-related indicators would drive universities to put more weight on research rather than teaching.

The *PAP* and *action plan* added more *public problems*. In addition to the problem of teaching content, the *PAP* and *action plan* added the problems of teaching methods. Teachers were expected to use English and new technology, such as multimedia, to teach (see changes regarding Recommendations Eight and Ten). Additional *public problems* involved targeting universities: the learning environment, strategic planning, and universities' characteristics and reputations.

4.4 The Political Agenda Setting of the Quality Project

4.4.1 Social Problems

Three years after the launch of the UTE, some problems caused by the UTE started to receive public attention through the media and therefore became social problems. Impression management was heavily criticised. Feng (2008), Liang, Zhang and Guo (2006), Wang and Zhang (2007) and Ye (2006) complained that universities asked teachers and students to change their normal behaviours to create a better image for

the inspection visits, for instance, asking students to clean the campus instead of attending lectures and, during the UTE inspection itself, replacing regular teachers with teachers who had a good teaching reputation.

The most extreme case of impression management involved the forging of documents, which received most criticism (Feng, 2008; Liang, Zhang and Guo, 2006; Wang and Zhang, 2007; Ye, 2006). However, the researcher would argue that this problem was caused by a UTE indicator concerning grading. Before the UTE, teachers graded students' examination papers and dissertations in their own individual ways. The UTE designed a particular grading format and universities were evaluated according to this format. In order to meet the UTE requirements, universities had to ask teachers to re-grade examination papers and dissertations for the previous three years according to this new format. Evaluating universities according to the newly introduced format was effectively encouraging universities to forge documents.

Another problem caused by the UTE was the homogenisation of universities. Feng (2008) questioned the appropriateness of using one set of criteria to evaluate different kinds of universities. Tan (2005), Wang and Zheng (2008) and Zhou and Zhu (2004) claimed that the UTE's indicators were based on the performance of research-intensive universities, which would thus motivate other universities to become research-intensive universities. This, it was argued, would harm the diversity of Chinese higher education.

In addition, Ji (2008) pointed out another problem – an excess of evaluations. Besides the UTE, there were the selection of National Key Disciplines, the evaluation of

degree-awarding powers, the evaluation of Project 211, the evaluation of Project 985 and so on. Ji (2008) called for an overarching design of evaluations to reduce unnecessary and overlapping evaluations.

Some of the *social problems* that emerged in the media were problems that the UTE was intended to tackle, such as the problems surrounding the learning environment, management and programmes (Lv and Zhang, 2004; Wang, 2006; Zhou and Zhu, 2004). This suggests that the UTE drew public attention to these issues but that it was not able to solve them.

Other *social problems* were those that had been discussed in the media before the UTE, including declining academic standards and the shortage of *infrastructure* and *human resources* (Lv and Zhang, 2004; Qu, 2008; Tan, 2005; Wang and Zheng, 2008; Zhang, 2006; Zhou and Zhu, 2004). The concern over the quality of higher education caused by the increase in participation in higher education persisted after the UTE.

4.4.2 Public Problems

Although the Ministry of Education did not admit that the UTE failed to assure the quality of undergraduate teaching, the following actions taken by the Ministry implied that it was aware of the failure of the UTE. The Ministry of Education did not start the second round of the UTE within five years, as had been its original intention (Ministry of Education, 2004a). Instead, in 2007 the Ministry of Education issued the *Opinions on the Quality of Undergraduate Teaching in Higher Education* (hereinafter referred to as the *2007 Opinions*) to announce the launch of a new project, the Undergraduate Teaching Quality and Teaching Reform Project, also called the Quality Project (Ministry of Education, 2007).

The *public problems* identified by the *2007 Opinions* were the same as those identified by the UTE, which were inappropriate programmes, dissatisfied student learning outcomes, a lack of qualified *human resources*, inappropriate teaching content and poor teaching methods. This is illustrated by the following quote:

> The quality of higher education has not met the demand for the development of the society and the economy. Programmes offered by many universities are not appropriate. Students' capability to practise and innovate has to be improved. The quality of teachers has to be improved. The way that students are educated, teaching content and teaching methods should be changed.
> (Ministry of Education, 2007, 1-2)

Besides these problems, in his speech the Minister of Education mentioned another three: a lack of *financial resources* for teaching, a problem in management, and the homogenisation of universities (Zhou, 2007). Except for the homogenisation of universities, the other two problems were identified by the UTE too – this identification of the same problems shows that the UTE was not able to resolve them. To sum up, except for the homogenisation of universities, the remaining the *public problems* identified by the Quality Project were the same as those selected by the UTE and most of the *social problems* raised by the public – with the exception of programmes –were ignored by the Ministry of Education.

4.5 The Policy Programming of the Quality Project

The *2007 Opinions* also acted as the *PAP* of the Quality Project, which intended to use awards, *the incentive tools*, as the *intervention instruments* to tackle the identified *public problems*, except for the problem of management (see Table 7).

Public problems	***PAP***
Programmes	The Featured Programmes Awards
Homogenisation of universities	The Featured Programmes Awards
Lack of financial *resources* for teaching	The Featured Programmes Awards, the Excellent Courses Awards and the Bilingual Courses Awards
Quality of *human resources*	The Distinguished Teacher Awards and the Teaching Groups Awards (T)
Management	
Teaching content	The Excellent Courses Awards and the Ten Thousand New Textbooks Project (T)
Teaching methods	The Excellent Courses Awards and the Bilingual Courses Awards (T)
Student learning outcomes	The Exemplary Experiment and Teaching Centres and the Undergraduates' Innovation Project

Table 7 A comparison of the *Agenda Setting* and the *PAP* of the Quality Project

(The public problems and the PAPs are extracted from the Opinions on the Quality on the Quality of Undergraduate Teaching in Higher Education (Ministry of Education, 2007); the PAP that targets for individual teachers is marked with 'T')

Given that this research is only interested in the impact of quality assurance policies on teachers, only the following awards targeting individual teachers will be discussed in this thesis: the Distinguished Teacher Awards, the Teaching Groups Awards, the Excellent Courses Awards, the Bilingual Courses Awards, and the Ten Thousand New Textbooks Project.

The *action plans* of these awards were the policy documents related to the selection of award-winners (for the full list of these documents, see Appendix 8). The award selection criteria written in these awards' application forms are summarised in Table 8.

	Excellent Courses Awards	Teaching Group Awards	Distinguished Teacher Awards	Bilingual Courses Awards	Ten Thousand New Textbooks Project
Degree	√	√	√	√	√
Title	√	√	√	√	√
Position	√	√	√	√	
Research	√	√	√	√	√
Teaching Experiences	√	√	√	√	√
English				√	
Diversity of team members	√	√			
Professional Development	√	√	√		
Teaching Content	√	√	√	√	
Pedagogy	√	√	√	√	
Textbooks		√	√	√	
Online Resources	√	√		√	
Facilities	√				
Reputation	√				√
Student Feedback	√			√	

Table 8 The selection criteria in the application forms of the Excellent Courses Awards, the Distinguished Teacher Awards, the Teaching Groups Awards, the Bilingual Courses Awards and Ten Thousand New Textbooks Project

As indicated in Table 8, besides the *public problems* mentioned above, the *action plans* added one more *public problem* – a lack of teamwork. Most criteria in Table 8 were teaching-related, so these awards should have been able to promote teaching. However, whether these awards achieved this or not will be assessed using teachers' accounts presented in the following chapters.

4.6 Summary

The analyses of the policy cycles of the UTE and the Quality Project above reveal a vital problem in policy-making, which is the government's ignorance of the *social*

problems that had emerged among the public. This problem is caused by the *national institutional rule* – the centralised control of the State Council. Without the need to consult other *actors*, the State Council is able to select problems that suit its interests – maximising the efficiency to be the *public problems*. Moreover, repeatedly identifying the same *public problems* in the *agenda settings* of the UTE and the Quality Project policy, as illustrated in Table 9, reflects the failure of the quality assurance.

Based on the analysis above, the *public problems* and the expected actions of teachers of the UTE and the Quality Project are summarised in Table 9.

	Public problems	Expected actions of teachers
UTE	Lack of qualified *human resources*	To obtain higher degrees and titles and the qualifications of the Principal Lecturers[13].
	Lack of *consensus* on the importance of teaching	To teach undergraduate courses; to demonstrate professional ethics and commitment
	Inappropriate teaching content	To update syllabi; to use or compile high quality textbooks
	Poor teaching methods	To utilise modern technology; to use English to teach
	Low quality teaching	To improve research performance
	Ineffective internal quality assurance mechanism	To comply to internal quality assurance policies
Quali -ty Proje ct	Lack of qualified human *resources*	To obtain higher degrees, titles, positions and more teaching experience; to improve research performance; to offer/participate in professional development
	Inappropriate teaching content	To update the teaching content; to use or compile high quality textbooks
	Poor teaching methods	To use English to teach; to employ appropriate pedagogies; to share resources online
	Lack of teamwork	To form teaching groups

Table 9 A summary of the UTE and the Quality Project

What teachers are expected to do to resolve the *public problems* will be compared and contrasted in later chapters with teachers' accounts in order to assess whether those expected actions have been performed by teachers or not. The solution – i.e. improve research performance – reflects an assumption held by the government, which is that good teaching should be evidenced in research. Whether this assumption is shared by teachers or not will be discussed later.

13 The qualification of Principal Lecturers was initiated by the UTE. According to the UTE requirement, all courses had to be taught by teachers with this qualification. This qualification was issued by universities based on teachers' degrees, titles and teaching experience.

5 The Project 985 University

This chapter starts by introducing the Project 985 University's *institutional rules.* The *institutional rules* affect how the Project 985 University and teachers react to the national policies. After presenting the existing internal quality assurance mechanisms, how the Project 985 University responded to the UTE and the Quality Project and how department heads and teachers perceive the impact of both external and internal quality assurance mechanisms will be elaborated in order to evaluate their contribution to the quality of teaching. Teachers' reflections on problems of the existing quality assurance mechanisms and factors contributing to the improvement of their teaching will be provided at the end.

5.1 The University Context

The term *university institutional rules* is adopted from Knoepfel et al.'s (2007; 2011) framework, instead of a term used by universities in their university policy documents. Similar to *national institutional rules*, *university institutional rules* are legal and social norms, formal and informal procedures and cultural values. The difference between *national and university institutional rules* lies in the *political-administrative authorities* who created the rules and the governing area. *National institutional rules* are made by the ruling party and its government agencies, and can affect all universities in China. As a result, *universities institutional rules* are partly created by *national institutional rules*. In addition, *universities institutional rules* are also set up by university leadership. Therefore, the researcher distinguishes two kinds of *university institutional rules*: the *institutional rules* inherited from the *national institutional rules* and those established by the university leadership. These

institutional rules limit teachers' access to resources, guide teachers' behaviours and define political and social legitimacy. *University institutional rules* can be understood as university context.

5.1.1 University Institutional Rules Inherited from the National Institutional Rules

The Project 985 University was established in the early 20th century and is a research-intensive university. Like in any other university in China, the control of the State Council or the CPC is evident at the Project 985 University. The State Council appoints the President of the University and the CPC Secretary of the University. The CPC Secretary is in charge of the University's CPC committee, which has a CPC branch office in every department. The CPC Secretary has a higher status than the President of the University. Any decision taken by the President has to be endorsed by the CPC Secretary.

After the State Council launched the policy to increase participation in higher education, the number of full-time students doubled from 2002 to 2010 (see Table 10). To cope with this sudden increase of student numbers, the Project 985 University resorted to establishing new campuses in 'university towns' to build new classrooms, student accommodation, canteens and sports facilities. Unlike the common denotation of term 'university town' in the UK, which refers to towns as Oxford and Cambridge, the term 'university town' in China denotes a dedicated area in suburbs where universities build new campuses. Although both central and local government provided funding for infrastructure, it was never enough to cover Project 985's expenditure on infrastructure. According to Project 985 University Yearbooks from 2002 to 2010, the average gap between the government funding and university

expenditure on infrastructure from 2002 to 2010 was ¥167,516,667 CNY (£16,751,666.7 British Pound Sterling (GBP)). The gap reached ¥485,940,000 CNY in 2008 and ¥460,150,000 CNY in 2010. Given the large gaps between government funding and university expenditure on infrastructure, the importance of generating additional incomes was stressed.

While the Project 985 University was trying to improve the infrastructure to cope with the dramatic increase of students, the increase of teachers lagged behind, as illustrated in Table 10. In fact, the Project 985 University is not responsible for this lag, because it does not have the power to determine either student or staff recruitment. Recruitment is determined by quotas allocated by the State Council.

Year	Number of full-time students	Increase of full-time students	Number of teachers	Increase of teachers	Student–teacher ratio
2002	25363		2403		11:1
2003	28600	12.8%	2413	0.4%	12:1
2004	34639	21.1%	2584	7.1%	13:1
2005	38797	12.0%	2602	0.7%	15:1
2006	43127	11.2%	2600	-0.1%	17:1
2007	46592	8.0%	2645	1.7%	18:1
2008	47834	2.7%	2652	0.3%	18:1
2009	48805	2.0%	2780	4.8%	18:1
2010	49475	1.4%	2875	3.4%	17:1

Table 10 The numbers of full-time students and teachers and the student–teacher ratio at the Project 985 University

Extracted from (Uni985 Yearbook, 2002; 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010)

Moreover, the majority of the Project 985 University's income comes from both central and local governments. According to the Project 985 University Yearbooks (2002; 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010), around 70% of income came from the governments. Of the 70% government funding, funding provided by Project 211 and especially Project 985 was very substantial. According to the President of the University (Uni985 Yearbook, 2009) and unofficial sources

(Anonymous, 2012) the Project 985 University received around ¥3,540,000,000 CNY (£354,000,000 GBP) from Project 985 alone from 1998 to 2013.

Around 30% of income was raised by the university. However, even the self-raised funding was highly controlled by the State Council. Around 70% of self-raised funding came from tuition fees. As mentioned above, student recruitment is restricted by the quotas allocated by the State Council. Moreover, the State Council regulates the tuition fees that universities are allowed to charge. The number of students determines the amount raised by the university from tuition fees.

The sources of the Project 985 University's income also highlight the importance of research. The Project 985 University Yearbooks from 2002 to 2010 show that income directly from government-funded and private-funded research projects constituted around 14% to 22% of total university income. Another source of funding labelled 'education funding', which constituted 25% to 50% of the total income, was also highly influenced by research. Education funding included subsidies and funding for various government initiatives, such as Project 985, Project 211 and the selection of National Key Disciplines. The inclusion of universities in Project 985, Project 211 and National Key Disciplines is based on universities' research performance. Hence, education funding is effectively determined by research performance.

The appointment of university leaders, quotas for student recruitment and funding are decided on by the State Council. The decision on the appointment of university leaders relates to *human resources* as defined by Knoepfel et al. (2007; 2011). Using quotas to determine how many students and staff universities are able to recruit relates

to *legal resources* as defined by Knoepfel et al. (2007; 2011). These serve as 'legal and regulatory bases' (Knoepfel et al., 2011, 69) to decide on student recruitment, appointment of university leaders and allocation of funding. Consequently, the *legal resources* affect funding, in other words the *financial resources* (Knoepfel et al., 2011). Given the fact that the State Council controls these essential resources, the State Council has a strong control over universities. This *national institutional rule* exerts its influence over the *university institutional rule*, so *University Institutional Rule One* is: to conform to government policies and requirements.

Given the amount of income generated directly or indirectly from research and the large gap in infrastructure, *University Institutional Rule Two* is: to work hard to improve research performance in order to obtain resources. Facing the pressure of increasing participation in higher education, *University Institutional Rule Three* is: to try to solve the problems caused by the dramatic increase of students. These are the *university institutional rules* formed by national policies.

5.1.2 University Institutional Rules Established by the University Leadership

The President of the Project 985 University (from 1999 to 2010) promoted a change from the prevalent bureaucratic culture in Chinese universities to a more collegiate culture. Two mottoes frequently mentioned by the President illustrated his beliefs: 'university is an academic community' and 'the university is professors' university' (Uni985 Yearbook, 2004; Uni985 Yearbook, 2005; Uni985 Yearbook, 2006; Uni985 Yearbook, 2007; Uni985 Yearbook, 2008; Uni985 Yearbook, 2009; Uni985 Yearbook, 2010). In many Chinese universities, senior administrative staff and managers were regarded as officials with higher status. To change this perception, the President often emphasised that the administrative staff and managers were meant to

assist teachers (Uni985 Yearbook, 2004; Uni985 Yearbook, 2007; Uni985 Yearbook, 2008; Uni985 Yearbook, 2010).

In order to restrict the university's administrative power, more power was given to departments. The President believed that departments should be given autonomy over *human* and *financial resources*. In 2003, the university launched reforms to realise this (Uni985 Yearbook, 2003). Departments were given a right to appoint and dismiss all personnel, within the limit of the allocated quotas. In terms of promotion, before 2003, the promotion decisions were made solely by the university. From 2003 onwards, promotion decisions were considered by department committees first and then passed on to the university committee. Before 2003, decisions on department expenditures were made by the university. After 2003, the university allocated certain amounts of money to departments and they could then decide where to spend that money. To ensure fairness and transparency, the university requested departments to establish various committees, for example a teacher appraisal committee, teaching committee and student recruitment committee, to devolve the responsibilities of decision-making from department leaders. The university would not interfere with the decisions made by departmental committees.

Although a considerable improvement had been made to promote a collegiate culture, the researcher would argue that the Project 985 University was still in the shadow of a bureaucratic culture. Firstly, after various committees were established, many teachers became committee members and participated in the decision-making process, but these committees were still headed by the President, the CPC Secretary, the department heads and departmental CPC secretaries. Secondly, the reform was carried

out in a top-down approach. Departments did not play any active role in it. They just received the power given by the university. In one word, *University Institutional Rule Four*, created by the President of the University, is to promote academic freedom and autonomy rather than the prevalent bureaucratic culture.

5.2 Internal Quality Assurance Mechanisms

Before the UTE and the Quality Project, there were a number of long-established internal quality assurance mechanisms implemented by the Project 985 University, which are: class observation by leaders or inspectors, the student evaluation of teachers and teacher appraisal.

5.2.1 Class Observation and the Student Evaluation of Teachers

Since 1993, the university has appointed retired teachers (around 13) who were good at teaching to be the university inspectors. Class observation by leaders was conducted before 2002. The student evaluation of teachers has been conducted since 1990s too but it was initially paper-based. From 2005, Project 985 started to conduct the online student evaluation of teachers. The student evaluation of teachers was compulsory for all students.

Evaluation forms for class observations by leaders and by teaching inspectors and the student evaluation of teachers have two sections, rating and open comments, but they are different in detail. For leaders and students, the rating section is more structured. The rating consists of four categories: teaching attitude, teaching content, teaching methods and teaching outcomes. In each category, there are several indicators. For each indicator, leaders and students had to select among ‘excellent’, ‘good’, ‘acceptable’, ‘fair’ or ‘poor’ to rate teaching. In the evaluation form for students, these results were assigned scores. One ‘excellent’ equalled 9.5 points; one ‘good’ equalled

7.5 points; one ‘acceptable’ equalled 6 points; and one ‘poor’ equalled 3 points. These scores could be added together to have a total score to compare teachers. Indicators in the teaching evaluation forms for students and leaders are compared in Table 11. The indicators for leaders and students have some subtle differences, which are highlighted in italics. The rating section in the evaluation form for teaching inspectors is less structured, containing only the overall rating.

	For students		**For leaders**	
Teaching attitude	Indicator 1	Elegant teaching manner	Indicator 1	Adequate preparation
		Role model		*Teach meticulously*
	Indicator 2	Adequate preparation	Indicator 2	Elegant teaching manner
		Teach proficiently		*Good manners*
Teaching content	Indicator 3	*Clarity in content*	Indicator 3	*Clarity in logic*
		Following syllabus		Following syllabus
	Indicator 4	Highlighting key points	Indicator 4	Highlighting key points
		Explaining thoroughly the parts that are difficult to understand		Explaining thoroughly the parts that are difficult to understand
	Indicator 5	Reflecting the latest development in the field	Indicator 5	Reflecting the latest development in the field
Teaching methods	Indicator 6	Relating to practice	Indicator 6	Relating to practice
		Using appropriate examples		Using appropriate examples
	Indicator 7	Thought provoking	Indicator 7	Thought provoking
		Interactive		Interactive
	Indicator 8	Teaching methods can achieve good results	Indicator 8	Teaching methods can achieve good results
Teaching outcomes	Indicator 9	Students can understand and master the knowledge and skills	Indicator 9	Students can understand and master the knowledge and skills
	Indicator 10	To improve students' self-learning ability and to promote their interests in learning	Indicator 10	To improve students' self-learning ability and to promote their interests in learning

Table 11 Teaching evaluation forms at the Project 985 University

Having examined the design of the evaluation forms, the researcher argues that these evaluation forms suffer from several problems. The first problem is that there can be two factors within one indicator. This in fact applies to the majority of the indicators. For example, the first indicator in the student evaluation form consists of ‘elegant

teaching manner' and 'role model'. They are two separate factors. Being able to show an elegant teaching manner does not necessarily mean that the teacher can serve as a role model. If a teacher is good at one factor while bad at the other one, leaders and students will have difficulty in pointing out the problem, so teachers will not be able to correct it.

The second problem of the design of the rating is related to the face validity of indicators. For instance, 'explaining thoroughly the parts that are difficult to understand', an indicator under teaching content, should be under the category of teaching methods. Problems in face validity will cause misunderstanding in interpreting results. For example, if a teacher received a lower score in teaching content, he or she would think that s/he had problems related to teaching content. However, if the low score was caused by this indicator – explaining thoroughly the parts that are difficult to understand – the actual problem may lie in his or her teaching methods. The lack of face validity can mislead teachers into identifying the wrong problems and subsequently make teachers fail to correct their problems.

The third problem is also related to the validity – whether the evaluation forms are able to reflect students' definition of high quality teaching. As illustrated in Table 11, the indicators for students and leaders are very similar. However, Cheng's (2011), Fisher, Alder and Avasalu's (1998) and Oliveira, Oliveira and Costa's (2012) studies show that, although students and staff agree on the indicators of quality teaching, they disagree on the weightings and combinations of these indicators. Hence, the evaluation forms may be inadequate to represent students' or staff's understanding of quality teaching.

The fourth problem is caused by the use of total scores in the student evaluation of teachers, which harms its validity. As elaborated in Chapter Two, Marsh (1984) opposes the use of total scores because total scores are unable to reflect the multidimensionality of quality and the information provided by total scores is less valid and difficult to interpret. Marsh (1984) stresses that, if total scores have to be used, weightings of indicators have to be tested empirically. However, in the evaluation form for students, the total scores are used without being weighted.

The fifth problem is associated with the reliability of the evaluation form for leaders. Many leaders admitted that they just randomly picked lessons to observe from time to time. It is unlikely for the same leader to observe the same teacher's class more than once. By just observing once or twice, it would be difficult for leaders to collect enough information to make sound judgement on whether 'teaching methods can achieve good results', 'students can understand and master the knowledge and skills' or teaching has improved students' self-learning ability and promoted their interests in learning.

The problems mentioned above can jeopardise the validity and the reliability of these internal quality assurance mechanisms, which will subsequently affect their impact on assuring the quality of teaching.

5.2.2 Teacher Appraisal

In 2003, Project 985 published a new teacher appraisal policy: *The Teaching Staff Appraisal and Allowance Allocation Regulations* (Uni985 Yearbook, 2003). This policy specifies the requirements for both teaching and research. In terms of teaching,

teachers must have a minimum teaching workload each year. The minimum teaching workload for professors is 270 teaching units, for associate professors 250 teaching units and for lecturers 220 teaching units. One teaching unit equals one 45-minute lecture. Teaching-related work, such as supervising students, can be classified as teaching units too.

The President also suggested that department committees consider the quality of teaching when making promotion decisions (Uni985 Yearbook, 2008). Quality of teaching is judged by class observation and the student evaluation of teachers.

In terms of research, *The Project 985 University Teaching Staff Appraisal and Allowance Allocation Regulations* requires teachers to have certain research workloads each year as well. Teachers have to either conduct a research project as a principal investigator or have a 'disciplinary research' publication. In China, research is broadly divided into 'disciplinary research' and 'pedagogical research'. Pedagogical research is aimed at improving teaching and learning in any subject area in order to advance the understanding of teaching and learning. Disciplinary research is aimed at advancing knowledge in a subject area. The requirements of teaching and research workloads apply to every single teacher at the Project 985 University.

In addition to the required teaching and research workloads, teachers who want to apply for promotion have to fulfil more requirements. They are appraised by their research achievements over the previous five years. Research achievements comprise publications, research projects and awards. Publications are measured by the number of disciplinary research articles published in influential journals in the field and the

number of books, including textbooks. Research projects are measured by the number of grants awarded by both central and local government. The value of these research grants and awards follows a hierarchy. Grants or awards awarded by the central government are regarded as the most prestigious, followed by those rewarded by the provincial government and the city council.

5.3 University's Reactions to the UTE and the Quality Project Awards

The Project 985 University's reactions to the UTE and the Quality Project will be divided into *agenda setting* and *programming*. The analysis will show that the objective of carrying out the UTE has changed from assuring the quality of undergraduate teaching to achieving a good result in the UTE during *agenda setting*, which subsequently guides the following *programming* and implementation of the UTE. This explains why teachers at the Project 985 University said that it had 'no impact'. Unlike the UTE, the Quality Project was absent from *agenda setting* and *programming*.

5.3.1 Agenda Setting

There were many problems faced by the Project 985 University. However, not all of them were able to receive attention and become top priorities. Quality of undergraduate teaching became a top priority only when the Project 985 University was facing UTE inspection. Even the Quality Project awards failed to make quality of undergraduate teaching a top priority.

The Project 985 University's Yearbooks list each year's top priorities. Table 12 summarises the top priorities from 2002 to 2010.

Year	Top priorities					
2002	Project 211[14]	Staff[15]	Ethics[16]	CPC[17]		
2003	CPC	University Town[18]	Increase Student Enrolment	Management[19]	University Anniversary Celebrations[20]	
2004	World-Class University[21]	Management	Project 985[22]	University Anniversary Celebrations	University Town	
2005	Ethics	CPC	Management	Quality of Undergraduate Teaching	Project 985	
2006	Project 985	Project 211	CPC	Management	University Town	Ethics
2007	CPC	Project 985	Project 211	Management	Quality of Undergraduate Teaching	Manage-ment
2008	CPC	Project 985	Project 211	Technology Transfer[23]	UTE[24]	University Town
2000	CPC	Technology Transfer	Project 985	Ethics	Nation Anniversary Celebrations[25]	Graduate Employ-ment[26]

14 Priorities related to Project 211 are about demonstrating accountability in order to secure the following rounds of funding from Project 211.

15 Priorities related to staff are about staff recruitment and appraisal.

16 Priorities related to ethics are about promoting certain values among senior management, teachers and students, for instance, promoting the idea that management is providing service to teachers and students among senior management; promoting self-discipline and professionalism among teachers; promoting trustworthy and citizenship among students.

17 Priorities related to the CPC are about holding meetings to discuss the CPC-related businesses and disseminating policies made by the CPC.

18 Priorities related to the university town are about the construction and the management of campuses in university towns and tackling problems caused by campuses at different locations.

19 Priorities related to management are university policy making and implementation, such as the teacher appraisal policy.

20 Priorities related to university anniversary celebrations are preparations made for the celebrations.

21 Priorities related to world-class university are about an agenda to establish the Project 985 University as a world-class university through improving its research performance.

22 Priorities related to Project 985 are about demonstrating accountability in order to secure the following rounds of funding from Project 985.

23 Priorities related to technology transfer are about encouraging collaboration with industries and establishing spin-off companies.

24 Priorities related to the UTE are about making preparations for the UTE reports and inspection.

9						
2010	CPC	Staff	Project 985	Research[27]	Teaching[28]	University Town
	Graduate Employ-ment	International Communi-cation[29]	Management			

Table 12 The Project 985 University's top priorities

Extracted from (Uni985 Yearbook, 2002; 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010)

Several priorities consistently appeared in university policies over the years. These were priorities related to the CPC, Project 211, Project 985, university towns and management. They became top priorities because of the *institutional rules* mentioned at the beginning of the chapter. The CPC-related agendas were not directly connected to teaching or research, but they often became top priorities because of *University Institutional Rule One* – the strong control of the government – which resulted in the obedience of the university. Project 211 and Project 985 became top priorities because of *University Institutional Rule Two* – generating income. University towns became a top priority because of *University Institutional Rule Three* – dealing with problems caused by dramatic increase in the numbers of students. Management became a top priority because of *University Institutional Rule Four* – the President's belief in academic autonomy and community.

Though mentioned with less frequency, quality of undergraduate teaching was among the top priorities in 2005 and 2007 (see Table 12). It was highly likely driven by the UTE, firstly because of the timing. According to the original plan, the Project 985 University would undergo the UTE inspection in 2006. Hence, in 2005 the Project

25 Priorities related to nation anniversary celebrations are preparations made for the celebrations.

26 Priorities related to graduate employment are measures to help graduates to find jobs.

27 Priorities related to research are agendas to improve research performance.

28 Priorities related to teaching are agendas to improve the quality of teaching.

29 Priorities related to international communication are about setting up collaborations with institutions abroad.

985 University made quality of education the top priority in order to prepare for the UTE inspection in 2006. However, the inspection date was changed to 2008. After the plan was changed, in 2007 the Project 985 University called for preparation for the UTE one year before the inspection. After the UTE inspection, the quality of undergraduate education was no longer mentioned in the university's top priorities, despite the launch of the Quality Project. Secondly, instead of quality of education, the focus was on quality of undergraduate teaching, the same as for the UTE, which also indicated the impact of the UTE.

The UTE became one of the top priorities, but the Quality Project did not. This was because of their different natures. The UTE was an *authority tool*, whereas the Quality Project was an *incentive tool*. The UTE was compulsory and had high stakes. The CPC Secretary of the University regarded the UTE as a test of university leaders' leadership skills (Uni985 Yearbook, 2007). Moreover, the CPC Secretary of the University thought that achieving a good result in the UTE would allow the university to negotiate more resources, such as funding or various quotas for student recruitment, teacher recruitment, professorships and so on. In contrast, applications for the Quality Project awards were not compulsory and it did not have wider implications either.

Nevertheless, the quality of undergraduate teaching was less important than the agendas related to the CPC, Project 211, Project 985 and university towns, because it had fewer *human resources, financial resources, organisation* and *the urgency of the problem* than other priorities had. Compared to the CPC, the quality of undergraduate teaching did not have dedicated *human resources* and *organisation*. The implementation of the UTE relied on existing *human resources* and *organisation.*

Compared to Project 211 and Project 985, the quality of undergraduate teaching was not able to provide the university with substantial *financial resources*. Compared to university towns, the quality of undergraduate teaching was mentioned less frequently as being urgent. Quality of undergraduate teaching became urgent only when the university was going to be inspected. In addition, the quality of undergraduate teaching did not have strong support from the university *institutional rules* listed above.

5.3.2 Programming and Implementation

As mentioned in the last section, only the UTE became a university top priority but not the Quality Project awards, so the Quality Project awards were absent from *programming*, the second stage of policy cycle (Knoepfel et al., 2011). In fact, there was no official *PAP* for the UTE either. Only the *objectives*, one of the elements of the *PAPs*, were mentioned by two speeches delivered by the President and the CPC Secretary of the University, but the objectives perceived by the two of them were different.

The CPC Secretary's understanding of the purpose of the UTE moved away from assuring the quality of undergraduate teaching. In the speech delivered by the CPC Secretary to announce the start of preparation for the UTE, he said:

> The UTE is a test of both university and department leaders' leadership skills, especially university leaders' leadership skills. Since it is a test, we have to perform well in the test…the UTE offers a great opportunity to present our university's achievements and a good result will enable us to compete for more resources and gain us the right to negotiate with the government.
>
> (Uni985 Yearbook, 2007, 97)

As illustrated in the quote, the CPC Secretary considered that the purpose of going through the UTE was to achieve a good result, because a good result would prove good leadership skills and the legitimacy to ask for more resources.

On the other hand, in the speech delivered by the President to announce the start of the preparations for the UTE, he made a contrasting point. Unlike the CPC Secretary, he considered that getting a good result in the UTE was not good enough:

> What the UTE asks for are just some basic requirements that any university should meet. It mainly focuses on basic facilities and infrastructure that a university should provide to students. Even if our university were rated 'A' in all indicators, it would only mean that our university is qualified to recruit students. It would not mean that our university is able to cultivate high quality students. As a university with a long history, we should feel ashamed in front of our university's founding fathers if we are only able to meet the UTE requirements…Hence, I think that going though the UTE should not be our aim. We should regard it as a new starting point of our ongoing journey of cultivating high quality students.
>
> (Uni985 Yearbook, 2007, 98)

The President might still feel disappointed even if the university were rated excellent in the UTE. The main objective perceived by the President was to produce high quality graduates.

As the *objective* of the UTE set by the Ministry of Education was to improve the quality of undergraduate teaching, the *objective* perceived by the President broadly fitted in with the government's *objective*. The objective perceived by the CPC Secretary shifted from improving the quality of undergraduate teaching to achieving a good result in the UTE. The divergence can be attributed to the nature of the UTE as an *authority tool*. The UTE was initiated and conducted by the State Council. As mentioned before, the State Council had control over many *resources*. Bad results in the UTE might be punished by a cut in funding, a dismissal of university leaders or a

decrease in quotas. Conversely, good results from the UTE might be used as a bargaining tool to negotiate more funding and more quotas. The CPC Secretary's concern was a reflection of the government's strong control.

After 2005, especially in 2007, the Project 985 University issued a series of documents to prepare for the UTE. They were: the *Notice on Launching '3×one' Teaching Quality Monitoring Mechanism* (Uni985 Yearbook, 2005), *UTE Preparation Work Schedule, the UTE Preparation Work List One and UTE Preparation Work List Two* (Uni985 Yearbook, 2007). These documents can be regarded as the university's *action plans* for the UTE. The '3×one' referred to three requirements for university leaders and the department heads. Each term they had to observe one lesson for undergraduates, as well as check one class's examination papers and one class's dissertations.

The *UTE Preparation Work Schedule* (Uni985 Yearbook, 2007) designated the *organisation*. It established dedicated groups at both university and department levels to take responsibility for the UTE preparation. The university dedicated group was headed by the President and the departmental dedicated group led by heads of the departments. The new *organisation* overlapped with existing administration structures. Moreover, the *UTE Preparation Work Schedule* did not mention *financial or human resources* used for the preparation. As a result, the existing *financial* and *human resources* and *organisation* were redirected to the UTE preparation and the UTE caused a distraction from the university's routines.

In the *UTE Preparation Work List One* and *Two* (Uni985 Yearbook, 2007), checking student examination papers and dissertations was given particular attention. Departments were required to check examination papers and dissertations from the past three years against the new standards for compiling, grading and reviewing examination papers and for grading dissertations (Uni985 Yearbook, 2006). Department and university leaders were requested to check a random selection of examination papers and dissertations.

The *UTE Preparation Work Schedule* (Uni985 Yearbook, 2007) also asked departments to check the status of laboratories and libraries and to collect data required by the UTE to compile the department reports and supporting documentation. These documents then became a component of the university report for the UTE. Every department was advised to plan in advance how the UTE evaluators would tour the individual departments and to make sure they were organised to show themselves at their best to impress the evaluators. The university made some of the departmental tours into a university-wide tour and evaluated these proposed departmental tours in advance of the UTE inspection. The UTE evaluators were then taken on these university tours on their arrival. In addition, department and university leaders were requested to select some teachers' classes to observe. Lastly, the *UTE Preparation Work Schedule* suggested an update of the UTE sections on the university website.

Comparing the *objectives* perceived by the CPC Secretary and the President with the *action plans*, the objective suggested by the President – taking the opportunity of the UTE to improve the quality of graduates – was left out. As the CPC Secretary

expected, the *action plans* centred around achieving a good result in the UTE by emphasising indicators that could be presented to evaluators. Student examination papers, dissertations, laboratories and libraries were concrete evidence that could be presented to evaluators. Other forms of evidence presented to the evaluators on their tours were: well-prepared lessons taught by distinguished academics, curricula, syllabi, teaching plans, textbooks, regulation documents, buildings, computers, sports facilities and so on. Evidence of more abstract indicators that could not be presented to the evaluators directly was documented in the university and the department reports. Therefore, the *action plans* proved the shift of the UTE *objective*.

The university report was designed to show that the Project 985 University had met the requirements of the UTE evaluation scheme. Its structure followed the structure of the evaluation scheme and all indicators were addressed by the report. Not surprisingly, the Project 985 University achieved 'excellence' in the UTE.

Just as agenda setting was absent from the Quality Project, it lacked *programming* too. The university passed the State Council's announcements on to departments, establishing university counterparts of the provincial and national awards and submitting the university award-winners' applications to the provincial education department. The awards received by teachers at the Project 985 University by the end of 2010 are summarised in Table 13.

	National	*Provincial*	*University*
Excellent Course Awards	52	81	160
Distinguished Teacher Awards	7	16	57
Teaching Group Awards	10	NA	NA

Table 13 The total number of the Quality Project awards that the Project 985 University received by 2010

(Data extracted from the Project 985 University website)

5.4 Perceived Impact of the UTE

In order to establish the perceived impact of the UTE, the department heads' accounts are presented first, followed by teachers' accounts.

5.4.1 Department Leaders' Perspectives

When asked about the impact of the UTE, most of the department leaders concluded that the UTE contributed to 'standardisation'. In fact, the standardisation referred to by the department heads was standardisation of documentation and procedures. Head A (DepHum), Head C (DepSc) and Head F (DepSoSc) reported that their departments had started to archive syllabi, examination papers and student examination results since the UTE. Besides archiving, Head A (DepHum) and Head C (DepSc) mentioned new procedures to manage examination papers. Teachers were required to design two examination papers for one exam. After the examinations, students' examination results were analysed to see if they fitted the normal distribution. Head D (DepSc) added standardised procedures to manage dissertations as well.

According to the proposed understanding of impact, standardisation of documentation or procedures can be regarded as the *consequences* resulting from the action of following standardised procedures to produce documents that met the UTE requirements, which were designated by the *2004 UTE Scheme* and university's *action plans* reviewed above. Whether standardisation of documentation would lead to improvement of quality was debatable among heads.

Some heads believed so. Head C (DepSc) equated standardisation with improvement in quality management. Head F (DepSoSc) held a similar view. She believed that,

although standardised documentation might not be necessary for individual teachers, it was necessary for the management of an institution because it could serve as evidence when investigating problems in teaching or hosting inspections. Head D (DepSc) considered that standardisation contributed to maintaining a national standard across China.

The researcher found flaws in the arguments of Head C, Head D and Head F. Head C mistakes standardisation of documentation for standardised management. Head D mistakes standardisation of documentation for maintenance of a unified standard across the university or the whole country. The common approach to maintaining a unified standard is developing benchmarks to describe student learning outcomes in order to achieve consistent learning outcomes across universities (Morley, 2003). Standardisation of documentation, the approach adopted by the UTE, cannot guarantee consistent benchmarks or learning outcomes. Take the case of the design of two examination papers for one exam as an example. Even if all departments have two examination papers for one exam, the content and the difficulty of these examination papers could vary from department to department. Head F mistakes standardised documentation for archiving documents. Standardised documentation means that all documents have to be produced in the same formats, while archiving documents can be achieved even when teachers write syllabi and lesson plans in their preferred ways.

On the other hand, some heads did not think that standardisation of documentation would lead to the improvement of quality. Head A (DepHum) was of the view that ‘the UTE was just a formality’. For Head A, the purpose of having standardised

documentation was getting ready for potential future inspections. It was just an administrative task. Head E (DepSoSc) commented that the impact of the UTE was that 'it (the UTE) just made teachers very busy at preparing documents in that period of time', so it was not able to make any substantial contribution to the improvement of teaching quality. Both heads' comments indicated that time and effort were consumed without making real contribution to the improvement of quality.

To assess whether standardisation of documentation would lead to improvement of quality, it is necessary to explore how teachers perceive the UTE and standardisation of documentation, which will be presented in the following sections.

5.4.2 Teachers' Perspectives

When asked whether the UTE had any impact on their teaching, all teachers answered that there was 'no impact', while they in fact did perceive a certain degree of impact. The discrepancy was caused by different interpretations of impact referred to by teachers. 'No impact' meant no impact on day-to-day teaching. This bold answer reflected a feeling of irritation caused by the UTE. Ironically, the feeling of irritation was caused by both the long-term and short-term impacts of the UTE.

The short-term impact was the influences during the period of the UTE. Lecturer B (DepSc) recorded a short-term impact:

> In that year, the whole university was quite anxious. The UTE did not have any impact on me. If at all, it influenced our dress code in that year. Usually when we are doing experiments, we will wear normal clothes, because our experiments are clean. However, that year all teachers were expected to wear white lab coats, while conducting experiments. In fact, it was unnecessary.
>
> (Lecturer B, DepSc)

This quote illustrates the discrepancy between the perceived and actual impact of the UTE. When Lecturer B said the UTE had no impact on him, he meant that the UTE no longer had any impact on his day-to-day teaching. However, the UTE had had some impact on his dress code during the time of inspection. This impact slightly annoyed Lecturer B, because wearing white lab coats was an unnecessary side effect and it would make no contribution to teaching quality. Therefore, the practice was discontinued after the UTE.

Another short-term impact was mentioned by Associate Professor A (DepHum):

> The UTE did not have any impact on me. As you might know, this kind of evaluation is just a formality… If there was any impact, it was that I spent a little extra time on preparing the lesson for that week. The impact of this kind of evaluation is quite limited.
>
> (Associate Professor A, DepHum)

This quote again illustrates the same discrepancy mentioned by Lecturer B. Making teachers spend extra time preparing lessons could be regarded as a positive impact of the UTE on teaching. However, this impact did not last long. Associate Professor A only invested extra time in preparing the lesson for that week in case the UTE inspectors randomly selected his class to observe. Afterwards, Associate Professor A reverted to the time it normally took him to prepare lessons.

Comparing the two short-term impacts mentioned by Lecturer B and Associate Professor A, they differ in the contribution to the quality of teaching. Wearing lab coats is just presentational without any substantial contribution to the quality of teaching, whereas spending extra time preparing lessons will impact on quality. Both practices were discontinued after the UTE. Lecturer B and Associate Professor A's accounts echo Head E's (DepSoSc) comment that 'it (the UTE) just made teachers

very busy at preparing documents in that period of time' and Head A's comment that 'the UTE was just a formality'. During the time of the UTE inspection, teachers put extra effort into presentation of their teaching. In the short term, the UTE is able to promote activities that are either superficial or substantial for the improvement of quality, but their sustainability is an issue.

During the preparation of the UTE, it seemed that the interviewed teachers were not too much burdened by the production of the documents. Professor C (DepSoSc) revealed that this work was mainly undertaken by administrative staff. Teachers' answers of 'no impact' also indirectly confirmed this (teachers at the other two universities had much stronger answers as they were involved in the production of the documents). However, since the UTE, the way in which the documents should be produced was inherited by the Project 985 University. Hence, both the *actions* and the *consequences* related to standardised documents and procedures became long-term impacts. These were exactly what irritated the teachers interviewed. Lecturer E (DepSoSc) expressed her irritation with the grading requirement introduced by the UTE.

> The UTE did not have any impact except making trouble for us. It made us grade the examination papers in a rigid way. We have to follow it point by point. Now we cannot write negative points on the examination papers. It makes us exhausted. The UTE was completely meaningless.
>
> (Lecturer E, DepSoSc)

This quote illustrates how much Lecturer E was irritated by the UTE requirement – writing sub-scores in a particular way. Like many other teachers, she preferred to use negative points to calculate the total score. She wrote down how many points the student lost in each section. In the end, all negative points added together and deducted from the full score (100) amounted to the students' total scores. After the

UTE teachers could not use this method any more. They had to use positive points by writing down the points that students gained in each section and then adding them together for the total scores.

However, this requirement introduced by the UTE was unable to prevent the error mentioned by Associate Professor D (DepSoSc). He revealed a common problem that was found during the preparation of the UTE:

> Following the checks of examination papers, the department has found many errors in grading student examination papers. When adding the sub-scores together to get a total score, many professors, including me, missed some of the sub-scores. Sometimes the shortfall amounted to ten points or more.
>
> (Associate Professor D, DepSoSc)

Errors in calculating total scores were critical mistakes. They prevented students from receiving the scores they deserved and resulted in unjust evaluations of students. This might have substantial consequences. For instance, if a student failed because of miscalculation, the student might need to spend extra time and money to take a course to compensate for the credits missed. Moreover, bad scores in the transcript might affect the student's employment prospects.

However, the UTE requirement still needed teachers to add up sub-scores, so the same error in calculation could happen again. Possible solutions to prevent this error would be a procedure to double check the calculation by the same teacher or by different teachers, or a procedure of writing either the positive points that students gained or the negative points that students lost in each section into a column and then adding numbers in this column to get total scores.

Unfortunately, none of these measures were adopted by the Project 985 University. When Associate Professor D was asked if measures were taken to eliminate calculation errors subsequent to the UTE, he answered:

> I am not sure that all the student examination results have been correctly calculated unless all examination papers are re-examined.
>
> (Associate Professor D, DepSoSc)

The quote indicates that after the UTE the student examination papers were not checked again. Neither the university nor department took measures to prevent the error from happening again.

5.4.3 Summary

The interviewed heads and teachers' perceptions of the impact of the UTE are summarised in Figure 4 and respectively Figure 5. As illustrated by the two figures, the heads and teachers confirmed the *implementation actions* of following standardised procedures and producing standardised documents according to the UTE's requirements. Moreover, the university's *action plans* also added new procedures, such as wearing lab coats, in order to impress the evaluators. This reflected the *objective* identified by the Project 985 University, which was to 'Excellence' in the UTE.

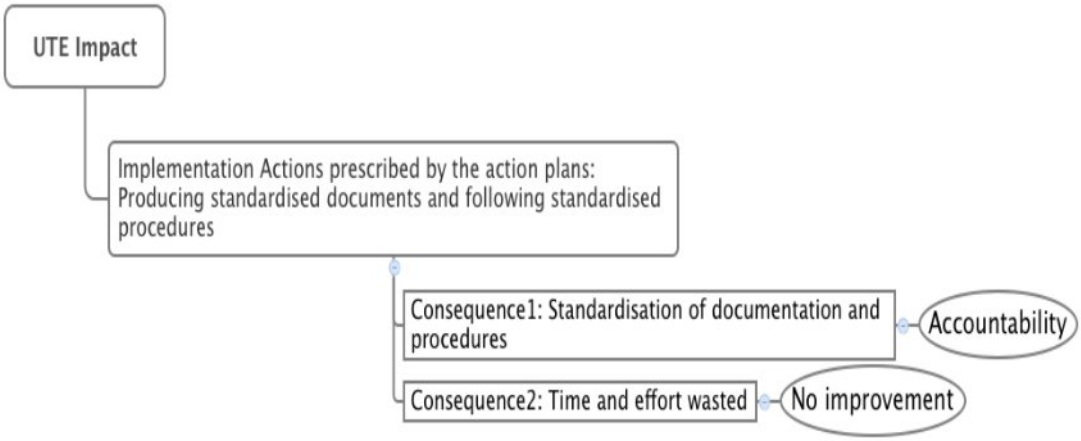


Figure 4 The impact of the UTE as perceived by head interviewees at the Project 985 University

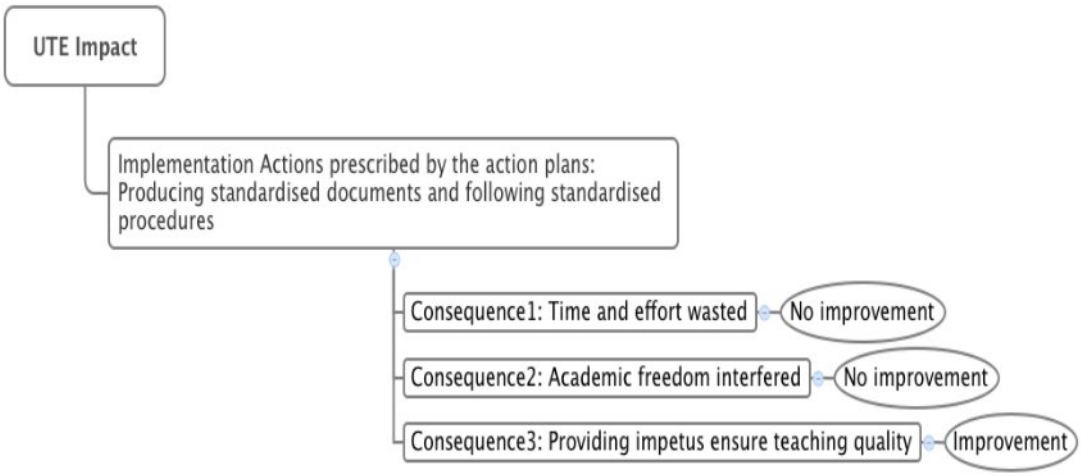


Figure 5 The impact of the UTE as perceived by teacher interviewees at the Project 985 University

Some heads and teachers agreed on a negative *consequence* of the UTE – a waste of time and effort. This was a *consequence* of the *implementation actions.* They considered it a waste of time and effort because they did not think that these actions alone could lead to the improvement of quality. In contrast, some heads were convinced that these *implementation actions* resulted in a positive *consequence* – the standardisation of documentation and procedures. They perceived it positively because it contributed to the fulfilment of accountability. However, none of the interviewed teachers considered the fulfilment of accountability to be the assurance of

quality. The disagreement lay in different understandings of quality. Those heads in favour of standardisation regarded 'quality as perfection and consistency', while other heads and teachers did not think so.

Some teachers mentioned two more *consequences*, the interference in academic freedom and the provision of impetus[30] to ensure teaching quality. As mentioned before, the Project 985 University's *institutional rule* promoted academic freedom. Therefore, when the particular method of grading examination papers was imposed by the UTE, teachers felt that their academic freedom was interfered. The provision of the impetus to spend extra time preparing lessons was a positive *consequence* because of its potential to contribute to the improvement of quality. Unfortunately, it disappeared with the UTE.

When comparing the *public problems* and the expected actions in Table 9 with the actions reported by teachers, it can be concluded that the UTE was unable to fulfil its purposes. The actions reported by the interviewed teachers covered some areas of the expected actions by the UTE, which are: to teach undergraduate courses, to demonstrate professional ethics and commitment and to update syllabi (the rest of the expected actions were not mentioned teachers at all). However, either the action had been taken long before the UTE or the UTE requirements were unable to generate the expected outcome. All the interviewed teachers had been teaching undergraduate courses for years (see Table 4) without the UTE. Spending extra time preparing lessons, mentioned by Associate Professor A, was extra commitment added on top of the commitment that Associate Professor A would normally make. Without the UTE,

30 Impetus is the driver to motivate teachers to ensure teaching quality.

Associate Professor A would also spend as much time and effort needed to prepare lessons. Regarding syllabi, what the heads and teachers mentioned was preparing documents according to the format required by the UTE, not updating the content of syllabi in order to improve them.

5.5 Perceived Impact of the Quality Project Awards

To evaluate the impact of the Quality Project awards, the department heads' accounts are presented first, followed by teachers' accounts in order to cross-check the impact of the Quality Project awards. Finally, the impact of the Quality Project awards is summarised.

5.5.1 Department Leaders' Perspectives

The Quality Project awards were *incentive tools*, which provided both financial and honorary incentives. The interviewed department heads at the Project 985 University mentioned two *consequences* as the impact generated by the Quality Project awards. The first *consequence* was the acknowledgement associated with the awards, especially national awards. As the winners of the national Excellent Course Awards, Head A (DepHum) and Head B (DepHum) confessed that it was a great honour for them to win national awards. It was a great honour for the department as well because teachers at the Humanities Department won three national Excellent Course Awards, one national Distinguished Teacher Awards and one Teaching Group Awards, despite there being only 40 teachers (see Table 14 for the comparison of awards received by the three case study departments).

The second *consequence* of the Quality Project was that it drew teachers' interest to teaching, which was mentioned by Head C (DepSc), but Head B (DepHum) and Head D (DepSc) disagreed. Head D acknowledged that the Quality Project was a

government gesture to put more weight on teaching, but he doubted whether teachers who were not interested in teaching would be lured into teaching. He pointed out that the financial incentive provided by the Quality Project awards was dwarfed by incentives to research, as illustrated in the quote below:

> There is a culture in China that emphasises research while neglecting teaching, especially in research-intensive universities. What the central government did (the launch of the Quality Project) was not enough to strike the balance. All in all, the central government invested little in teaching. [If the central government wants to maintain a balance between teaching and research,] at least it should treat those who win the Distinguished Teacher Awards and the Excellent Teaching Awards and those who perform well in research equally. To strike a balance, the department lacks the capacity; the university may have the capacity; it has to rely on the central government.
>
> (Head D, DepSc)

This quote echoes *University Institutional Rule Two* that the teachers should work hard to improve their research performance in order to generate more income for the university. Those who performed well in research were able to receive funding from various government initiatives, such as research grants, Project 985 and Project 211, whereas the only source providing funding for teaching was the Quality Project awards. Head D believed that the responsibility lay with the central government, because it controlled the *financial* and *legal resources*. It could allocate equal *financial resources* to teaching or tailor the rules to grant teaching equal status.

Furthermore, Head B (DepHum) contended that, instead of being prospective – cultivating teachers' interest in teaching – the nature of the Quality Project awards was retrospective – providing financial rewards to teachers who had been interested in teaching and had been able to demonstrate their teaching achievements.

The silence of both heads at the Social Sciences Department also indicated the lack of impact on non-winners or winners of awards given by the university. As mentioned

earlier, the central government was not alone in granting awards; provincial governments and universities were required to offer matching awards as well. Table 14 summarises the awards won by the three departments.

		DepHum	DepSc	DepSoSc
Excellent Course Awards	National	3	4	0
	Provincial	5	7	0
	University	6	17	2
Distinguished Teacher Awards	National	1	1	0
	Provincial	2	1	0
	University	5	3	0
Teaching Group Awards	National	1	1	0

Table 14 The number of awards of the Quality Project at different departments at the Project 985 University

(Extracted from the Project 985 University website)

The Humanities and the Sciences Departments won university, provincial and national awards. However, the Social Sciences Department only received two university Excellent Course Awards. The fact that neither head at the Social Sciences Department mentioned the impact of the Quality Project awards suggested that university awards alone were not substantial enough to create impact.

In terms of whether the Quality Project awards were able to improve the quality of teaching, some heads interviewed at the Project 985 University expressed their doubts. Head B (DepHum) said, 'it is impossible to expect [the Quality Project] to bring any substantial change in teaching'. Head E added:

> As I said before, improvement of teaching is a continuous process. Those various 'projects' are like waves of movement. After the waves passed by, things went back to before. I don't think they are good ways to improve teaching. However, the bureaucrats, who govern higher education, have money and they want to control. They like to do those kinds of things.
>
> (Head E, DepSoSc)

Head B did not think the Quality Project awards would improve the quality of teaching, because teachers who were interested in teaching would continue putting

effort into teaching as they had been doing before the Quality Project; teachers who were not interested in teaching would not be changed by the Quality Project awards. After witnessing many government initiatives, Head E was tired of the shifts of government policies. He attributed these shifts to political needs. The government needed to update its policies to refresh its influence over higher education.

This section has shown that there were mixed opinions among the heads towards the impact of the Quality Project awards. To assess heads' opinions, teachers' perspectives on the Quality Project awards will be presented below.

5.5.2 Teachers' Perspectives

The interviewed teachers' accounts suggested that only award-winners, but not all, could perceive the impact of the Quality Project. Echoing Head A and Head B, Associate Professor A (DepHum), a winner of a provincial Excellent Course Award and a provincial Teaching Group Award, also mentioned acknowledgement as a *consequence* generated by the Quality Project awards. He thought that the acknowledgement offered by these awards provided extra motivation for him to improve teaching. However, not all winners perceive awards as having an impact. Professor B (DepSc), the winner of two national awards for textbooks, did not think the acknowledgement brought by the two awards was able to provide enough motivation for him to continue compiling textbooks.

Why were some winners motivated by awards while others were not? The researcher would argue that it was determined by the substantiality of financial and honorary incentives. The Excellent Course Awards that Associate Professor A won provided a substantial amount of financial incentive (¥100,000 CNY or £10,000 GBP), while the

awards for textbooks that Professor B received did not provide any financial reward. In addition, the Excellent Course Awards brought more honour to winners because they had more exposure in public. There was a national website to exhibit all award-winning courses. Universities were requested to build dedicated websites to bring exposure to all award-winning courses too. These websites provided space for each course to present its syllabus, lesson plans, coursework and student feedback. For national awards for textbooks, there was just a national website, listing awards winners' names and the titles of the textbooks. Hence, the financial and honorary incentives provided by even national awards for textbooks were not substantial enough to motivate winners.

Associate Professor A (DepHum) added another *consequence.* When recalling the action of filing the application, he had mixed opinions on whether the action would contribute to improvement. He said:

> Filing the application is just a formality, which does not have any direct impact on teaching. The only benefit from doing it is that it also provides an opportunity to reflect my teaching.
>
> (Associate Professor A, DepHum)

On the one hand, Associate Professor A thought the action of filing the application would not improve teaching directly; on the other hand, the reflection that happened in the process of filling out the application form can help him to improve teaching. However, Professor C (DepSoSc), who had also filed an application, did not mention the benefit of reflection. This indicated that award winners tend to acknowledge the positivity of filing applications.

The teachers' accounts also confirmed the retrospective nature of the Quality Project awards. Associate Professor A (DepHum) informed the researcher that he and his

colleagues had been working on this course for many years before they received the provincial Excellent Course Award and Teaching Group Award. Associate Professor C (DepSc) was teaching a course that received the National Excellent Course Award and the teaching group that taught this course also received the National Teaching Group Award. He said that the course and the teaching group had existed for two decades. The retrospective nature meant that the awards were not designed for cultivating interest in teaching. In addition, this proved that the awards were not a motivating factor for those teachers to invest in teaching, because they had been doing so for many years without any awards.

As a teacher who applied for textbook-related Quality Project awards but did not win, Professor C (DepSoSc) did not try to learn from the award-winners to improve, as the policy-makers wished. Instead, he doubted the reliability of the awards. Based on an extensive review of textbooks in his subject area, he was confident about the quality of his textbook. He explained:

> Before writing the textbook, I conducted a review of all the textbooks in this subject that were being used at that time in China. I found that most of them were quite similar to each other, because they copied ideas from Taiwanese textbooks. This applied even to those textbooks awarded by the Ministry of Education. Nine out of ten ideas came from Taiwan…Hence, at that time I started to introduce new ideas from other countries [by compiling my textbook].
>
> (Professor C, DepSoSc)

Professor C attributed the failure to the bias of the reviewers who were appointed by the Ministry of Education. He was convinced that the reviewers were biased towards the textbooks published by the press affiliated to the Ministry of Education. This was just Professor C's suspicion but it reflected a wider phenomenon, a lack of confidence in the fairness of government practice. The lack of confidence resulted from the lack of transparency. Only the winners' names and titles of books were published. The

information about the applicants, reviewers and how winners were selected remained unclear. Not surprisingly, there were doubts about the reliability of the Quality Project awards.

For those teachers who did not apply for awards, they were either unaware of them or doubted their validity. If teachers were not aware of the Quality Project awards, the awards would not have impact on them. The interviewed teachers who had not heard of the Quality Project were lecturers, as they were excluded from the awards. This is the same reason that Associate Professor D did not apply for awards. He said:

> I don't think the Quality Project had any impact on me. The Distinguished Teacher Awards only targeted professors who have established academic reputations. It has been revealed recently that those who receive Distinguished Teacher Awards are often academic officials (academics who assume some senior managerial roles)…As a result, the Distinguished Teacher Awards are only a formality. It is not about the quality of teaching.
>
> (Associate Professor D, DepSoSc)

This quote explains the reason that many lecturers and associate professors were excluded from the Quality Project awards. They believed that awards were biased towards those who had an established reputation, which was established through research performance, or those who assumed senior managerial roles, such as the presidents of universities or department heads. This belief is supported by Wang's (2010) research on the criteria of Distinguished Teacher Awards, which finds that teaching-related criteria are outweighed by research-related criteria.

5.5.3 Summary

The impact of the Quality Project awards perceived by the heads and teachers is summarised in Figure 6 and Figure 7 respectively. With regard to the Quality Project awards, the *implementation action* is filing the application and the *interaction* is receiving awards.

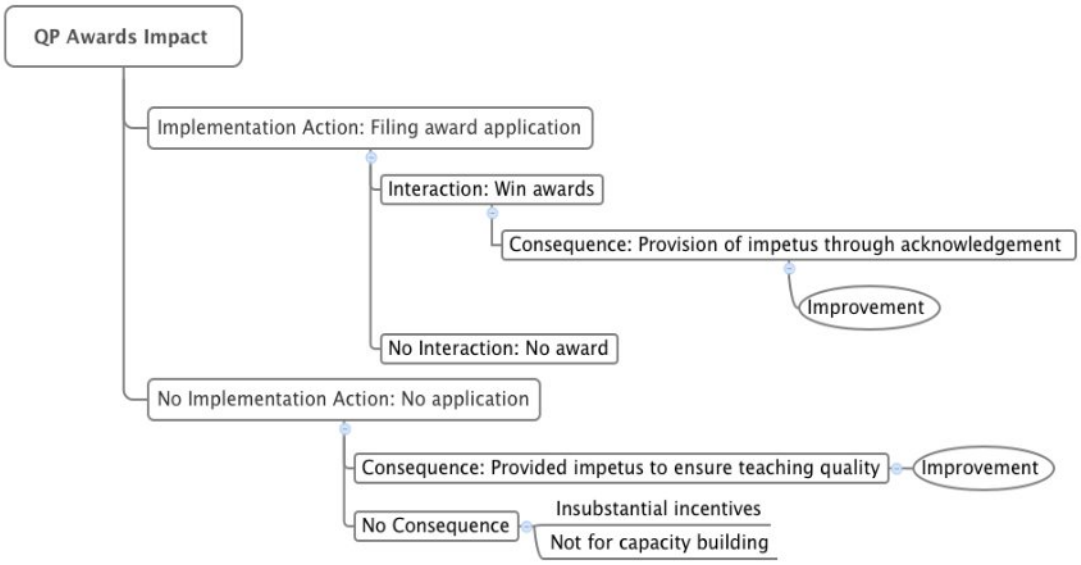


Figure 6 The impact of the Quality Project (QP) awards as perceived by head interviewees at the Project 985 University

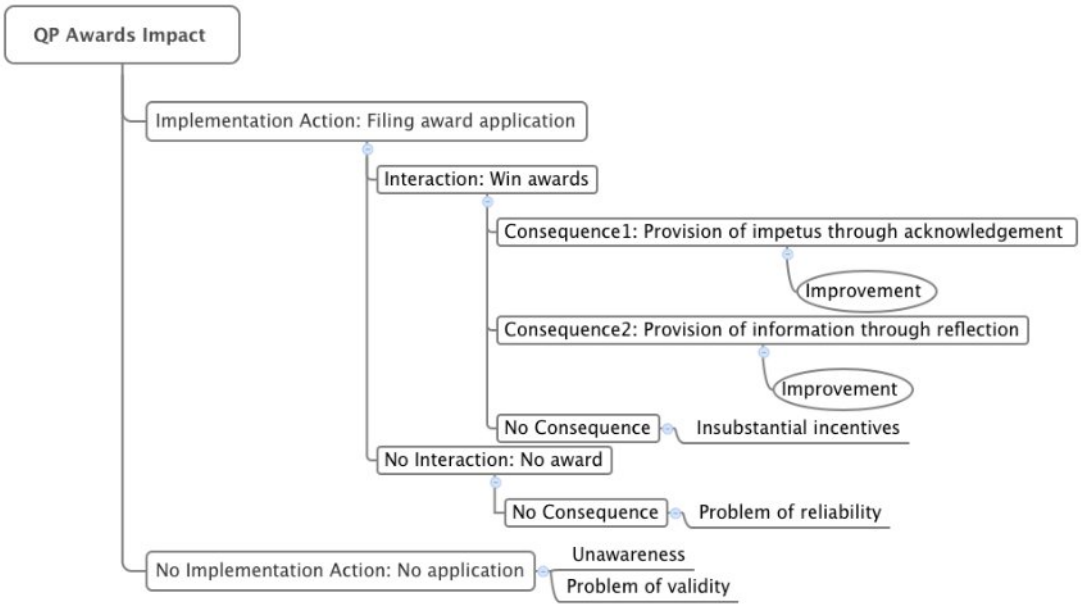


Figure 7 The impact of the Quality Project (QP) awards as perceived by teacher interviewees at the Project 985 University

Those heads and teachers who had received Quality Project awards agreed on the acknowledgement offered by the awards as a *consequence*. However, the silence of

the heads at the Social Sciences Department suggested that university awards were not able to offer substantial incentives for acknowledgement.

The interviewed heads at the Project 985 University disagreed on whether the awards were able to offer incentives to draw interest to teaching or not. None of the interviewed teachers reported that their interest in teaching was developed due to the incentives offered by these awards. The retrospective nature of the awards suggested by Head B (DepHum) was confirmed by Associate Professor A (DepHum) and Associate Professor C (DepSc). The retrospective nature of the awards decided that these awards were unable to develop the interest in teaching.

There were more discrepancies between the heads' and the teachers' perceptions. Head C (DepSc) thought that insubstantial incentives accounted for the lack of *implementation action*, whereas Professor B's (DepSc) perception indicated that insubstantial incentives resulted in the lack of *consequences*. The interviewed teachers added two more reasons for the lack of *implementation action*, which were not mentioned by the heads. Several interviewed lecturers were unaware of the Quality Project awards and Associate Professor D (DepSc) pointed out that this was attributed to the problem of validity. Besides offering acknowledgement, Associate Professor A (DepHum) provided one more positive *consequence*, which was generating information through reflection. It was the *consequence* of the *implementation action* of filing an application.

In terms of the expected actions listed in Table 9, none of the interviewed teachers mentioned them. This suggested that the Quality Project awards were unable to generate the expected actions and thereby unable to resolve those *public problems*.

5.6 Perceived Impact of Class Observation

Unlike the external quality assurance policies imposed by the Ministry of Education, departments at the Project 985 University have more room to manoeuvre on internal quality assurance policies, especially class observation, which stemmed from peer review. There are three kinds of class observation implemented at the Project 985 University: class observations by leaders, by university inspectors and by department inspectors. There are some variations in how the three kinds of class observation are implemented at the three departments. Consequently, how helpful teachers considered them in contributing to the improvement of teaching quality also varies.

5.6.1 Department Leaders' Perspective

Class observation by leaders was carried out at the Humanities Department and the Social Sciences Department, but not at the Sciences Department. The way in which it was conducted at the two departments was different. At the Humanities Department, the department heads and teachers who had teaching-related administrative responsibilities would go together to observe teachers' lessons several times a term. After the lesson, observing heads and teachers and the teacher who conducted the teaching would sit together to discuss the feedback.

At the Social Sciences Department, only the head who was responsible for undergraduate teaching, who was Head E in this case, was required to conduct observation. He observed around seven times a term. For him, it was just an administrative task to complete. Head E did not provide feedback to teachers unless

there was a serious problem. This was because he thought that observing teachers' lessons was interference, so '[he] should try not to bring any influence to bear on teachers' (Head E, DepSoSc).

Whether class observation by leaders was implemented and how it was implemented were determined by the departmental context, in other words, departmental *institutional rules*. The practice at the Humanities Department that a group of heads and teachers went to observe a teacher's lesson and turned the observation into a discussion illustrated that there was a strong sense of community. It also indicated that heads and teachers prioritised their time for class observation. This could happen only when they believed that class observation was an important exercise to improve teaching and they were responsible for helping each other. Moreover, being able to turn class observation into a discussion reflected that they regarded each other as equal as peers and had trust in each other. The heads and teachers who had teaching-related administrative responsibilities were not regarded as having more authority because of their status. They could also talk frankly without worrying that the teacher would misinterpret their comments as attacks.

At the Social Sciences Department, the fact that only Head E went to observe teachers' classes indicated that class observation was not considered an important exercise that needed the involvement of other heads or teachers. Head E explained that he did not communicate with teachers because of the concern about interference. This reflected that the head was regarded as an outsider and there was a lack of mutual trust between each other. Class observation by leaders was an administrative exercise for accountability.

Class observation by university inspectors was implemented at all three departments, but in different ways in different departments. The Sciences and the Social Sciences Departments reacted passively. Their heads would act only when a teacher's teaching was rated as 'poor'. At that time, the university would contact the department and then the heads would summon the teacher to have a conversation.

In contrast, the Humanities Department cooperated actively. It often recommended candidates to the university to make sure that one of the university inspectors was from the department. The reason was explained by Head A:

> Inspectors from other disciplines may not be able to fully understand our discipline. Inspectors from our discipline know what the key points are and are able to point out major problems.
>
> (Head A, DepHum)

As illustrated by the quote, the Humanities Department intended to ensure that the inspector had the expertise to provide reliable feedback. The second reason was that the heads would be able to receive frank feedback directly from the inspector. By doing so, the Humanities Department was trying to avoid the following incident from happening again. As Head A described:

> One of the incidents happened when the university inspector was not from our department. The inspector might not know how lessons were taught in our department. Then he used one of the teachers at our department as a bad example in a teaching-related meeting. Hundreds of people attended that meeting. Our department felt very bad and the teacher felt very angry. Of course we know that the teacher felt angry on the one hand; on the other hand, he would make some changes to improve his teaching. The exercise still has some value.
>
> (Head A, DepHum)

The quote first indicates a distrust of the judgement made the university inspector. Head A thought that it was the disciplinary barrier made the inspector unable to appreciate the teacher's teaching. Second, the description of the incident is full of

sympathy for the teacher. Head A felt that it was inappropriate for the inspector to point out the name of the teacher and department, because it humiliated the teacher and the department in front of hundreds of people. This reflected a strong sense of community, which made the teachers feel for each other. Third, the quote shows a respect for different opinions. Head A defended the teacher but he also believed that criticisms would help the teacher to improve.

What was consistent at the three departments was there was no direct communication between university inspectors and teachers, so in most cases teachers had no idea of what the university inspectors thought about their teaching. This could be problematic. Without feedback, teachers could not benefit from it. Class observation by university inspectors would become an exercise purely for accountability with little concern for teachers' development. Moreover, the fact that only those who performed badly were informed about the university inspectors' comments would create a negative image of the exercise.

Class observation by department inspectors was implemented at the Humanities and Sciences Departments and it was conducted in a similar way. The departments asked inspectors to inform heads of their comments after the observation. Then, the heads would pass on the feedback to teachers. The only difference was that at the Humanities Department the head would have a chat with the teacher, informing the teacher what the inspector thought about his or her teaching and asking what the teacher thought about the comments, while at the Sciences Department this would only happen when the teacher was rated 'poor' by department inspectors. Both heads emphasised that inspectors' comments were just suggestions. Teachers had the

autonomy to evaluate inspectors' comments and to decide whether or not to accept inspectors' suggestions. After having these kinds of chat with teachers, Head A (DepHum) and Head C (DepSc) witnessed teachers' improvement.

5.6.2 Teachers' Perspectives

Only class observation by department inspectors was perceived positively by the interviewed teachers at the Project 985 University, because inspectors' feedback made teachers aware of their strengths and weaknesses in their teaching, which consequently contributed to improvement. This is illustrated by Associate Professor B (DepHum) and Lecturer B's (DepSc) experiences:

> After the observation, they (department inspectors) told me their opinions and suggestions. I thought that they were very helpful. Their comments made a lot of sense. First, they discussed their observations with me from the perspective of students. Second, they shared with me their experience. Many inspectors were just associate professors until they retired. They were not the best in terms of their research but it was well recognised that they were very good at teaching. I respected them and would like to learn from them.
>
> (Associate Professor B, DepHum)

> I did not notice there was an inspector in the first place. Suddenly, I realised that the inspector was sitting in the last row in the class…He was a very experienced teacher from our department. After the observation, he gave me very good advice. Before they make any suggestion, our department inspectors always start by praising our teaching. They also score us, but they do not let us know the scores. If you know the inspector personally, he or she will have a chat with you. If you do not know the inspectors personally, he or she will chat with you if you ask him or her.
>
> (Lecturer B, DepSc)

The first quote demonstrates that Associate Professor B was very grateful for department inspector's suggestions. Two factors made class observation a positive experience. First, the inspector had a good knowledge of the subject area, a lot of experience in teaching and a reputation of being good at teaching. These enabled the inspector to provide valuable feedback and win teachers' respect. Second, the inspector acted like a peer, communicating and discussing with the teacher his

opinions with the best intention to help him to improve teaching. As a result, Associate Professor B appreciated his help and regarded the observation as an opportunity for development rather than an inspection for accountability.

The second quote shows that Lecturer B' perception was still positive but the degree of positivity was slightly less, compared to Associate Professor B's. He was aware of the dual functions of the classroom observation: for accountability and for development. The fact that Lecturer B mentioned scoring indicated that he was aware of the function for accountability. At the same time, he also thought the inspector's suggestions were helpful for improving teaching. In addition, the quote implies that the relationship between colleagues at the Sciences Department were not as close as those at the Humanities Department. It could be the case that the inspector and the teacher did not know each other. (This was partly attributed to the fact that there were more than 200 teachers at the Sciences Department and only around 40 teachers at the Humanities Department.) In this case, the inspector would not provide feedback to the teacher, unless the teacher asked. If Lecturer B did not, the exercise would be solely for accountability. When Lecturer B asked, the inspector was frank with him, praising his strengths and pointing out his problems by giving suggestions. In this case, the exercise became developmental.

On the contrary, class observation by university inspectors was perceived negatively by the interviewed teachers, because of the problem of reliability and the violation of the *university institutional rule*. The reliability was jeopardised by having inspectors who did not have the expertise in the subject area observe teachers who specialised in it. There were only around 13 inspectors at the Project 985 University. Each inspector

was only specialised in one discipline. It was still unavoidable for inspectors to observe teachers from other disciplines, despite the Humanities Department's best efforts to ensure that one of the university inspectors was from this department.

Without expertise, the interviewed teachers believed that the university inspectors were not able to provide valuable feedback. The following quote represents this common opinion shared by many teachers:

> They (university inspectors) are only able to observe some superficial things, such as whether the teacher has made slides or if the teacher answers a call when teaching.
>
> (Professor A, DepHum)

The university inspectors' feedback was not valuable because they could not make valid comments on the substance of teaching, so teachers did not think the exercise could help them to improve teaching. Under this circumstance, class observation by university inspectors became an exercise purely for accountability.

As mentioned earlier, *University Institutional Rule Four* encouraged academic freedom. As a result, the accountability-only exercise particularly irritated teachers at the Project 985 University, because they considered it a violation of *university institutional rule.* The degree of irritation is illustrated by the incident described below:

> It happened when I was still an associate professor. In one class, when I was criticising the university, suddenly I realised that there was an old gentleman sitting at the back of the classroom. I asked what he was doing here. He said that he came to observe the class. I asked him why he did not ask my permission to observe my class. I told him that what he did showed disrespect to me and he should have introduced himself first.
>
> (Professor C, DepSoSc)

According to university regulation, inspectors did not need to ask for teachers' permission to observe their lessons or introduce themselves. Nevertheless, Professor

C still challenged the university inspector on the basis of introduction as a matter of courtesy. He explained why he resisted class observation by university inspectors. He said:

> Class observation by inspectors shows a lack of trust of teachers. If a teacher is employed to teach, this should have suggested a certain degree of trust that the teacher is able to teach well. Hence, the teacher should not be watched. The teacher should be trusted. Class observation by inspectors delivers a feeling of mistrust and it seems to suggest that the inspectors teaching better than teachers.
>
> (Professor C, DepSoSc)

Professor C pointed out the potential damage the accountability-oriented class observation by inspection could make. Academic freedom relies on self-discipline enforced by trust among academics. As academic freedom was embedded in the culture of the Project 985 University, teachers also cherish the trust. Without trust, teachers would lose the motivation to discipline themselves.

Along with class observation by university inspectors, teachers resisted class observation by leaders. The resistance is illustrated by the following incident. It happened at a time when Professor A (DepHum) was the department head and class observation by leaders was conducted in a more accountability-oriented manner. Professor A described the incident as follows:

> [After I entered the classroom,] the teacher told me not to observe his class. He said that he could not teach if I was there, so I left the classroom with a blush.
>
> (Professor A, DepHum)

Professor A did not feel upset about the teacher's response. On the contrary, she understood it. She interpreted the teacher's action as a protest against her official status – the department head.

The two cases described above were rare cases. They represented how much teachers could be irritated by class observation by leaders or inspectors. In addition, they reflected the culture of respect of academic freedom at the Project 985 University, so that a teacher could openly express dissatisfaction without concern about potential consequences.

On the other hand, some teachers understood the need for classroom observation by leaders or inspectors. Lecturer B (DepHum) explained:

> Sometimes some teachers do not take teaching seriously. There were cases that teachers asked some graduate students to replace them to teach optional courses. They were found out by inspectors… [Class observation by leaders or inspectors] could deter those who do not take teaching seriously.
>
> (Lecturer B, DepHum)

This quote shows that there were cases where some teachers did not fulfil their responsibility and these cases were revealed because of the classroom observation.

Due to the lack of data, it was unclear whether class observation succeed in deterring teachers from being irresponsible or not. Moreover, as mentioned above, class observation by university inspectors was believed to be able to check the form of teaching but not the quality of teaching. Even if class observation made teachers go to and teach, it would not be able to guarantee the quality of their teaching necessarily. On the other hand, this quality assurance mechanism could potentially demotivate teachers who are diligent and who cherish the trust of the university, such as Professor C.

5.6.3 Summary

How the interviewed heads and teachers perceived the impact of class observation is summarised in Figure 8 and Figure 9 respectively. While class observation by

university inspectors was implemented at the three departments, class observation by department inspectors was implemented at the Humanities and the Sciences Departments and class observation by leaders was implemented at the Humanities and the Social Sciences Departments. As the two figures indicated, the heads and teachers took different stances. The heads took the stance of the observers when they were referring to class observation by leaders; they took the stance of the observed when they were referring to class observation by inspectors. By default, teachers were the observed, so teachers did not have to take part in the *implementation action* – observing.

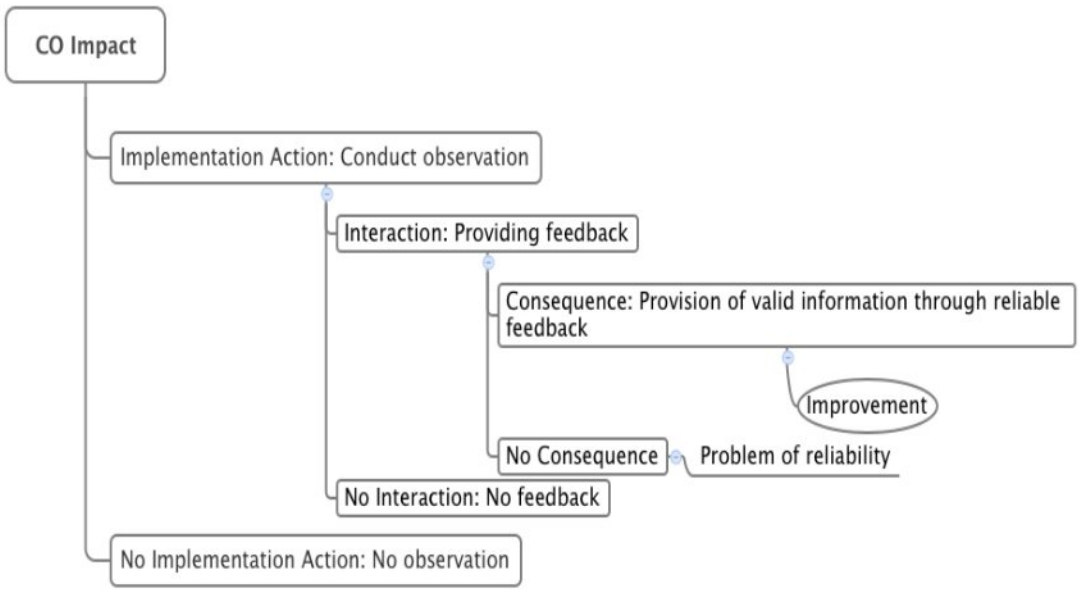


Figure 8 The impact of class observation (CO) as perceived by head interviewees at the Project 985 University

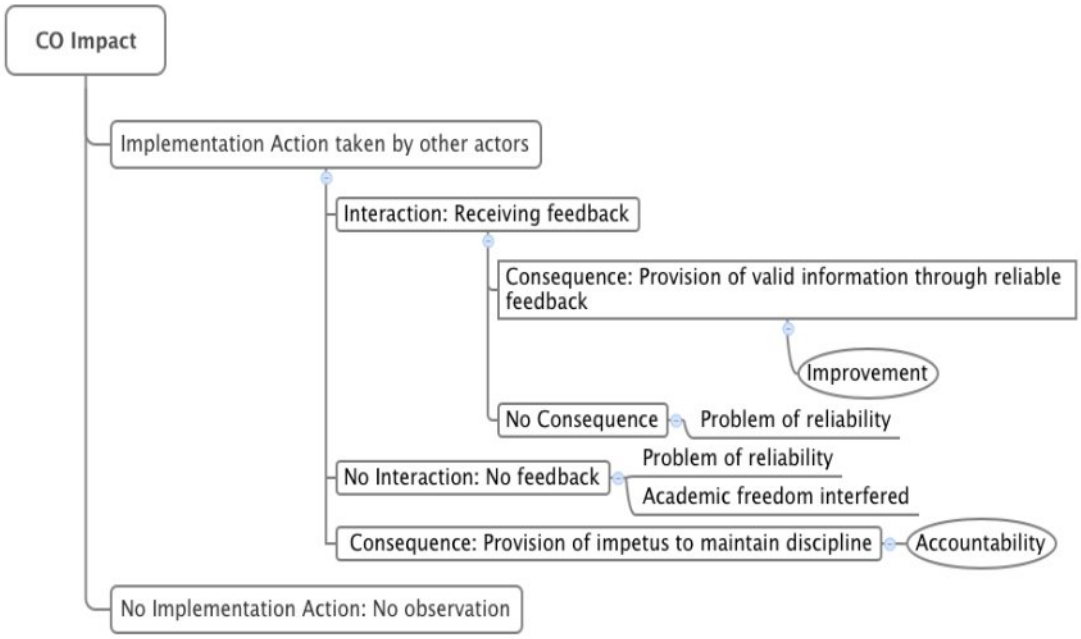


Figure 9 The impact of class observation (CO) as perceived by teacher interviewees at the Project 985 University

The interviewed teachers and heads reached a consensus on the three conditions for class observation to have impact. First, there should be an *implementation action*. In other words, class observation had to be conducted. Second, there should be an *interaction*, i.e. teachers should receive feedback. Third, the observers should be reliable, i.e. inspectors have subject expertise. Once the three conditions were met, the class observation would have a *consequence* of providing teachers with the information regarding their strengths and weaknesses in teaching. This information would enable teachers to improve their teaching.

Furthermore, teachers' accounts provided explanations of why some interviewed teachers were not interested in knowing the feedback. The experiences of Professor A (DepHum) and Professor C (DepSoSc) suggested that the *university institutional rule* of academic freedom could prevent the class observation by leaders or university

inspectors from having impact. Again, the problem of reliability could keep teachers from being interested in knowing the feedback.

Besides the positive *consequence* of providing valid information through reliable feedback, class observation could have another potential positive *consequence*, which was maintaining basic discipline. To have this *consequence*, no *interaction* was required. As long as teachers expected to be observed, class observation would have this *consequence*.

5.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal

The discussion of the student evaluation of teachers and teacher appraisal is combined in this section, because the two internal quality assurance mechanisms are related to each other at the Project 985 University. The heads' accounts are presented first to explain how the results of the student evaluation of teachers affect appraisal decisions, followed by their perceptions of the reliability of student evaluation. Secondly, teachers' accounts are elaborated to investigate the impact on teachers. This chapter ends with a summary section.

5.7.1 Department Leaders' Perspective

The interviewed heads' accounts revealed that departments had adopted the President's suggestion that they should take the quality of teaching into account when making promotion decisions. They used the results of the student evaluation of teachers as the indicators of the quality of teaching. The President's suggestion was implemented as an informal university policy, as is revealed by Head E:

> There is no formal university policy. However, if a teacher ranks among the bottom 30% in the student evaluation of teachers, it will affect his or her

> promotion. I have not seen any written documents about this rule, but in practice it is very important.
>
> (Head E, DepSoSc)

According to the quote above, at the Social Sciences Department the threshold was 30%. Moreover, Head E added that the 30% should be calculated based on the university-wide ranking and normally none of the teachers at the Social Sciences Department would rank among the bottom 30%. Consequently, the pressure of student evaluation on teachers at the Social Sciences Department was not very high. Due to an absence of official policy document, there were some variations in the implementation of this informal policy at different departments.

At the Humanities Department, the threshold was 10% and it was calculated on the basis of the department-wide ranking. Head A elaborated:

> If a teacher ranks among the bottom 10% [of teachers at our department] in the student evaluation of teachers, the teacher will be automatically disqualified from promotion. This has happened at our department…
>
> (Head A, DepHum)

This quote indicates that the pressure of student evaluation on teachers at the Humanities Department was higher because there would be some teachers ranked in the bottom 10% at the department.

The Sciences Department did not use a threshold percentage. Head C explained how the results of the student evaluation of teachers were used:

> The result of the student evaluation of teachers is one of the criteria for promotion…Say there were five teachers applying for promotion. Among them, the one who had the lowest score in the student evaluation of teachers would not get promoted for sure.
>
> (Head C, DepSoSc)

This quote shows that at the Sciences Department teachers had to compete with their colleagues who applied for promotion at the same time. If all applicants had low

scores, a teacher with a higher score among the low scores would win. If all applicants had high scores, even a teacher with a relatively high score among all teachers at the department could lose. The standard was variable.

As shown above, the results of the student evaluation of teachers were only able to affect promotion decisions but not regular appraisal decisions, in other words, decisions to dismiss teachers. Only Head D at the Sciences Department related the results of the student evaluation of teachers to regular appraisal decisions. He said:

> If the teachers always rank in the bottom 10% or 30%, the department will postpone their promotion or change their positions through the so-called elimination mechanism. At Chinese universities, it is very difficult to completely fire somebody. We did a little bit but the academic mobility is still very low. We only fired a very small percentage of these teachers. We only dealt with a dozen of teachers who had issues related to teaching. Some of them were fired. For others, they were transferred to other positions. Those who were not suitable for teaching were asked to work in the labs or were transferred to administrative positions.
>
> (Head D, DepSc)

According to Head D, some teachers who were consistently at the bottom of the ranking could be punished by dismissal or transfer to non-teaching positions, but they made up just a small percentage of those teachers. The reason that it was very difficult to dismiss teachers was because there was no such tradition in China. The positions of university teachers were regarded as permanent positions. This also explained why at the Humanities Department and the Social Sciences Department the results of student evaluation were only related to the appraisal for promotion but not regular appraisal. However, the potential problem of only associating the results of the student evaluation of teachers with teacher appraisal for promotion would be difficulty in motivating teachers who had no intention to seek promotion, mainly professors.

Given the importance of the results of the student evaluation of teachers for promotion, the student evaluation of teachers should be reliable. However, the perception of the reliability of the student evaluation of teachers varied among the heads. Head C (DepSc) had the most confidence. Firstly, he believed that about 80 percent of students would take the evaluation seriously and complete it carefully. Secondly, in addition to the scores, most students at the Sciences Department also provided commentary feedback, so the department could know more about what students thought about their teachers' teaching. Thirdly, in his experience those who taught very well also had high scores in the evaluation. Fourthly, the department analysed the past three years of students' examination scores to see if there was grade inflation and they did not find it. Head C judged: '[the analysis shows that] teachers did not give students higher scores because of the student evaluation of teachers, so there was no grade inflation.' Hence, the student evaluation of teachers should be reliable, he thought.

The researcher found flaws in some of Head C's accounts, especially the last point. Head C conflated the reliability of the student examination with the reliability of the student evaluation of teachers. The analysis of student examination scores can only show the reliability of examinations. The reliability of the student evaluation of teachers has to be analysed by the scores that students give to teachers. They are separate issues. The third perception of the reliability is inadequate. Head C felt that the teachers who taught well also had high scores in the student evaluation of teachers. However, he also admitted that if teachers had slightly low scores it did not necessarily mean that they taught badly. It could be that those teachers were strict with students. In this case, Head C said that they would use students' commentary

feedbacks to try to find out what was the reason that students gave the teacher a low score. Overall, Head C concluded that:

> Promotion is an issue of teachers' interests, so we would not just look at one indicator. However, if the scores in the student evaluation of teachers are too low, it will have a bad effect on teachers' promotion.
>
> (Head C, DepSc)

This quote reveals that Head C was aware of the problem of reliability of the student evaluation of teachers. Hence, the department would use indicators from other sources, such as class observation by leaders or inspectors, to judge teachers' teaching, if the scores in student evaluation were not very low. However, if the scores in student evaluation were very low, Head C would trust student evaluation and believe that there were some serious problems in those teachers' teaching.

At the same department, Head D (DepSc) adopted a slightly more critical view. He summarised his attitude towards the student evaluation of teachers as: 'in general, it is reliable but there were some individual cases requiring more consideration', because after a few years' practice he found that the sample size was big enough. When some teachers ranked in the bottom 10% to 30%, department leaders would talk to teachers and students to investigate the reasons. He also became aware of the problem that students gave low scores to teachers who were strict with them. If this were the cause of the problem, department officials would recommend that teachers to change their approach. If it was a problem of teaching, they would ask other teachers to help the poorly rated teachers to improve teaching, but these teachers would not be promoted before they demonstrated improvement.

At the Humanities Department, like Head C, Head A had confidence in the student evaluation of teachers, although he had some doubts about how students rate their teachers:

> Overall, the results of the student evaluation of teachers match the results of class observation by leaders. We observed many teachers' classes. There were some teachers that we thought their teaching was not great. The results of student evaluation showed the same. However, when it comes to an individual teacher, I don't know how a student rated his or her teacher. Was it because the teacher taught very well [so the student gave the teacher a high score]? Or was it because the teacher criticised the student before or the teacher gave the student a low score in his or her previous examination, so the student gave the teacher a low score? I really don't know. We did not investigate. As long as the overall trend is ok, we think it should be fine.
>
> (Head A, DepHum)

This quote shows that, because of the match between results of the student evaluation of teachers and the results of class observation by leaders, Head A trusted the reliability of the student evaluation of teachers, despite his suspect that some students might be biased.

At the same department, Head B (DepHum) was more aware of the potential biases of the student evaluation of teachers and class observation by leaders or inspectors, so he had confidence in them but only when the results were consistent. He explained:

> The unreasonable elements of these mechanisms (class observation by leaders or inspectors and the student evaluation of teachers) are obvious…It is not the case that, when a teacher receives a bad result of the student evaluation of teachers in one term or he or she receives a bad comment from one inspector, the teacher cannot get promoted. At this university, we look at the overall performance of teachers over a period of time – whether the results of student evaluation and class observation are consistent and whether they are consistent in the past several years. If they are all bad, it shows something. If some are good and some are bad, it is normal.
>
> (Head B, DepHum)

These unreasonable elements referred to by the Head B were the potential biases introduced by class observation by leaders or inspectors and the student evaluation of teachers. Class observation was biased to leaders or inspectors' preference; the

student evaluation of teachers was biased to students' preference. Hence, teachers' performance should be evaluated based on their consistent and overall performance.

The interviewed heads at the Social Sciences Department had the most negative opinions towards the student evaluation of teachers. Head E questioned the idea of using students' preferences to define quality teaching. He sensed that students liked teachers who were able to make teaching fun because they had unreasonable expectations of university:

> During secondary education, there was a university entrance examination to pass, so students were able to accept the kind of education that forces them to study. After they entered tertiary education, they want to compensate the fun that they lost during their childhood, so they are not very interested in learning new knowledge. Hence, students tend to appreciate teachers who make classes fun or even gimmicky. I sense it not only when I go to observe teachers' classes, but also when I talk with colleagues. This sense is quite common among teachers.
>
> (Head E, DepSoSc)

Head E believed that students' judgements were not reliable because they were looking for fun, instead of high quality teaching. The unreasonable expectations were caused by the larger context of Chinese education system, especially the university entrance examination.

Under this circumstance, when the results of the student evaluation of teachers were associated with teacher appraisal for promotion, Head E (DepSoSc) argued that students' preference for fun affected the quality of teaching. Teachers would become concerned about the results of student evaluation, so some of them compromised their teaching to please students. He observed that:

> For example, some teachers do not care about the academic rigor or the systematic completeness of knowledge. Instead, they care more about how to teach some superficial knowledge in a funny way. Also, in relation to the requirements for students, some teachers did not ask students to do many

> assignments or reading; when grading students' examination papers, some teachers tend to maximise the scores that they could give.
>
> (Head E, DepSoSc)

What those teachers did is putting the cart before the horse. They shifted the focus from delivering high quality teaching to please students so that they would achieve a high score in the student evaluation of teachers. Due to these teachers' behaviours, Head E became suspicious of the reliability of the student evaluation of teachers.

Head F (DepSoSc) also had some doubts about the reliability of the student evaluation of teachers. She complained that the students did not take the evaluation seriously:

> Some students have very casual attitudes towards student evaluation of teachers. They do not read the content of each different section. They just fill a number of scores in, based on how much they like the teachers.
>
> (Head F, DepSoSc)

Because she suspected that students rated teachers based on their fondness for the teachers' personalities, instead of the quality of their teaching, Head F considered the results unreliable. Head F added that the casual attitudes to the evaluation were partly caused by the fact that the evaluation was compulsory for students. When students were forced to rate teachers, it might lead to carelessness.

Problems of the student evaluation of teachers mentioned by some interviewed heads were related to its reliability and they attributed it to the students' inability to provide fair responses. Head C and Head D of the Sciences Department found that the students tended to give low scores to teachers who were strict. Head A at the Humanities Department suspected that students might give low scores to teachers who gave students low scores in examinations. Head E of the Social Sciences Department observed that students gave high scores to teachers who made teaching fun during teaching. Head F of the same department sensed that students rated teachers carelessly.

The underlying assumption of their arguments was: if the students took the evaluation seriously and conducted it objectively, the evaluation should be able to reflect the quality of teaching.

However, none of these heads touched on the issue of validity. The issues related to the validity of the student evaluation of teachers that the researcher raised in the previous section were not mentioned by any head. In Section 5.2.1, the researcher pointed out the problem of multiple factors in each indicator in the student evaluation of teachers; the problem of face validity; the problem of whether the design of the student evaluation of teachers is able to reflect students' definition of teaching quality; the problem of using total scores.

Finally, the practice of the student evaluation of teachers was only aimed at accountability. The results of evaluation were mainly used as references to make promotion decisions. Only at the Sciences Department were teachers' rankings in the university and the departmental rankings published. However, teachers were not informed of the students' commentary feedback, which would be more useful to help teachers to improve. At the Humanities and Social Sciences Department, teachers were not informed of any result of the student evaluation of teachers. Even if teachers were informed of the total scores or ranks, they were not very helpful for improving teaching.

5.7.2 Teachers' Perspectives

The perceived impact of student evaluation of the interviewed teachers varied according to teachers' experiences, academic titles and affiliated departments. Associate Professor D, one of the most popular teachers at the Social Sciences

Department, was the only teacher who had a very positive attitude towards student evaluation, because he benefited from it. He shared his experience:

> I got my associate professorship thanks to the student evaluation of teachers. I was nearly eliminated from the promotion selection. However, in the university committee meeting, the director of human resources mentioned that throughout the years I was highly regarded by students. The President said, 'We should encourage the promotion of teachers who are good in teaching. This should be a rule'. The President also praised the director of human resources for mentioning this in a meeting and enabling a teacher good at teaching to be promoted.
>
> (Associate Professor D, DepSoSc)

Associate Professor D's personal experience showed that, although there was an absence of official university policy, student evaluation did play an important role in promotion at the Project 985 University. As a result, he considered that the student evaluation of teachers was a powerful quality assurance mechanism and he was motivated to perform his best.

Professors interviewed did not care about the results of the student evaluation of teachers, because there were no practical implications. Professor B (DepSc) and Professor C (DepSoSc) confessed that:

> I know there is a student evaluation of teachers, but I do not care about it too much. If I teach well, the results should be good.
>
> (Professor B, DepSc)

> The quality assurance mechanisms that I mentioned above (the student evaluation of teachers and class observation by inspectors) cannot influence me at all. First, I am a professor. Second, I am a professor who can supervise doctoral students. Third, the contract that I signed with the university was an indefinite contract. As long as I don't make any political mistake, the university cannot fire me even if I cannot teach well.
>
> (Professor C, DepSoSc)

Professor B did not care about the student evaluation of teachers because he was confident in his teaching. Professor C did not care about it because, even if he had bad results in student evaluation, it would not affect his academic title and job.

Professor A (DepHum) did not care about student evaluation because she did not have confidence in its reliability. Once she received 80 (out of 100) in student evaluation, which was considered bad because most teachers scored more than 90. She thought that she received the low score because she gave students low scores in their examinations. When grading student examination papers, it was very rare for Professor A to give students scores above 90 (the full score is 100). She believed that 80 was already a decent score at the Humanities Department, while at other departments 80 could be a very low score. She did not care what scores other departments or other teachers gave and insisted on being strict in grading. She normally gave 80 or 70 or even failed students. In addition, Professor A said that the low score could be attributed to the fact that she tended to teach deep and difficult content in order to challenge students and she had no intention to make her teaching funny.

In terms of the impact of student evaluation perceived by other interviewed associate professors and lecturers, only those at the Humanities Department reported its impact on them, whereas those at the Sciences and Social Sciences Departments did not care about it. Associate professor and lecturers at the Humanities Department were concerned with student evaluation. On the one hand, the concern reminded them of their responsibility to deliver high quality teaching. This impact is illustrated by Associate Professor B's (DepHum) quote. He said, 'It gives me a sense of awe, so I would not dare to mess it (teaching) up'.

On the other hand, due to the concern, there was a tendency to not be too strict with students in order to avoid low scores, which are illustrated by the following quotes:

> Before I had a concern. I did not dare to be too strict with students, but now I am strict with students when it is necessary. The student evaluation of teachers is like a coin with two sides. [On the negative side,] if teachers are too strict with students, students may take revenge… Not literally revenge. I mean that students give these teachers low scores.
>
> (Associate Professor A, DepHum)

> Maybe when some teachers are quite strict with their students, the students will give them low scores…Although I am concerned about receiving low scores in the student evaluation of teachers, I still teach them in the way that I should and request them to do what they should do.
>
> (Lecturer A, DepHum)

Interestingly, their concerns about being strict with students did not stem from personal experience. Both Associate Professor A and Lecturer A admitted that they had not encountered these kinds of incidents themselves, as they did not know their results of student evaluation. They had just heard from their colleagues about the cases where some teachers were strict with students and then they received low scores in the student evaluation of teachers, which prevented them from being promoted. However, both teachers claimed that they resisted this tendency.

In contrast, other associate professors and lecturers interviewed at the Science and Social Science Departments did not take the student evaluation of teachers seriously, because they did not perceive any consequence of having low scores in student evaluation. As Head E (DepSoSc) explained, it was unlikely for teachers at the Social Sciences Department to rank the lowest 30% in the university ranking. At the Sciences Department, there was no threshold percentage to disqualify teachers from promotion.

In addition, some teachers also expressed no confidence in the validity of the student evaluation of teachers. This view is illustrated by the quotes below.

> I cannot see the results of the student evaluation of teachers. I am only informed of the total scores. I think it is meaningless. My teaching is open. My students can [give me feedback] any time. The total scores [of the student evaluation of teachers] are meaningless. At the end of the course, I collected students' feedback by myself. The feedback that they gave to me was not just one or two lines. It was one or two thousand words and four or five pages. This feedback came from their hearts. Their comments are very specific. I think feedback like this is valid.
>
> (Associate Professor C, DepSc)

> The scores in the student evaluation of teachers are untrustworthy…[To know students' feedbacks,] I rely on the communication with students. I find it out myself and I can feel it myself, instead of official student evaluation.
>
> (Lecturer E, DepSoSc)

Associate Professor C and Lecturer E did not regard the scores from the student evaluation of teachers as valid student feedback. They needed narrative feedback to inform them of what students thought about their teaching, while total scores offered little help to improve teaching. Because the student evaluation conducted by the university failed to provide needed information, they had to collect student feedback by themselves.

Although Head D (DepSc) mentioned the application of the results of student evaluation to regular appraisal, none of the interviewed teachers at the Sciences Department mentioned its impact. This suggested that the teachers were unaware of this department policy.

5.7.3 Summary

Figure 10 and Figure 11 summarise the impact of student evaluation perceived by the heads and teachers respectively. Similar to class observation, teachers did not need to take part in the *implementation action*, so teachers were absent from it. The heads

interviewed across the three departments confirmed an informal university policy that the results of the student evaluation of teachers would affect teachers' promotions. Therefore, the interviewed heads were certain that student evaluation would have *consequences*. However, they disagreed on whether student evaluation provided impetus to ensure high quality teaching or to please students, depending on their perceptions of the reliability of student evaluation. The former would lead to improvement whereas the latter would not.

Figure 10 The impact of student evaluation (SE) as perceived by head interviewees at the Project 985 University

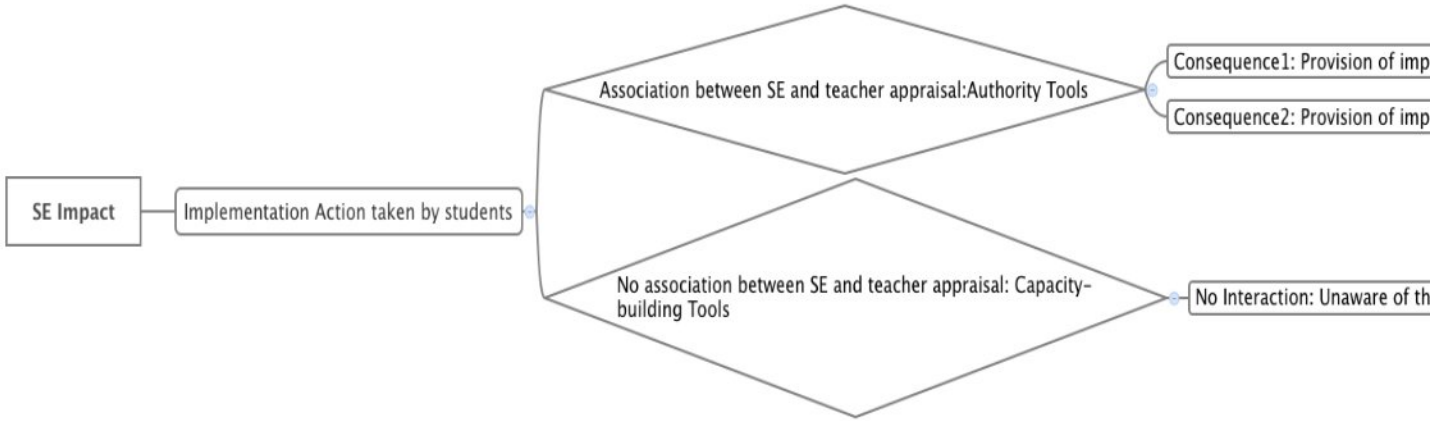


Figure 11 The impact of student evaluation (SE) as perceived by teacher interviewees at the Project 985 University

The interviewed teachers' accounts showed that some teachers did not think that their results in student evaluation could influence their promotions. These teachers did not know their results of student evaluation because either they did not want to or they were not able to. The lack of practical implications and problems of reliability and validity kept these teachers from being interested in knowing their results. Some of the interviewed teachers mentioned the validity issue of student evaluation. Given the limited information provided by total scores, these teachers preferred to collect student feedback by themselves.

5.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching

Based on interview data, this section answers several fundamental questions in relation to quality assurance: what are the problems of the existing quality assurance mechanisms perceived by heads and teachers' perspectives? what drives heads and teachers to invest in teaching? and what actually helps heads and teachers to improve teaching along the way? The answers to these questions will provide some insight into how to improve current quality assurance policies.

The first unexpected finding in relation to these questions is that heads and teachers have the same perspective. This shows that, as the frontline leaders, they are aware of the problems of current quality assurance mechanisms and, as teachers themselves, they also agree with other teachers on what motivates and helps them to improve teaching. However, because they are at the bottom of administrative hierarchy, what they could do to change the current quality assurance mechanisms is very limited.

5.8.1 Problems of Current Quality Assurance Mechanisms

Although the Project 985 University was a research-intensive university, the President had been actively promoting the belief that the undergraduate teaching was a core activity of the Project 985 University (Uni985 Yearbook, 2004; 2005; 2006; 2007; 2008; 2009; 2010). Providing high quality undergraduate education had become a characteristic that distinguished it from other research-intensive universities. After he was transferred from a research institution to the Project 985 University, Professor B (DepSc) chose to teach undergraduate courses, because he felt ' It would be a little bit weird if professors did not teach undergraduate courses at this university'. Professor C (DepSoSc) felt very proud of the fact that two thirds of the undergraduate courses were delivered by either professors or associate professors at his department. He believed that they formed the backbone of the quality of undergraduate education. Considering the results of the student evaluation of teachers when making promotion decisions was also regarded by many interviewed teachers as evidence of the university's endeavour to emphasise undergraduate teaching.

Nevertheless, a large number of interviewed heads and teachers repeatedly voiced their concerns that the money and reputation brought by research tended to offset the importance of teaching established by these quality assurance mechanisms across the three departments. The superiority of research was firstly reflected in promotion, which is illustrated by the following quote:

> At this university, teaching is a necessary condition for promotion, but research is the key. (Head C, DepSc)

Head C's comment confirmed the decisive role that research played in promotion. Although teaching was important for teacher appraisal for promotion, it was insufficient for promotion.

Secondly, the superiority of research was established by the honorary and financial rewards that could be brought by research, as illustrated by following quotes:

> Although the university claims that the undergraduate teaching is the first and foremost thing, it does not give teachers any rewards when teachers perform excellently in teaching. On the other hand, when teachers publish some excellent articles, the university writes news articles about them and holds news conferences for them… Sometimes they even receive free houses as rewards. After witnessing this, of course I would like to put more effort into research.
>
> (Lecturer C, DepSc)

> We have to conduct both teaching and research. These are the responsibilities of teachers at this university. Which one weighs more? In a research-intensive university like this, it relies on our research performance to compete for various research grants, evaluations and rankings. Hence, departments also offer research awards and match research grants. Research awards are money offered to teachers who publish in prestigious journals. Matches of research grants are extra research grants offered to teachers who have won the bidding processes of national or provincial research grants.
>
> (Head F, DepSoSc)

Lecturer C and Head F's quotes reflect the lucrative return of research. Research potentially was able to bring teachers both honorary and financial rewards and the amount of financial rewards, such as houses, research grants provided by the government and extra grants provided by the university, was very substantial.

Head C (DepSc) pointed out the difference between the approaches to assure the quality of teaching and research. He said 'the approach to assure the quality of teaching is by requirement, while the approach to assure the quality of research is by incentive'. This can be explained by differences in policy tools. Quality assurance policies mainly use *authority tools* and rarely use *incentive tools.* The Quality Project is the only case when the *incentive tools* are used for quality assurance purpose. On the other hand, the quality of research is often assured by using *incentive tools*.

Research grants, National Key Disciplines, Project 985 and Project 211 all provide teachers, departments or universities with substantial financial and honorary rewards in order to assure the quality of research. As a result, teachers receive more and stronger incentives for research rather than for teaching.

In order to provide some incentive to teaching, both Head C (DepSc) and Head D (DepSc) suggested financial rewards for teaching, as illustrated below:

> The university could allocate some money as teaching bonuses, which is paid to teachers based on the amount of teaching they conducted. Say, ¥100 CNY for one lesson. I don't think it is too much. If a teacher teaches 300 lessons one year, the teacher will receive ¥30,000 CNY. This is nothing. If teachers take some part-time jobs to teach outside the university, they will earn much more than this.
>
> (Head D, DepSc)

> The university should have a rule like this. If a teacher completes his or her teaching workload and receives very positive comments, the university should offer some research grants to the teacher, so he or she does not have to spend a lot of time on applying for research grants.
>
> (Head C, DepSc)

The incentives suggested by Head C and Head D differed in forms. One incentive would be direct bonuses for teaching. The amount of rewards would be determined by the amount of teaching conducted. It would not differentiate between teachers who teach well and teachers who teach badly. The second incentive would be given to teachers who teach well and whose their teaching is recognised by students, colleagues, leaders or inspectors. In addition, providing research grants to teachers who teach well could help to maintain a balance between teaching and research.

The second problem of the current quality assurance mechanisms was that they cause harm to autonomy. Professor A (DepHum) mentioned the harm to individual teachers. She said:

> We expect ourselves to be self-motivated to perform our best. We could perform better with our sense of honour rather than government interference.
> (Professor A, DepHum)

Self-motivation and a sense of honour were values underpinning academic autonomy. However, the implementation of quality assurance policies made Professor A feel that her self-motivation and sense of honour were replaced by the requirements of quality assurance policies.

Current quality assurance policies and mechanisms also made department leaders lose their initiative in reflecting and finding out quality assurance approaches that suited the departmental contexts. When Head B was asked what he would like to see in terms of further improvement in quality assurance policies, he answered:

> Now governments and the university are asking us to do too many things, it becomes quite difficult for us to think what actually we need to do.
> (Head B, DepHum)

This was the harm to autonomy of leadership. Consistently driven by various government and university policies made Head B too engaged in implementation, so that he lost sight of what was actually needed to improve quality.

Due to these limitations of current quality assurance mechanisms, Head E concluded that 'these quality assurance mechanisms were only able to maintain the quality of teaching above the minimum standard'. In other words, these assurance mechanisms were unable to assure teaching to a high standard.

5.8.2 What Drives Teachers to Invest in Teaching?

As mentioned previously, some teachers were motivated by the student evaluation of teachers. However, among all the interviewees only two teachers mentioned it. The

drivers frequently mentioned by teachers were: passion for teaching, dedication and responsibility.

Teachers used words such as ‘enjoyment’, ‘interest’, ‘liking’, ‘affection’ or ‘pleasure’ to describe their passion for teaching. They enjoyed the process of teaching, feeling happiness and a sense of satisfaction. For instance, Lecturer E described her passion for teaching in the following quote:

> When I entered the classroom for the first time and started teaching, I felt this is my world. I never felt nervous about teaching. Instead, I was looking forward to meeting students to discuss material and share my thoughts with them. (Lecturer E, DepSoSc)

As illustrated by the quote, it was in Lecturer E’s nature to enjoy teaching.

Moreover, students’ reactions were able to reinforce teachers’ passion. Professor C shared his experience. He said:

> I like to teach students… Also, when students applaud from the first lecture to the last lecture, of course I become more and more motivated. (Professor C, DepSoSc)

Applauding teachers in lectures was not what commonly would happen. This unusual reaction motivated Professor C a lot. The more common reactions mentioned by the interviewed teachers were asking teachers questions. Even subtle reactions, such as students’ expressions showing that they were following teachers were able to motivate teachers.

The second driver was dedication. The term ‘dedication’ used in this research refers to a self-driven will to achieve the personal best. For example, when Lecturer A (DepHum) was asked what drives her to improve teaching, she answered that it was ‘a continuous pursuit of perfection’.

Dedication can be learnt from colleagues. Associate Professor A (DepHum) was impressed by his colleague's dedication:

> Dedication can be passed to each other…For example, there are teachers who prepare new lessons until three or four in the morning and go to class on that morning. They can do it every year. Although I cannot do this, their actions or sprits drive me to do things better.
>
> (Associate Professor A, DepHum)

Although Associate Professor A could not do exactly the same as his colleague, his colleague's dedication motivated him to perform better.

The third driver was a sense of responsibility. The term 'responsibility' used by the researcher in this research refers to a sense of fulfilment of duties and a will to meet others' expectations – not letting others down. Besides being generated by teachers themselves, a sense of responsibility can be stimulated by students, as mentioned by Associate Professor B (DepHum):

> When facing so many energetic students who asked me many naïve questions and who gazed at me with eagerness to learn, a sense of responsibility arose spontaneously.
>
> (Associate Professor B, DepHum)

In this case students helped Associate Professor B to develop a sense of responsibility.

Other people's opinions or the image in others' minds can also motivate teachers to be responsible. Professor B (DepSc) confessed that one important driver was:

> When I teach students, I cannot allow this happen – people think a professor's teaching cannot match a lecturer's teaching…My responsibility is to teach students well. If I could not achieve it, it would be a disgrace.
>
> (Professor B, DepSc)

The title of professor made Professor B believe that he should perform better in every aspect, including teaching. His concern over failing to do so became the driver for maintaining high quality teaching.

Overall, these accounts above indicated that investment in teaching was mainly determined by teachers' personal attitudes, the influence of colleagues and the interaction with students, which had little connection to the formalised quality assurance mechanisms reviewed earlier.

5.8.3 What Helps Teachers to Improve Teaching?

Despite the tension between teaching and research discussed above, many interviewed teachers considered that research helped them to improve teaching. They were convinced that research was what distinguished them from primary and secondary school teachers and made them better-qualified teachers. How research helped teaching is illustrated by the following quote:

> As I have conducted research in different areas related to this course, teaching just goes smoothly. When I saw the title in the slides, I knew what was following. I don't even need slides. All of these areas I have done research and I have published hundreds of articles.
>
> (Professor B, DepSoSc)

This quote shows Professor B's strong belief that research made him a better teacher. He added that his research experience enabled him to provide more insights to students.

These teachers' beliefs that research helps to improve teaching seems to be in conflict with the tension between teaching and research mentioned above. In fact, the tension discussed above is about drivers. Both government and university policies practically provide more incentive to motivate teachers to do research, instead of teaching. In

contrast, the discussion on what helps teachers to improve teaching is discussed under the assumption that teachers have been motivated to teach.

Even when teachers are motivated to teach and conduct research, *time* is an issue. When Associate Professor A (DepHum) was asked what helped him to improve teaching, he said that research was one of the important factors, but he regretted becoming a teacher when he did not have *time* for research. In the first two years, Associate Professor A spent most of his *time* on teaching, because he was new to teaching. He was struggling to find *time* for research. At this point, he had regrets. However, after the first two years' hard work, he established the foundation of the course and started to be able to maintain a balance between teaching and research. Moreover, he found that concentrating on teaching in one term and then on research in the other term was beneficial. He praised the support from his department. If teachers requested it, the department would try its best to meet teachers' needs.

Associate Professor A's experiences reflected that the key to enable teaching and research to be compatible was *time*, when teachers were equally motivated to do both. The university policy should allow teachers to have sufficient *time* for both. The arrangement of dividing teaching and research by terms was very helpful. Like Associate Professor A, many interviewed teachers at other departments also praised their departments for helping them to make this arrangement if it was possible. In addition, many interviewed teachers at the three departments said that the current teaching workload was the maximum. If they took on a greater teaching workload, it would affect the quality of both teaching and research.

Students also played a very important role in helping teachers to improve their teaching. Many interviewed teachers shared the feeling that 'as time goes by, I am improving my teaching by learning from students' feedback' (Head F, DepSoSc). Head F collected students' feedback from questionnaires she administered by herself. She distributed sheets with a few open questions in the classes and asked all students to write comments anonymously. Other interviewed teachers also collected students' feedback through more casual occasions, such as eating in the canteen together with students.

Another source of help came from colleagues, which was mainly mentioned by those interviewed teachers from the Humanities Department. Head A (DepHum) mentioned the mentoring for new teachers. When he was a new teacher, he had observed an experienced teacher's classes for one year. The tradition to provide mentoring to new teachers was continued nowadays. In fact, Head C (DepSc) also mentioned the mentoring at the Sciences Department. Not only systematic mentoring, but also casual chatting, was perceived as helpful. Lecturer A (DepHum) mentioned that chatting with colleagues in the canteen or on the bus also inspired new ideas to improve teaching. At the Humanities Department even retired teachers were happy to offer help. Associate Professor B (DepHum) described how easily he received help from retired teachers:

> At our department there are many lovely old gentlemen. Even after they retired, they often come to the department. For instance, Professor Cao often arrives at the department around nine or ten in the morning. He will read some newspapers and magazines first and sit downstairs. He will approach you and have a chat with you. There are many retired teachers like this.
>
> (Associate Professor B, DepHum)

Because these retired teachers were available and approachable, Associate Professor B leant a lot from them and really appreciated their help. Teachers still would like to

go to the department to help young teachers to improve teaching even after retirement, reflecting a culture of teaching at the Humanities Department.

5.9 Chapter Conclusions

There were three features that distinguished the Project 985 University from the other two case study universities. Firstly, a respect for academic freedom was shared by the university leadership and teachers. Secondly, there was a paradox that the university policies maintained and mediated the tension between teaching and research. Thirdly, the Humanities Department at the Project 985 University could serve as a model of quality culture.

Empowered by respect for academic freedom, there were cases that teachers at the Project 985 University protested against some quality assurance mechanisms, whereas no form of protest was reported by teachers at the other two universities. The point that academic freedom nurtured self-discipline to assure quality was only mentioned by interviewees at Project 985 University. The implication of having more academic freedom was that the teachers were less likely to have a cynical view of quality assurance policies in comparison to their counterparts.

Although it did not appear in the top priorities, the President had actively promoted a belief that undergraduate teaching was a core business of the university. The minimum teaching workload was an example of it. The Project 985 University was the only university where the minimum teaching workload was effectively implemented among the three case study universities. The President also instigated an informal university policy that associated the results of student evaluation with promotion. Although it was an informal policy, it was more effective than the similar

formal policies at the Project 211 University and City University, because a case illustrating the implementation of this policy was reported. This belief that undergraduate teaching was a core business of the university was shared among many interviewed teachers. Even professors considered teaching undergraduates to be an essential part of their work. Providing high quality undergraduate teaching had become a characteristic that distinguished the Project 985 University from other two.

On the other hand, research was another important core activity of the university, which was reflected by *University Institutional Rule Two*: to work hard to improve research performance in order to obtain resources. The rewards for research included both the financial and the honorary rewards offered by the national policies, such as research bidding, the Project 985 and the National Key Disciplines. Not only were they the major sources of the university's revenue and of teachers' research grants, but also they provided solid indicators of the university and teachers' reputation.

The tension between teaching and research was reconciled by a *consensus* that delivering quality teaching was essential for survival and conducting quality research was essential for thriving. Moreover, another *resource* offered by the Project 985 University to help teachers to mediate the tension was *time*. The university gave teachers the freedom to allocate their *time*. Many of the teachers interviewed confessed that they preferred to allocate teaching and research to different terms, which allowed them to concentrate on teaching in one term and on research in another term. Otherwise, they would not have achieved what they had achieved in term of both teaching and research. On the contrary, the freedom to allocate their *time* was not enjoyed by teachers at the other two universities.

The Humanities Department at the Project 985 University was a special case. There was a quality culture at this department. As mentioned above, the heads and teachers interviewed trusted in and were willing to help each other. This nurtured a culture of constructive criticism, which enabled the heads to turn class observation by leaders into peer review and teachers to provide frank feedback. There was also a real sense of collaboration among teachers, which made them willing to help and seek to inspire each other. Furthermore, teachers at this department were devoted to teaching but at the same time they were excellent researchers too. Therefore, the Humanities Department won more Quality Project awards than any other case study department.

The researcher would argue that there were several factors contributing to the quality culture at the Humanities Department. Firstly, the number of teachers at the department was very small, which allowed teachers to develop a close relationship with each other. Secondly, the Project 985 gave the departments a high degree of autonomy and the decision-making process was transparent. This entailed that the majority of teachers could participate in the decision-making process on many issues. This transparency and participation in decision-making resulted in the establishment of mutual trust. The close relationship and mutual trust nurtured the culture of constructive criticism. Thirdly, there was less distraction for teachers at the Humanities Department from part-time offers outside the university and research commissions from businesses. Therefore, teachers could concentrate more on teaching and research.

6 The Project 211 University

The Project 211 University's *institutional rules*, i.e. the context, and existing internal quality assurance mechanisms will be introduced first, followed by the Project 211 University's responses to the UTE and the Quality Project. Department leaders and teachers' perceptions on the impact of these external and internal quality assurance policies will be elaborated to assess whether those policies have contributed to the improvement of the quality of teaching. Teachers' reflection on the problems of these quality assurance policies and factors that contributed to their improvement of teaching will be presented at the end of this chapter.

6.1 The University Context

The university context or *university institutional rules* are created under the influence of the *national institutional rules* and university leadership, which will be described below.

6.1.1 University Institutional Rules Inherited from the National Institutional Rules

Under the same influence from the State Council, the Project 211 University's *institutional rules* inherited from the *national institutional rules* are the same as those acquired by the Project 985 University. They are: *University Institutional Rule One*: to conform to government policies and requirements; *University Institutional Rule Two*: to work hard to improve research performance in order to obtain more resources and *University Institutional Rule Three*: to attempt to solve the problems caused by the significant increase in student numbers.

The Project 211 University was re-established in the 1970s by the State Council. Like the Project 985 University, the President and the CPC Secretary at the Project 211 University are appointed by the State Council. As the State Council or the CPC has the power over appointing and dismissing university leaders, there is a natural tendency for university leaders to obey the government. Moreover, around 40% of the university's total income came from the State Council. In comparison, around 22% of the total income comes from tuition fees (Uni211 Yearbook, 2008; Uni211 Yearbook, 2009; Uni211 Yearbook, 2010; Uni211 Yearbook, 2011). As mentioned in the last chapter, the number of students that universities can recruit and how much universities can charge in tuition fees are also determined by the State Council. This means that around 60% of the Project 211 University's income relied on the State Council. The control of *legal*, *human* and *financial resources* contributed to *University Institutional Rule One*: to conform to government policies and requirements.

University Institutional Rule Two: to work hard to improve research performance in order to obtain more resources is attributed to the great financial pressure that the Project 211 University is facing. According to the Project 211 University Yearbooks from 2007 to 2011, the Project 211 University was running a deficit, except in 2009. Despite the surplus in 2009, the average deficit of each year from 2006 to 2010 was ¥23,765,040 CNY (£2,376,504 GBP). At its biggest the deficit reached ¥102,370,000 CNY (£10,237,000 GBP) in 2007.

To increase income, working on improving research performance is very rewarding. The income of the Project 211 University includes 'central government funding',

‘undertaking income’, ‘operational income’ and ‘other income’. The central government funding and the undertaking income constitute around 70% of the total income of the Project 211 University (Uni211 Yearbook, 2008; Uni211 Yearbook, 2009; Uni211 Yearbook, 2010; Uni211 Yearbook, 2011). The central government funding is subject to the affiliation to Project 211 and the number of students. The undertaking income is revenue generated from teaching and research, such as tuition fees and research grants. As mentioned before, both the affiliation to Project 211 and the student recruitment quota are determined by the State Council considering universities’ research performance. As the majority of the income of the Project 211 University is determined by its research performance, improving research performance appears to be the most sensible decision that the university can make to reduce its deficit.

University Institutional Rule Three: to attempt to solve the problems caused by a significant increase in student numbers is set out by the dramatic increase of student numbers. After the implementation of the national policy of increasing participation in higher education, the number of undergraduate and graduate students more than doubled between 2000 to 2010, while the annual increase of teachers was either 3% or 1% between 2007 and 2010 (see Table 15).

	Number of undergraduate and graduate students	Increase of undergraduates and graduate students	Number of teachers	Increase of teachers	Student–teacher ratio
2000	11,785				
2001	13,261	13%			
2002	15,680	18%			
2003	18,359	17%			
2004	20,490	12%			
2005	21,611	5%			
2006	22,012	2%			
2007	23,497	7%	1,650		14:1
2008	24,530	4%	1,707	3%	14:1
2009	26,493	8%	1,718	1%	15:1
2010	27,771	5%	1,768	3%	15:1

Table 15 The numbers of undergraduate and graduate students and teachers and the student–teacher ratio at the Project 211 University

Extracted from (Uni211 Yearbook, 2007; 2008; 2009; 2010; 2011)

The Project 211 University is neither responsible for the significant increase in student numbers nor the commensurate slow increase in teacher numbers, as the recruitment of both students and teachers are controlled by government quotas. In order to cope with the increase in student numbers, the Project 211 University established another three campuses.

6.1.2 University Institutional Rules Established by the University Leadership

The fact that the Project 211 University was selected for Project 211 but not Project 985 indicates that the position of the Project 211 University among all universities in China is above average but not as high as those selected for Project 985. Without the affiliation to Project 985, the Project 211 University receives much less funding from the local or central government compared to the Project 985 University. From 2001 to 2010, Project 211 provided ¥216,720,000 CNY (£21,672,000 GBP) to the Project 211 University. At the same time, the Project 985 University received ¥3,540,000, 000 CNY (£354,000,000 GBP) from Project 985 only. Moreover, the Project 211 University is less competitive in obtaining research grants from local or central

government. From 2007 to 2010, on average the research funding allocated to the Project 211 University was only 16% of what was granted to the Project 985 University. Nevertheless, the gap between the Project 211 University and the Project 985 University is not unbridgeable. If it tries hard, the university can become a top university, the President of the Project 211 University believes. Hence, the goal set by the President is 'becoming a university well-known at home and abroad' (Uni211 Yearbook, 2008, 55). To achieve this goal, the university should have 'innovative research groups with high potential and departments that are first class in China and well-known abroad to generate significant research output' (Uni211 Yearbook, 2008, 56). This ambition led to *University Institutional Rule Four:* to work hard to improve research performance in order to become a top university in China.

Like the President of the Project 985 University, the President of the Project 211 University expressed the wish to promote academic freedom and autonomy. In his inaugural speech, the Project 211 University President announced three mottos, which were 'wisdomless, 'capabilityless' and 'actionless' (Uni211 Yearbook, 2007). He explained:

> 'Wisdomless' refers to not imposing our opinions on others. A university is a place where academic freedom should be respected. We should give space for different ideas to develop. 'Capabilityless' reminds us to stay humble even if we are capable. We should not think we are always more capable than others. Although I am the President of the University, I should not assume that I am literally the head of the university and my status is high at the top. Being 'actionless' can maximise others actions. As the President, I should only deal with the issues within the responsibilities of the President.
>
> (The President of the Project 211 University, 2006, 2)

The three mottoes express the President's support for academic freedom and his will to change the prevalent bureaucratic culture. He also revealed his intention to give more power to departments. He proposed that:

> Departments should be mainly responsible for the running of the university. They should be given more rights and take more responsibilities.
> (Uni211 Yearbook, 2008, 43)

Although the President made some steps forward in sharing power with departments, the power over making final decisions was still in the university and department leaders' hands. Several university committees were formed at the Project 211 University but, unlike those at the Project 985 University, they were not decision makers. They could make proposals but whether or not to accept their proposals was decided by the university leaders. For example, unlike their counterparts at the Project 985 University, departments' budget plans had to be authorised by the university. Similarly, teachers were involved in decision-making but they did not have the power to make a final decision either. At each department, departmental committees were established to make proposals, but final decisions had to be made by the department heads and the departmental CPC secretaries.

6.2 Internal Quality Assurance Mechanisms

The existing internal quality assurance mechanisms implemented at the Project 211 University before the UTE and the Quality Projects were class observation, the student evaluation of teachers and teacher appraisal. The details of these mechanisms will be provided in the following sections.

6.2.1 Class Observation and the Student Evaluation of Teachers

In 1993 the Project 211 University started 'triple evaluations', i.e. class observation by leaders and by inspectors and the student evaluation of teachers. In 2006, the Project 211 University began officially appointing around 12 university inspectors. The indicators used by the three internal quality assurance mechanisms are listed in Table 16 (the indicators adopted by all three internal quality assurance mechanisms

are in italics). The indicators in class observations by leaders and inspectors are almost identical. The only difference is that there is one more indicator in the evaluation forms for leaders, which is Indicator 11. The way in which the leaders and the inspectors rate teachers is the same. For each indicator, they choose A, B, C, D or E to rate. ‘A’ represents ‘highly agree’; ‘B’ represents ‘agree’; ‘C’ represents ‘uncertain’; ‘D’ represents ‘disagree’; and ‘E’ represents ‘highly disagree’. After rating all the indicators, leaders or inspectors have to give an overall rating of ‘excellent’, ‘good’, ‘acceptable’, ‘pass’ or ‘fail’.

For the student evaluation of teachers, students rate each indicator by A, B, C or D (‘A’ represents ‘excellent’; ‘B’ represents ‘good’; ‘C’ represents ‘acceptable’; ‘D’ represents ‘poor’). However, these letters, which indicate an ordinal rank, are then transferred into nominal numbers to give each teacher a total score in order to compare teachers. All indicators are classified into four categories: teaching attitude, teaching content, teaching methods and teaching outcomes. Each category is allocated a weight. Within each category, each indicator is allocated a weight. The highest weight goes to Indicator 2 – strictness with students and respect for students’ opinions. It is unclear how these weights are determined, but it is counterintuitive that this indicator is weighted more heavily than other indicators. For instance, Indicator 10 enables students to master the knowledge and the ability to cultivate students’ skills and should be weighted more heavily.

For students			**For leaders and inspectors**	
Teaching attitude	Indicator 1	Punctual	Indicator 1	Teach with passion
Weight 0.2	Weight 0.4	Submit to the discipline		Full of energy
		Role model	Indicator 2	Teach with infectious enthusiasm
	Indicator 2	Strict with students		Able to draw student interest
	Weight 0.6	Respect students' opinions	Indicator 3	*Easy to understand*
Teaching content	Indicator 3	Clear viewpoint		*Thought provoking*
Weight 0.35	Weight 0.3	Clear concepts	Indicator 4	Explain in a concise and precise manner
	Indicator 4	Carefully selected content		*Highlight key points*
	Weight 0.3	*Updated knowledge*		Clarity in logic
	Indicator 5	Relate to practice	Indicator 5	Familiar with the content
	Weight 0.2	Use appropriate examples		Able to organise the content into a coherent manner
	Indicator 6	*Highlight key points*	Indicator 6	Have substantive content
	Weight 0.2	Well-organised		Provide ample information
Teaching methods	Indicator 7	Student-tailored teaching	Indicator 7	*Reflecting the latest development in the field*
Weight 0.3	Weight 0.35	*Thought provoking*	Indicator 8	Able to provoke students' thinking, association and creativity
	Indicator 8	Vivid language	Indicator 9	*Able to arouse student interest*
	Weight 0.35	*Easy to understand*		Able to make class lively
	Indicator 9	Multiple teaching methods	Indicator 10	Effectively use multimedia
	Weight 0.3	Clear writing on the blackboard	Indicator 11 (for leaders only)	Classroom discipline
Teaching outcomes	Indicator 10	Enable students to master the knowledge		
Weight 0.15	Weight 0.5	Able to cultivate students'skills		
	Indicator 11	Able to provoke thinking		
	Weight 0.5	*Able to promote student interest*		
	Extra points	Able to use textbooks written in English and teach in English (excluding English courses) 20/15/10/0 points		

Table 16 Teaching evaluation forms at the Project 211 University

The Project 211 University's evaluation forms suffer from similar problems as those of the Project 985 University. The first problem is having multiple factors within one indicator. For example, Indicator 1 in the evaluation form comprises three factors: 'punctual', 'submit to the discipline' and 'role model'. This indicator becomes problematic when teachers fulfil just one or two factors of the same indicator – for example, when a teacher is punctual and can submit to the discipline but cannot act as a role model for students.

The second problem is the use of total scores in student evaluation. Although the Project 211 University applies weightings to different indictors, the weightings are not tested empirically. Even the logic of the weightings is questionable. Despite all the weightings, the indicator that is given the highest weighting actually pertains to the 'extra points'. On top of the total score (100), another 20 points are given to teachers who use textbooks written in English and teach in English. This makes the student evaluation of teachers biased towards teachers who are good at English. However, teachers who can speak good English are not necessarily good teachers.

6.2.2 Teacher Appraisal

In 2006, the Project 211 University published an updated teacher appraisal policy: *The Project 211 University Quantitative Performance-based Assessment for Staff in Teaching and Research Posts* (Project 211 University, 2006b). This policy specifies how to calculate the teaching and research workload.

The quantitative performance-based appraisal is considerably more in favour of research than teaching. In terms of teaching, one lecture roughly equals one point. For example, if a professor teaches 240 lectures in one academic year (8 lectures per

week), he or she will receive approximately 240 points from teaching for the whole year. In terms of research, any published article counts for at least 50 points. For an article in a prestigious journal, the points the article can earn are equal to 600 times the 'impact factor' of the journal. Say, if the impact factor were 3, the points would be 1800. The points that research projects can gain are calculated by the amount of grants. For example, a sciences research project awarded with a ¥300,000 CNY grant from the national government would count for 1500 points. Even a project without funding can be rewarded 50 points. This appraisal policy encourages teachers to conduct research, because research could generate more points than teaching.

If teachers would like to apply for promotion, they have to fulfil additional requirements according to *the Project 211 University Professional Qualification Review Criteria* (Project 211 University, 2006a). The requirements are also divided into teaching and research-related requirements. Teaching-related requirements are the number of lectures given by the applicants in the previous five years and teaching evaluation results. For professorship applicants, they should have given 240 lectures every academic year, for associate-professorship applicants it is 200 lectures and for lectureship applicants it is 10 lectures. In terms of teaching evaluation results, professorship applicants' average scores should be at least 70 points out of 100 in the student evaluation of teachers and associate professorship applicants should achieve at least 75 points for the previous five years. For lectureship applicants, their teaching should be considered satisfactory by department inspectors.

Research-related requirements comprise publications, research projects and awards. Publications are measured by the number of disciplinary research articles and books.

Eligible disciplinary research articles have to be published by journals recognised by the university. Articles that are published by journals outside this category are regarded as low quality articles, so they are not used for promotion. Textbooks are not eligible either. Research projects are measured by the number of grants received from government, other organisations and businesses. Awards granted by local and central government for either pedagogical or disciplinary research are recognised. The status of these grants and awards follows a hierarchy. Grants or awards made by the national government are regarded as the most prestigious, followed by those awarded by the provincial government and the city council. Grants given by other organisations and businesses have the lowest status.

The Project 211 University allows for exceptional promotion, which rewards teachers who have made exceptional research achievements, such as publishing in *Nature* or *Sciences* or receiving the most prestigious national research awards in China.

As illustrated, the Project 211 University's promotion criteria emphasise research performance, especially disciplinary research performance. In order to provide promotional opportunities for teachers who have made contributions to teaching, the Project 211 University adds an alternative type of professorship – teaching-path professorship. Applicants for this professorship should have given at least 360 lectures per year (12 lectures per week) on average for the previous five years. Alternatively, they have to have been granted Distinguished Teacher Awards.

Furthermore, applicants should have published a certain number of pedagogical research articles. These articles should be published by journals that are recognised by the university. A published textbook constitutes one article. Moreover, applicants

should have conducted research projects on teaching and learning funded by the national or the provincial government or they should have received national or provincial awards for designing courses or writing textbooks.

The criteria for teaching-path professorship suggest that this alternative professorship is designed for teachers who conduct good pedagogical research, but not for teachers who are only good at teaching. The criteria are almost the same as those of regular professorships. They all require applicants to have publications, research projects and awards. The difference between the regular and teaching-path professorships lies in whether they are disciplinary or pedagogical research achievements. Another difference is that teaching-path applicants have to have undertaken a much larger teaching workload.

6.3 The University's Reactions to the UTE and the Quality Project Awards

The Project 211 University's reactions to the UTE and the Quality Project will be divided into *agenda setting* and *programming*. The analysis will show that the objective of carrying out the UTE has changed from assuring the quality of undergraduate teaching to achieving a good result in the UTE during *agenda setting*, which subsequently guides the following *programming* and implementation of the UTE. This is exactly why teachers at the Project 211 University were irritated by the UTE. The Quality Project was also implemented at the Project 211 University but with less intensity.

6.3.1 Agenda Setting

The Project 211 University faced many problems at the same time. However, not all of them could receive attention and become top priorities. The President's speech at

the beginning of each year announced the top priorities for the year. Top priorities from 2006 to 2010 are summarised in Table 17.

Year	Top Priorities				
2006	University Anniversary Celebrations[31]	Undergraduate Education	Graduate Education	Project 211[32]	Chinese Language[33]
2007	Undergraduate Education	Graduate Education	Management[34]	University Culture[35]	
2008	Undergraduate Education	Research	Graduate Education	University Characteristics[36]	
2009	Human Resources[37]	Management	University Town[38]	Research	Graduate Education
	Project 211				
2010	Management	Undergraduate Education	Graduate Education	Research	Human Resources
	University Town	International Communication[39]			

Table 17 The Project 211 University's top priorities

Extracted from (Uni211 Yearbook, 2007; 2008; 2009; 2010; 2011)

Table 17 shows that undergraduate and graduate education constantly featured among the top priorities. Other frequently mentioned top priorities were research, Project 211, management and the university town. They became top priorities because of the *institutional rules* mentioned at the beginning of the chapter.

31 Priorities related to university anniversary celebrations are preparations made for the celebrations.

32 Priorities related to Project 211 are about demonstrating accountability in order to secure the following rounds of funding from Project 211.

33 Priorities related to Chinese language are about teaching Chinese to overseas students or overseas in order to promote Chinese language and culture.

34 Priorities related to management are university policy making and implementation, such as the teacher appraisal policy.

35 Priorities related to university culture are about promoting a scientific, people-oriented and harmonious university culture.

36 Priorities related to university characteristics are about clarifying the characteristics that distinguish the Project 211 University from the rest of Chinese universities.

37 Priorities related to human resources are about staff recruitment and development.

38 Priorities related to the university town are about the construction and the management of campuses in university towns and tackling problems caused by campuses at different locations.

39 Priorities related to international communication are about developing collaborations with overseas institutions.

It was because of the national policies that undergraduate education became a top priority. In his 2006 and 2007 speeches, the President mentioned undergraduate education when he talked about the UTE inspection in 2007. In his 2008 and 2010 speeches, the President mentioned undergraduate education when he talked about the Quality Project. This suggested that the UTE and the Quality Project had a strong influence on the Project 211 University's policies, which reflected the impact of *University Institutional Rule One*: to conform to government policies and requirements.

Project 211 became a top priority because of *University Institutional Rule Two*: to improve research performance in order to obtain more resources. As mentioned earlier, Project 211 provided a generous amount of money to the university, so of course it was a top priority for the university to ensure teachers generated enough research output to demonstrate their accountability to Project 211.

Graduate education and research were identified as top priorities because of the ambition to become a top university in China, which was *University Institutional Rule Four*. The President regarded having a large number of graduate students as an important indicator of university status. To be able to increase the quota for graduate student recruitment, the Project 211 University had to apply for 'graduate school' status from the Ministry of Education. Gaining this status was considered an honour. Only the most prestigious universities had authorised graduate schools. In order to apply to gain this authority, Project 211 had to improve its research performance.

The university town became a top priority because of *University Institutional Rule Two*, which tried to solve the problems caused by the dramatic increase in student numbers. Two of the newly established campuses were around a two-hour drive away from the main campus. How to manage multiple campuses and ensure the same standards across the four campuses was a challenge for the university.

Management was mentioned when the President was expressing his intention to share more power with departments. As mentioned earlier, this intention was partly realised. Another issue related to management was the change of teacher appraisal policy. Under the existing points-based teacher appraisal system, the President sensed that teachers were only willing to do the work that earned them points, so he wanted to change the appraisal to a post-based appraisal system (Uni211 Yearbook, 2010). In the post-based appraisal system, it would become a teacher's duty to undertake all teaching and research-related tasks.

6.3.2 Programming and Implementation

The *objective*, one of the elements of the *PAP*, of going through the UTE was mentioned in a speech by the President. The *objective* identified by the President was to achieve an 'excellent' result. The President said:

> The UTE is very crucial to our university's development, so we have to take it very seriously. If we cannot achieve 'excellence', the consequences will be serious.
>
> (Uni211 Yearbook, 2007, 45)

As the Project 211 University was endeavouring to establish itself as one of the most prestigious universities, the President considered that it would seriously damage its reputation if it could not achieve 'excellence' in the UTE.

The President allocated all existing *human resources* to the preparation for the UTE. The strategy developed by the President was:

> All staff should be mobilised to take part in UTE preparation; contracts should be signed in order to make sure that leaders and teachers are aware of their responsibilities; weaknesses should be identified and rectified within a certain timeframe in order to make sure they meet the UTE requirements.
> (Project 211 University, 2006g, [no pagination])

The requirement of signing the contracts reflected that the President took the preparation for the UTE very seriously. The allocation of all *human resources* to the UTE preparation made it an extra task added to the existing workload of staff and as such would divert their attention from their daily work to the UTE and cause disruption to day-to-day teaching.

In order to fulfil the *objective,* the Project 211 University made an *action plan*, composed of a 'Box Project', a 'Site Project' and a 'Good Attitude Project' (Project 211 University, 2007e). The term 'Box Project' was used to symbolise documentation and archiving. The archived documents were put into boxes, so it was called a 'Box Project'. In addition to all the certificates that they received, teachers had to submit their lesson plans, syllabi, PowerPoint slides, published or unpublished textbooks, published articles and the official confirmations of conducted research projects over the previous three years. The 'Site Project' was centred around the UTE evaluators' visit to the Project 211 University with an aim to impress the evaluators during their visit. The 'Good Attitude Project' required teachers and students to demonstrate good attitudes when the UTE evaluators visited the university.

The Project 211 University took the 'Box Project' very seriously. In various meetings university leaders repeatedly reminded department leaders and teachers to put their

effort into documentation (Project 211 University, 2006c; Project 211 University, 2006d; Project 211 University, 2007c; Project 211 University, 2007e; Project 211 University, 2007f). In order to ensure that the documents met the UTE indicators, documents were checked again and again, firstly by external evaluators invited in by the Project 211 University (Project 211 University, 2007j), secondly by university leaders (Project 211 University, 2007h), thirdly by departments (Project 211 University, 2007d) and then in the mock evaluation (Project 211 University, 2007g).

The 'Site Project' centred on impressing evaluators. A documentary and an exhibition were prepared (Project 211 University, 2007b; Project 211 University, 2007g). Furthermore, university leaders selected some departments as recommended places for evaluators to visit and they visited these departments beforehand to check their preparations (Project 211 University, 2007i). Several pre-evaluations were conducted to make sure the preparation was sufficient. In October 2006, the university appointed 20 inspection groups composed of leaders and teachers from the Project 211 University to evaluate departments' progress in preparing for the UTE (Project 211 University, 2006e). In June 2007, the Project 211 University invited some external evaluators who had had experience in evaluating other universities to evaluate its own progress (Project 211 University, 2007j). In September 2007, the State Council and the local government sent some evaluators to give the Project 211 University a mock evaluation (Project 211 University, 2007g). The effort that Project 211 invested in reflected its determination to achieve 'excellence'.

As 'good attitude' was an intangible concept, the 'Good Attitude Project' was mainly implemented through observing teachers' classes (Project 211 University, 2007g;

Project 211 University, 2007i). Moreover, demonstrating discipline was also considered a part of a good attitude. During the UTE preparation, teachers and students were asked to develop extra awareness in their behaviour (Project 211 University, 2006d).

Shortly before the UTE inspection, the Project 211 University submitted the university report. The university report was designed to show that the Project 211 University had met the requirements of *the 2004 UTE Scheme*. Its structure followed the structure of the evaluation scheme and every single indicator was addressed by the report. In the end, the Project 211 University, as it had so badly wanted, achieved its 'excellence' in the UTE.

The same motivation used by the Project 211 University to achieve 'excellence' in the UTE motivated the Project 211 University to participate in the Quality Project. The Project 211 University considered the Quality Project awards, especially national awards, the evidence of the high quality teaching offered by the Project 211 University and winning these awards was expected to would boost its reputation significantly (Project 211 University, 2006f).

The *objectives* for implementing the Quality Project were set by the *Project 211 University's Decisions on Implementing of the Quality Project, Furthering Teaching Reforms and Improving the Quality of Teaching* (Project 211 University, 2007a). The *objectives* were: establishing a certain number of university awards and winning a number of provincial and national awards. Similarly to the UTE, the *2007 Decisions*

allocated all *human resources* to the Quality Project, which asked all departments and staff to get involved in the Quality Project (Project 211 University, 2007a).

Although the Project 211 University had the *objectives* and allocated *human resources*, it did not have *action plans* for implementation. Similarly to the Project 985 University, it just disseminated national policy documents to departments. The different approaches to implementing the UTE and the Quality Project were attributed to the different nature of the two national policies. As an *authority tool*, the implementation of the UTE was mandatory. On the contrary, the Quality Project was an *incentive tool*, so the application was optional. Teachers did not have to submit applications, if they did not want to.

The awards received by teachers at the Project 211 University by the end of 2010 are summarised in Table 18.

	National	*Provincial*	*University*
Excellent Course Awards	8	37	164
Distinguished Teacher Awards	1	2	NA
Teaching Group Awards	1	NA	NA

Table 18 The total number of the Quality Project awards that the Project 211 University received by 2010

(Data extracted from the Project 211 University website and the Quality Project website)

6.4 Perceived Impact of the UTE

To evaluate the impact of the UTE, the accounts of the department heads are presented first, followed by teachers' accounts in order to compare and contrast their accounts. Finally, the impact of the UTE on teachers is summarised.

6.4.1 Department Leaders' Perspectives

The impact of the UTE perceived by the interviewed heads at the Project 211 University constituted *implementation actions* and *consequences*. The *implementation actions* mentioned by the heads were: producing documents and adopting a set of procedures to prepare for the UTE inspection. Head G mentioned the production of teaching-related documents:

> Each course should have a curriculum and an introduction in both Chinese and English etc., a whole set of documents…Even if we did not have these documents, I had to make up for the shortfall in documents in order to meet the university's requirements.
>
> (Head G, DepHum)

Head H (DepHum) introduced the new format of grading examination papers:

> Say, one section has five or ten questions. When I am grading the examination paper, I cannot just write a score that the student gained in this section without writing down the score of each question. I have to write the score of each question and the total score in this section…In the end, I have to sign it off.
>
> (Head H, DepHum)

Head H (DepHum) also mentioned a new procedure that teachers had to carry out:

> During the UTE preparation, every teacher was asked to carry the whole set of teaching documents with him or her to the classroom, including lesson plans, textbooks, teaching calendars, curricula etc., many documents…
>
> (Head H, DepHum)

These *implementation actions* were prescribed by university's *action plans*. The first two quotes illustrate the *implementation action* of the Box Project and the last quote demonstrates the *implementation action* of the Good Attitude Project. Making up for the discrepancies in documentation that were mentioned by Head G caused some controversies among teachers, which will be explained in the following section.

All interviewed heads at the Project 211 University mentioned 'standardisation' as a *consequence* of the UTE and they all took a positive view of it. However,

'standardisation' meant different things to different heads. Head K (DepSoSc) referred to the standardisation of documentation as 'standardisation':

> The overall impact of the UTE is positive. I do not care what other people say about the UTE. I only talk about its impact on my department, which is positive. The first and foremost impact is standardisation. Before the UTE, teachers did things their own way. Sometimes, they had even mislaid their teaching materials. After the UTE, every teacher had become aware of the documentation of lesson preparation, teaching and examinations and now considered this is part of his or her job.
>
> (Head K, DepSoSc)

Head H (DepHum) used the term 'standardisation' to mean standardisation of procedures:

> During the preparation for the UTE, we all had to do things in particular ways [such as grading examination papers according to the UTE requirement and carrying the whole set of teaching documents to the classroom]. After the UTE, teachers continued to follow the same procedures and they had become standardised procedures. I think it is desirable.
>
> (Head H, DepHum)

Heads K and H expressed positive attitudes towards the UTE. Their positive attitudes towards the UTE originated from their understanding of quality. Heads K and H viewed quality as 'perfection and consistency'. For them, having documents in the same format and the procedures for teaching were indicators of quality. Once documents were produced or procedures were adhered to, the indicators of quality were met and therefore quality was assured.

What Head G (DepHum) called 'standardisation' was halfway to establishing standards, which is illustrated by the quote below:

> In order to write the curriculum, we had to think it over and over and revise it again and again…We put a lot of effort into designing the aims and the content of each lesson. (Head G, DepHum)

Designing the aims and the content for each lesson was a step towards establishing the minimum standards of the course, i.e. the sub-concepts of 'quality as exceptional'

defined by Harvey and Green (1993). However, Harvey and Green (1993) maintain that whether an education provision has met the standards or not is determined by student learning outcomes. Therefore, it was essential to have a description of student learning outcome in order to establish the standards. Without the benchmark of student learning outcome, the standards would be unsound.

While acknowledging the 'standardisation' was a contribution made by the UTE, Head G (DepHum) and Head H (DepHum) did not think that the 'standardisation' would lead to the improvement of the quality of teaching. Head H (DepHum) confessed, 'However, these standardised procedures contribute little to the improvement of teaching quality'. Head G added:

> During the preparations for the UTE in 2007, all of us were very busy…The costs of the UTE inspection were too high and we paid too much for it. There were things that were done for formality only.
>
> (Head G, DepHum)

The 'costs' referred to by Head G were not only financial but also opportunity costs, i.e. the excessive amount of time and energy spent on the preparations for the UTE. Considering the financial and opportunity costs of the UTE, Head G did not think having the standardised curriculum was worthwhile. As the researcher argued above, at best it was an unfinished product of the minimum standards. These heads' quotes reflected that there were different definitions of quality co-existing in their minds. On the one hand, they defined quality as consistency and perfection or minimum standards; on the one hand, they defined it as transformation.

6.4.2 Teachers' Perspectives

In contrast to the interviewed heads' positive attitudes to the UTE, the majority of interviewed teachers expressed negative opinions on it, or more specifically irritation.

For example, Associate Professor E (DepHum) said 'the UTE was a distraction, a meaningless distraction, a hindrance, seriously!' Both its *implementation actions* and *consequences* irritated teachers. The *implementation actions* were the actions taken by teachers to implement the *action plans* designed by the Project 211 University. The first quote below illustrates the *implementation action* of the Box Project; the second and third quotes illustrate the *implementation actions* of the Site Project; the fourth quote illustrates the implementation *action* of the Good Attitude Project:

> The UTE was completely useless, just a waste of time for us, because it required us to reproduce syllabi, lesson plans and so on and so forth. We had had these documents but not in such great detail, because many years teaching experience had enabled us to memorise those details. However, the UTE requested us to put everything down in a great detail. Moreover, the UTE had a particular way of grading student examination papers. We had been grading student examination papers in our own way before the UTE. When the UTE came, we had to re-grade examination papers in that particular way. It was a great bother!
>
> (Lecturer H, DepSc)

> The UTE did not have any influence, except that the university forced me to attend a PowerPoint training. I already knew how to use PowerPoint…[but] I had to go. The university expected every teacher to go, regardless of whether he or she knew how to use it or not…
>
> (Professor E, DepHum)

> We were asked to carry five documents, such as the syllabus and the call roll with us to the classroom. This was just a formality and had nothing to do with the quality of teaching.
>
> (Associate Professor G, DepSc)

> At that time, we had to perform the stand-up greeting at the beginning of each class. (The stand-up greeting is often performed in primary or secondary schools.) All students stood up and said 'good morning' or 'good afternoon' to us and we replied and asked students to sit down. Wearing flip-flops to attend class was banned.
>
> (Associate Professor H, DepSoSc)

The first quote describes the troubles created by the Box Project. Teachers had to reproduce documents according to the new requirements introduced by the UTE, which were the same as those mentioned by teachers at the Project 985 University. The second and third quotes reflect Project 211 University's obsession with

presentations. In order to ensure that all slides were presentable, Project 211 University forced teachers to go to PowerPoint training; in order to show that teachers were well prepared, the University required teachers to bring all teaching-related documents with them. The fourth quote suggests the eagerness to impress the evaluators. To demonstrate good discipline, the University borrowed a secondary school practice.

Not only time-consuming, but also *the implementation action* of the Box Project, i.e. the reproduction of documents was regarded unethical. Professor G pointed out:

> In the past, we had not had those required teaching-related documents. During the preparation for the UTE, we created them. Creating what had never existed was a fraud, wasn't it?
>
> (Professor G, DepSoSc)

What Professor G said was bold, but it was true that faking that those documents had existed was a fraud. Instead of the Project 211 University, the UTE should take the responsibility for this fraudulent production of documents. The UTE checked documents for the preceding six years (later cut to the preceding three years) to see if they met the UTE requirements. Checking documents against the requirements made after the production of the documents lured universities to reproduce these documents.

Besides the negative impact, Associate Professor E (DepHum) reported a positive impact of the UTE. Associate Professor E (DepHum) said that during the preparation for the UTE many leaders went to check teachers' classes. Sometimes, they caught teachers arriving late to class or found that some students were absent from class. This could serve as a warning to remind teachers and students that they should arrive on time.

6.4.3 Summary

The impact of the UTE is summarised in Figure 12 and Figure 13. The impact of the UTE constituted both *implementation actions* and *consequences*. The interviewed heads and teachers mentioned the production of documents according to the UTE requirements as the *implementation actions*. Moreover, the interviewed teachers added more procedures, such as being forced to attend unnecessary training and performing secondary school greetings. The action of preparing documents was required by the *2004 UTE Scheme*, but the rest of these actions were added by the Project 211 University in order to impress the evaluators. These *implementation actions* illustrated eagerness to achieve the Project 211 University's objective of achieving 'excellence' in the UTE.

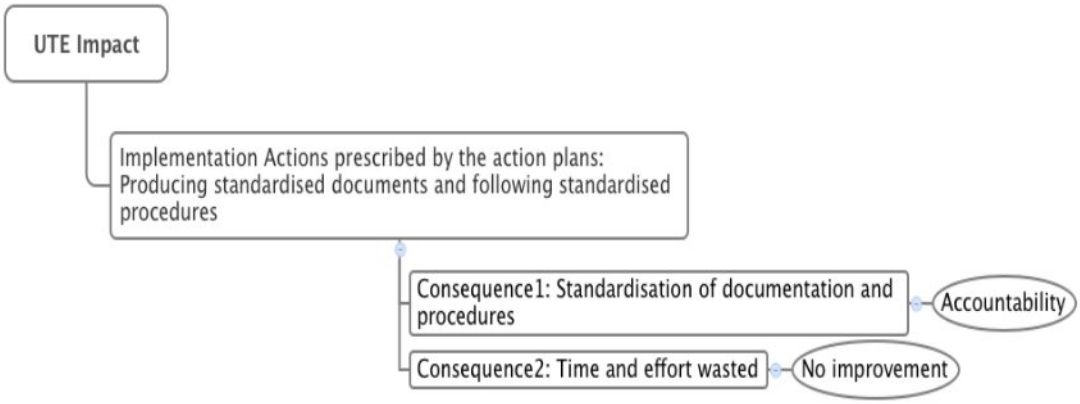


Figure 12 The impact of the UTE as perceived by head interviewees at the Project 211 University

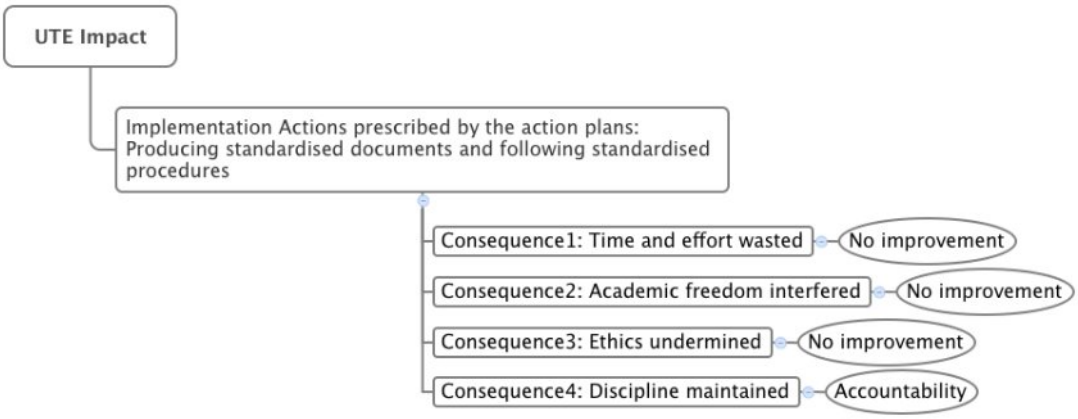


Figure 13 The impact of the UTE as perceived by teacher interviewees at the Project 211 University

Some of the interviewed heads and teachers agreed on a negative *consequence* – a waste of time and effort. It was a waste because they did not think it would contribute to the improvement of quality. Notwithstanding this, all interviewed heads maintained that it had positive *consequences*, i.e. the standardisation of documentation and procedures. These *consequences* were positive because they contributed to assurance of 'quality as perfection and consistency', in other words, accountability. The opposing idea held by some heads reflected a subscription to two different definitions of quality at the same time. In contrast, none of the interviewed teachers mentioned the standardisation of documentation or procedures as *consequences*.

Almost all interviewed teachers, on the other hand, mentioned negative *consequences*: the interference in academic freedom and the undermining of ethics. These were caused by the same *implementation actions* – the preparation of documents and the adoption of procedures according to the university's *action plans.* However, one positive *consequence* – maintenance of discipline – was mentioned by an interviewed teacher. This was the *consequence* of the *implementation action* of checking classes. This *consequence* was perceived positively because it contributed to the assurance of

accountability. While the *consequence* of checking classes can be perceived positively, the *consequence* of the preparation of documents was not perceived positively by any interviewed teachers. Why was it this case? Maintenance of discipline could serve as a necessary condition for improvement. However, preparing documents could serve neither as a necessary condition nor as a sufficient condition. In other words, good discipline, i.e. starting teaching on time, had a potential to improve teaching, whereas preparing documents in those particular ways did not have such potential.

Again, when comparing the *public problems* and the expected actions in Table 9 with the actions reported by teacher respondents, a similar conclusion can be drawn that the UTE was unable to fulfil its purposes.

6.5 Perceived Impact of the Quality Project Awards

The awards won by the three case study departments at the Project 211 University were listed in Table 19.

		DepHum	DepSc	DepSoSc
Excellent Course Awards	National	0	0	1
	Provincial	0	1	1
	University	4	13	6
Distinguished Teacher Awards	National	0	0	0
	Provincial	0	0	1
Teaching Group Awards	National	0	0	0

Table 19 The number of the Quality Project awards won by different departments at the Project 211 University

(Extracted from the Project 211 University website)

The Social Sciences Department won one national award and two provincial awards. The Sciences Department won one provincial award. The Humanities Department won neither national nor provincial awards.

6.5.1 Department Leaders' Perspectives

The heads interviewed at the Project 211 University reported an *implementation action* and *consequences* as the impact of the Quality Project. Head G (DepHum) confirmed that his department implemented the university *action plan* by encouraging teachers to apply for the Quality Project awards. The heads at the Social Sciences Department and the Sciences Department mentioned two *consequences*: awards-winners would motivate other teachers and can serve as models for other teachers. For instance, Head H (DepSc) said:

> No matter in which profession, people who demonstrate excellence can always motivate others. The same applies here. If a teacher received a university or a provincial Distinguished Teacher Award, other teachers would admire him or her and would try hard to match him or her.
>
> (Head H, DepSc)

Head K (DepSoSc) added that, as all materials of the award-winning courses were shared online, other teachers were able to access them at anytime and use them as models to improve their teaching.

6.5.2 Teachers' Perspectives

The impact of the Quality Project awards perceived by the interviewed teachers included an *implementation action* and *consequences*. Teachers' perceptions of the impact varied according to their involvement in the Quality Project.

Associate Professor H (DepSoSc), who filed an application together with her colleagues and won a National Excellent Course Award together, considered that the Quality Project awards offered both honorary and financial acknowledgement of teachers' effort spent on teaching. However, Lecturer G (DepHum), a non-award winner, pointed out that the majority of teachers' efforts in teaching remained unacknowledged, because only a small number of teachers could win awards.

Conversely, all research performance was recognised by the Project 211 University, because the university appraisal policy rewarded research publications by points.

Lecturer I (DepSoSc) was teaching a course that received a provincial Excellent Course Award. She appreciated the teaching materials provided by the award-winning course:

> Those teachers have put a lot of effort into making the application for the provincial Excellent Course Award. They have produced many materials. When I started teaching this course, I could use these materials directly.
>
> (Lecturer I, DepSoSc)

As a less-experienced teacher, Lecturer I was very pleased to have well-thought teaching materials ready for her to use, given that she might not able to produce materials with such quality.

However, Lecturer H (DepSc), who was also teaching a course that won a provincial Excellent Course Award, was not content with just having well-prepared teaching materials. She complained:

> … It (the course that won the provincial Excellent Course Award) is like an empty shell… They (other teachers who took the major responsibilities in application) do not give us any guidance in teaching. They do not tell us what standards we should match in teaching in order to live up to the expectations of the provincial Excellent Course. They do not make us aware of how other award-winning courses are taught.
>
> (Lecturer H, DepSc)

Lecturer H was expecting more guidance and help with day-to-day teaching. However, once the award was granted, work on improving the course stopped. The different perceptions of Lecturer I and Lecturer H were caused by different expectations. Lecturer I regarded the winning of Excellent Course Award as an end in itself, whereas Lecturer H regarded it as a beginning. This suggested that giving awards inspired individual teachers to produce well-considered teaching documents for the

application, but did not inspire teachers to work actively together to help each other to improve teaching sustainably.

Many interviewed teachers expressed a lack of interest in any of the official awards. The majority of those were professors. The following quote illustrates how irritated teachers could be:

> I don't think any kind of quality assurance mechanism works for me. For example, I never applied for Excellent Course Awards or Distinguished Teacher Awards. Whenever the university sends me these kinds of emails, I will just delete them as soon as I see the subject box.
>
> (Professor D, DepHum)

Lecturer J (DepSoSc) also expressed a lack of interest in government initiated quality assurance policies. He explained why he was so disdainful:

> All of these [quality assurance policies] in China are the product of a bureaucratic culture, including university policies. They are made by policy-makers just tapping their heads. The real purpose for policy-makers to initiate new policies is to use those new policies to demonstrate their achievements. Their decisions are not based on research. They do not care whether those policies are practical or implemented. They just make programmes with some specifications and require all universities to implement them. These programmes always end up with no real impact.
>
> (Lecturer J, DepSoSc)

Lecturer J's quote illustrates a strong distrust of government policies. He first pointed out that teachers were excluded from the policy-making process. He believed that policy-making was not evidence-based and many policies were made due to policy-makers' self-interest not consideration of the development of higher education. Because of the power of those policy-makers, they imposed their will on teachers and asked teachers to implement their policies. Hence, Lecturer J was not willing to be involved in those policy initiatives.

Another reason accounted for the lack of *implementation action* was insubstantial incentives. Professor D (DepHum) said:

> First, it is a big hassle. Second, it consumes a lot of time. In addition, I am not interested in this kind of thing. It is meaningless to me.
>
> (Professor D, DepHum)

In addition to the contempt mentioned earlier, this quote also indicates a lack of substantial incentives. The financial and honorary rewards offered by those awards did not match the amount of time and effort demanded for the applications. In addition, as a high-achieving professor, Professor D could earn more financial and honorary rewards from research, instead of teaching.

Professor F (DepSoSc) did not apply for the Quality Project awards because she did not think these awards were reliable. She said:

> Those who receive the Quality Project awards are not determined by teachers nor students. They are determined by officials.
>
> (Professor F, DepSoSc)

The 'officials' mentioned by Professor F were not government officials. They were leaders at universities. Who received awards was decided by committees, which were composed of leaders from various universities. Because teachers and students were excluded from selecting award-winners, Professor F did not think these awards were reliable.

6.5.3 Summary

Figure 14 and Figure 15 illustrate the impact of the Quality Project awards as perceived by the interviewed head and the teachers respectively. The impact of the Quality Project awards included: an *implementation action*, an *interaction* and *consequences*. The *implementation action* was filing the application. When a teacher

received an award, the feedback loop would be completed, so there was an *interaction*. If a teacher did not receive an award, there would be no *interaction*.

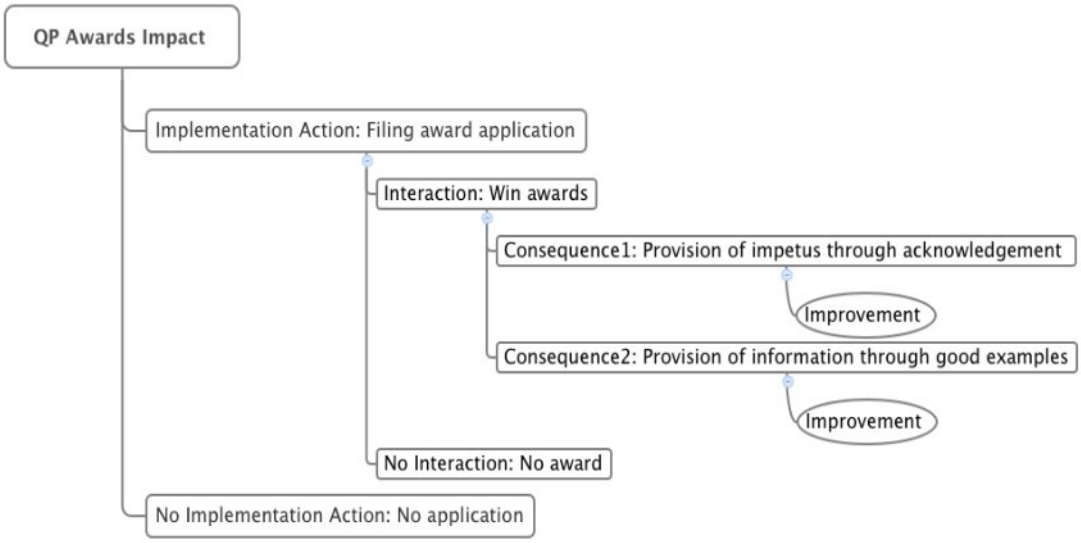


Figure 14 The impact of the Quality Project (QP) awards as perceived by head interviewees at the Project 211 University

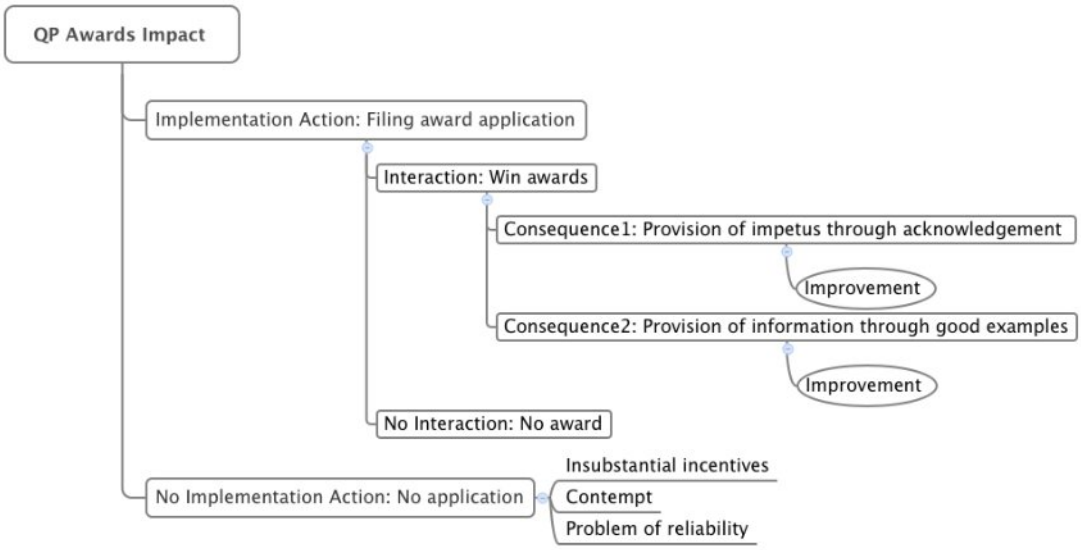


Figure 15 The impact of the Quality Project (QP) awards as perceived by teacher interviewees at the Project 211 University

Although the interviewed heads and teachers agreed on the *consequence* as being the provision of information, the *target groups* were different. The heads included

teachers who did not apply and who did not win awards in the *target groups*, stating that their interest would be drawn towards teaching and they would be motivated to learn from the information provided by the award winners. However, the teachers' accounts suggested that only those who were involved in award-winning courses were able to perceive this *consequence*. In addition, a teacher pointed out a potential side effect on non-award winners. Instead of being motivated to learn from the award winners, those non-award winners who had been interested and invested in teaching would be disappointed. However, this is an unavoidable condition for awards; the nature of awards determines that only a minority of teachers, i.e. the very best ones, receive the awards. Teachers who were the second best or average would not receive awards.

If there was no *implementation action*, there would *no consequence*. Teacher respondents explained the reasons for the lack of an *implementation action*, which were a lack of substantial incentives, contempt and the problem of reliability. The fundamental reason was the lack of substantial incentives. As an *incentive tool*, the Quality Project awards relied on incentives to motivate teachers. However, the incentives provided by the Quality Project awards were dwarfed by the incentives offered to research provided by other departments within the State Council.

In fact, the most frequently referenced reason was contempt with several interviewed teachers expressing their contempt for official awards. Why were these teachers contemptuous of the awards offered by the Ministry of Education but not of the rewards for research, such as research grants or the National Key Disciplines, offered by other departments within the State Council? The researcher would argue it was

likely to be that the rewards for research were too substantial to ignore. Without research grants the teachers would not be able to conduct research. Without the titles bestowed by the National Key Disciplines, the departments would not be able to attract high quality students to apply and not qualify for other rewards, such as the Project 211. In contrast, without the Quality Project awards, teachers were still able to teach and winning or losing would not affect teachers' chances of winning research-related rewards. Moreover, the contempt was generated as a protest against the lack of academic freedom. When implementing a policy using an *authority tool* such as the UTE, the Project 211 University added more requirements on top of what was required by the national policy and demanded teachers implement them using authority. Under the influence of a bureaucratic culture, the Project 211 University – unlike the Project 985 University, which shared decision-making power with teachers – did not take other measures to meet the demand for academic freedom either. Therefore, when having the freedom to make their own choices, teachers reacted passively and with contempt so as to show their dissatisfaction. The final reason was the problem of reliability caused by the selection of judges. As suggested, the reliability could be improved by involving more regular teachers (i.e. teachers did not hold managerial positions) and students in.

With regard to the expected actions listed in Table 9, none of the interviewed teachers mentioned them. This suggested that the Quality Project awards were unable to generate the expected actions and thereby unable to resolve those *public problems*.

6.6 Perceived Impact of Class Observation

The interviewed heads' perceptions of class observation are introduced. Their perceptions are then compared and contrasted with the interviewed teachers' to evaluate the impact of class observation.

6.6.1 Department Leaders' Perspectives

Interviews with the heads revealed that there were four kinds of class observations at the Project 211 University: class observations by university inspectors, department inspectors, university leaders and department leaders. University inspectors were retired teachers who were appointed by the university solely to observe teachers' classes across the university (Head I, DepSc). None of the departments appointed dedicated teachers as department inspectors. When the departments needed teachers to observe other teachers' classes, they asked some experienced teachers to take this role temporarily (Head G, DepHum; Head H, DepHum; Head K, DepSoSc). The departments mainly asked experienced teachers to observe new teachers and teachers who had applied for promotion, who were going to apply for teaching awards and who had received student complaints (Head H, DepHum; Head I, DepSc). University leaders only went to observe teachers' classes under special circumstances, such as 'Quality Month', a month dedicated to inspecting teaching (Head G, DepHum; Head J, DepSc). Department leaders selected one or two lessons to observe per term (Head G, DepHum; Head J, DepSc).

In terms of the impact of class observation by department inspectors, different heads interviewed had different opinions. Head H (DepHum) and Head K (DepSoSc) believed that, because inspectors would 'communicate' (Head K's word) or have a 'discussion' (Head H's word) with teachers, department inspectors provided teachers

feedback to help them to improve teaching. However, Head I (DepSc) disagreed with this. He believed that the problem of class observation by department inspectors is: 'Even when the inspector found problems with the teacher's teaching, the inspector was afraid that he or she might embarrass the teacher by pointing out the problems.' Hence, in Head I's opinion, the practice of class observation by department inspectors did not help teachers to improve.

Because of this concern, Head I (DepSc) was in favour of class observation by university inspectors. He commented:

> The university inspectors are retired, so they are not afraid of offending teachers. There is no conflict of interest between the inspectors and the teachers.
>
> (Head I, DepSc)

However, Head I was also aware of the potential bias of retired teachers' opinions. He acknowledged that 'their concept of teaching may be outdated'.

In Head H (DepHum) and Head K's (DepSoSc) opinions, the strength of using the university inspectors did not lie in providing honest opinions about teachers' teaching but in deterring teachers from violating some basic disciplinary norms, such as changing the time of lectures without submitting an application or giving a lecture without sufficient preparation.

In relation to class observation by leaders, the interviewed heads perceived class observation by department leaders as being more helpful. Head G (DepHum), Head H (DepHum) and Head J (DepSc), when talking about their experiences of observing teachers' classes as heads, emphasised that they would give teachers immediate feedback in order to help teachers improve their teaching. However, class observation

by university leaders was less helpful, because 'the university leaders mainly check discipline issues, while department leaders and inspectors are able to check teaching method and content and to assess whether teaching is able to facilitate learning' (Head J, DepSc).

In summary, the impact of class observation constitutes the *consequences*, i.e. providing information by pointing out the awareness of teachers' strengths and weaknesses in teachers' teaching and maintenance of discipline.

6.6.2 Teachers' Perspectives

Lecturer F (DepHum) and Associate Professor I (DepSoSc) agreed with Head H (DepHum) and Head K (DepSoSc) and they believed there was a need for class observation to maintain some basic discipline. Associate Professor I (DepSoSc) mentioned cases where teachers cancelled lectures before or after holidays. Lecturer F (DepHum) said that she was shocked when she heard cases of teachers being irresponsible.

Lecturer F (DepHum) and Associate Professor I (DepSoSc) suggested that class observation by university inspectors did not have any impact on them because these were not reliable. They said:

> The university inspectors do not have the expertise to understand my teaching.
> (Lecturer F, DepHum)
>
> Our university inspectors have a fault-finding attitude.
> (Associate Professor I, DepSoSc)

The first quote points out the problem of reliability caused by the lack of expertise. The second quote reveals a reliability issue caused by the focus on accountability.

Due to the focus on accountability, class observation by university inspectors created the image of fault-finding and teachers developed a distrust of university inspectors.

Some teachers interviewed at the Project 211 University acknowledged the benefit of class observation by department inspectors. Associate Professor F (DepSc) was observed by a department inspector who was an experienced teacher. The inspector suggested that Associate Professor F use more examples. Associate Professor F said ' I think this feedback is quite useful'.

Class observation by department leaders was considered helpful too. Lecturer F (DepHum) was observed by a department leader, who commented that Lecturer F did not have enough interaction with students during teaching. She said 'I suddenly realised this could be a problem…I think this is indeed a help'.

6.6.3 Summary

How the interviewed heads and teachers perceive the impact of class observation is summarised in Figure 16 and Figure 17. Teachers were absent from the *implementation action*. When teachers received the feedback, the *interaction* happened.

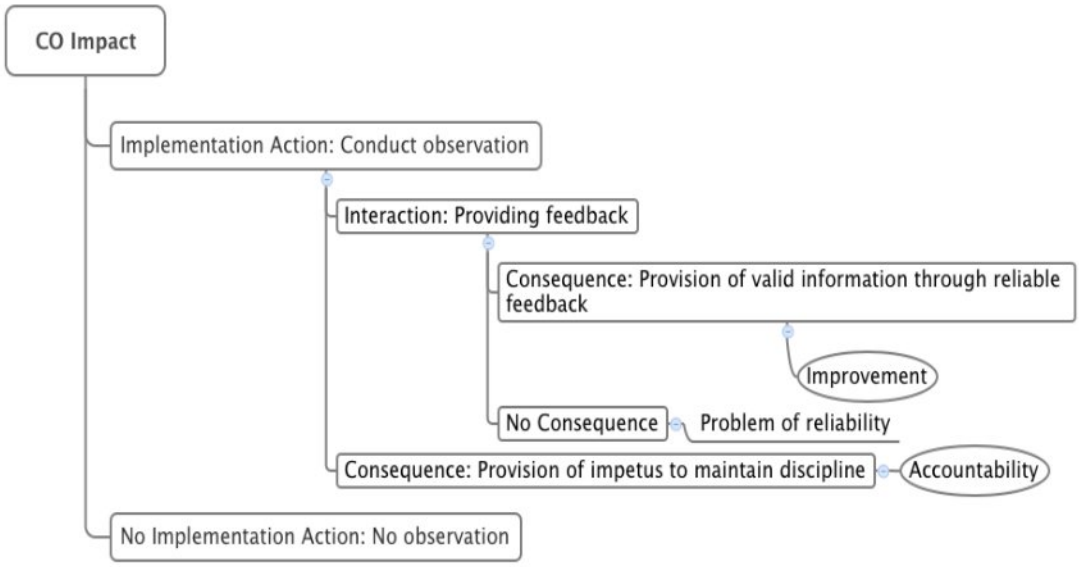


Figure 16 The impact of class observation (CO) as perceived by head interviewees at the Project 211 University

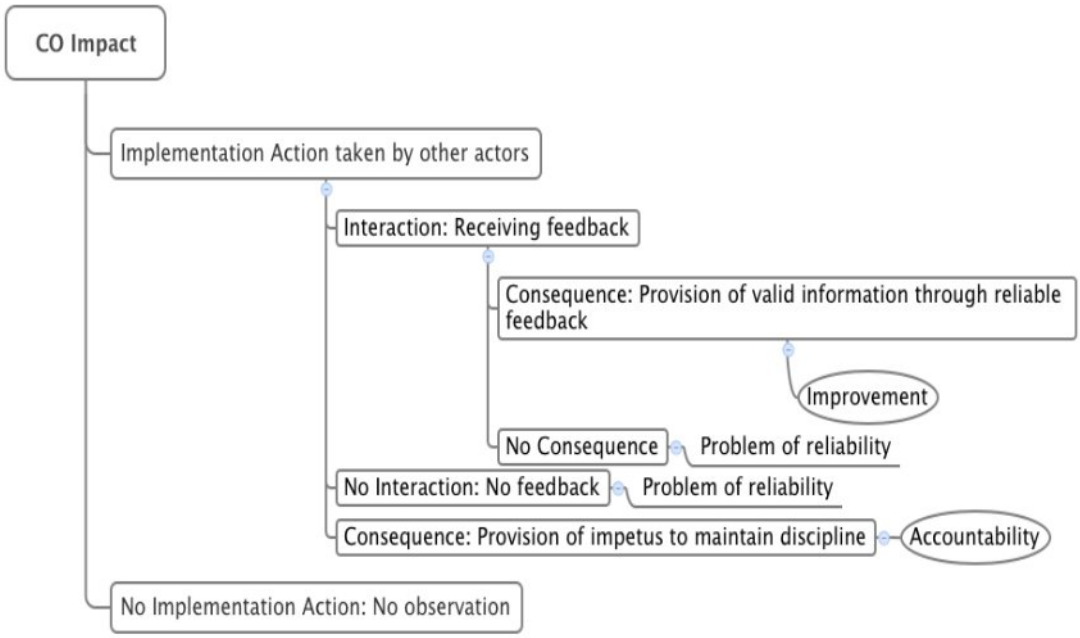


Figure 17 The impact of the class observation (CO) as perceived by teacher interviewees at the Project 211 University

The interviewed heads and teachers agreed on most of the impact of class observation and the reasons attributed to the lack of impact. However, there were two subtle differences. Firstly, the interviewed heads and teachers had different concerns regarding the reliability. A head was concerned that the department inspectors were

unable to honestly point out the problems. Teacher respondents were concern about university inspectors' lack subject knowledge. This was also the reason that prevented these teachers from asking university inspectors for feedback. Secondly, the teachers' accounts implied that they did not always acquire feedback, but the heads assumed that teachers normally would receive feedback.

6.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal

This section draws a comparison between how interviewed heads and teachers perceive the impact of the student evaluation of teachers and its relation to teacher appraisal. A summary will be provided at the end.

6.7.1 Department Leaders' Perspectives

The interviewed heads confirmed that the university policies had been implemented at their departments and they restated these policies. For example, Head I (DepSc) repeated that to be eligible for promotion of professorship, applicants' results in the student evaluation of teachers should be above 70 and the results of the student evaluation of teachers would affect the point-based calculation of teaching workload, which in turn would affect teachers' incomes.

The major concern of the heads was the reliability of student evaluation, which is illustrated by the quote below:

> We have found that those teachers who check student attendance or who are strict in grading student assignments and examinations receive lower scores in the student evaluation of teachers…This in turn contributes to another practice. Some teachers tend to maximise the scores that they could give to students in examinations in return for higher scores in the student evaluation of teachers. This results grade inflation in student examination results. After I took over the position [of the department head], I found that in a class of 40 students there were too many students scoring 90 and over (the full score was 100).

(Head H, DepHum)

The problems of the student evaluation of teachers perceived by the heads interviewed at the Project 211 University are consistent with those perceived by their counterparts at the Project 985 University. At the Project 211 University, the interviewed heads' confidence in the student evaluation of teachers varied according to the perceived degree of grading leniency. Head G (DepHum), Head I (DepSc) and Head K (DepSoSc) thought grading leniency happened in a small number of cases, so they had broad confidence whereas Head H (DepHum) and Head J (DepSc) believed that grading leniency happened quite often, so they considered student evaluation to not be a very reliable instrument.

The benefit of the student evaluation of teachers perceived by the heads interviewed was that it could collect student feedback for teachers who could adjust their teaching accordingly. However, the heads were not clear as to how many teachers checked the information, so they could only talk about their own personal experiences. Head G (DepHum), Head H (DepHum) and Head I (DepSc) said that they used information provided by student evaluation to improve their teaching but Head J (DepSc) confessed that he would not check his results because he did not think they were reliable.

6.7.2 Teachers' Perspectives

Although in theory the student evaluation of teachers should affect every single teacher at Project 211, the interviewed teachers' accounts revealed that in practice it could only influence a small number of teachers because the majority of teachers received very high scores. Associate Professor I (DepSoSc) said that when the student evaluation of teachers was conducted for the first time, her score was 90 (out of 100).

It was one of the highest scores at that time. However, now the situation had changed. She said:

> All of us have similar scores now. Before, those teachers who, like me, had a good reputation as teachers scored just about 90. Now, very few teachers cannot reach 90. I think this shows there are some problems with the current practice of the student evaluation of teachers.
>
> (Associate Professor I, DepSoSc)

Associate Professor F (DepSc) informed the researcher that the Dean of Academic Affairs of the Project 211 University validated this.

Differently from their counterparts at the Project 985 University who did not know their results of student evaluation, the teachers interviewed at the Project 211 University confirmed that they were able to access their results online, not only the total scores, but also the scores of each indicator. Associate Professor E (DepHum), Lecturer F (DepHum) and Associate Professor H (DepSoSc) considered these scores validated student feedback, which could help teachers to improve their teaching. Hence, the first *consequence* of student evaluation was providing information, i.e. student feedback.

The connection between student evaluation and appraisal provided teachers impetus to ensure high quality teaching and to please students. Maintaining high quality teaching could be achieved through carefully preparation or continuous improvement. Pleasing students could involve lowering academic standards or requirements for students. The former adopted an ethical approach but the latter was unethical. Associate Professor E (DepHum) confessed that student evaluation had the two *consequences* on him. Moreover, Associate Professor I (DepSoSc), Lecturer J

(DepSoSc) and Professor G (DepSoSc) confirmed the under-the-table deal mentioned by the heads, which is illustrated by the following quote:

> … some teachers explicitly said or implied to students that, if students gave them high scores, they would give them high scores in return. This is what I was told by some students.
>
> (Associate Professor I, DepSoSc)

As mentioned earlier, the majority of teachers received scores higher than 90, so this unethical deal could be offered by the minority of teachers who received very low scores in the student evaluation.

Most interviewed teachers claimed that student evaluation did not have any *consequence* on them: they either doubted the reliability and the validity of student evaluation or did not care about it. Most of them doubted the reliability of student evaluation and they suspected that students gave high scores to teachers who were not strict with them (Professor D, DepHum; Associate Professor F, DepSc; Associate Professor H, DepSoSc; Associate Professor I, DepSoSc; Lecturer I, DepSoSc).

Professor E (DepSc) pointed out the validity issue. She said:

> I think that teaching should not only be evaluated by the content but also the thinking behind it…The current student evaluation questionnaire only focused on knowledge but not on the thinking behind it.
>
> (Professor E, DepSc)

This point can be validated by the student evaluation form. As presented in Table 16, fourteen factors focus on the content or delivering knowledge and only two are about critical thinking. Furthermore, Lecturer H (DepSc) maintained that the lack of commentary feedback is another validity problem suffered by the student evaluation.

Moreover, Professor D (DepHum) and Professor G (DepSoSc) had contempt for the student evaluation of teachers due to the loss of dignity and the unethical conduct. Professor D encountered several cases of students chatting in his lectures. He was concerned that, if he confronted those students, they would be annoyed and give him very low scores in the student evaluation. He asked his colleagues for advice and many of his colleagues advised him to ignore the students and carry on teaching. However, Professor D could not stand just letting these students disturb his teaching, so he confronted them. Professor D commented:

> Because of this system (teachers are evaluated by students and the results affect teacher appraisal and subsequently salaries), many colleagues do not dare to displease students. I do not care much about my scores in the student evaluation of teachers and how much they affected my salary. I act according to my conscience. As long as I have done my work properly, I do not care how students rate me.
>
> (Professor D, DepHum)

Professor G was upset about the unethical conduct:

> There are teachers acting in some disgraceful ways to get high scores in the student evaluation of teachers. I strongly oppose the practice of teachers pleasing students by making them aware in advance of potential exam questions. I would never do this. I do not mind how the students rate me.
>
> (Professor G, DepSoSc)

Professor D pointed out the disadvantage of associating the student evaluation of teachers with teacher appraisal. This made student evaluation a high-stake assessment and some teachers sacrificed dignity and ethics for good results. Professor D and Professor G made a statement to defend their dignity and ethics by disdaining student evaluation.

6.7.3 Summary

The interviewed heads' and teachers' perceptions of the impact of student evaluation are summarised in Figure 18 and Figure 19. Because teachers were absent from the *implementation action*, the impact of student evaluation constitutes an *interaction* and

consequences. The *interaction* was teachers checking the results of their student evaluation. The interviewed heads did not articulate the impact made by the association between the teacher appraisal for promotion and student evaluation, just hinting at it being relevant in the background. However, the interviewed teachers' accounts clearly showed the difference. Associate professors and lecturers looking for promotions perceived the impact, despite only a minority of teachers could risk receiving scores lower than the minimum requirement in the student evaluation. The perceived impacts were *consequences* of providing impetus to ensure high quality teaching and to please students. The unethical behaviours of certain teachers were reported by a number of interviewed teachers.

Figure 18 The impact of student evaluation (SE) as perceived by head interviewees at the Project 211 University

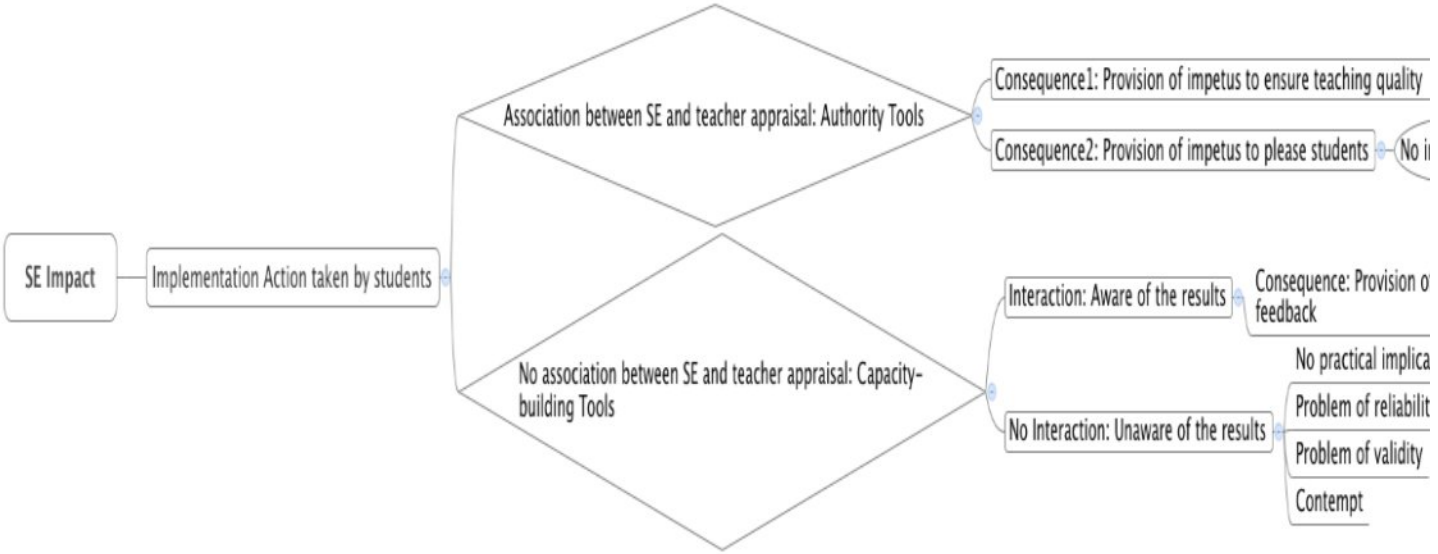


Figure 19 The impact of student evaluation (SE) as perceived by teacher interviewees at the Project 211 University

When teachers were able to ignore the connection between student evaluation and appraisal for promotion, whether student evaluation would have *consequence* or not was determined by if there was an *interaction*. When teachers checked the results of their student evaluation, they could use the provided information to help them to improve. This *consequence* – i.e. the provision of information – was shared among the interviewed heads and teachers. When teachers did not check their results, they did not intend to utilise the information provided by student evaluation. Therefore, there would be no *consequence*. The interviewed heads and teachers also agreed on one of the reasons accounting for the lack of *consequences*, which was the problem of reliability. The reliability was harmed by a prevalent belief that students gave low scores to teachers who were strict with them. In addition, the teachers added more reasons: no practical implications, the problem of validity and contempt.

6.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching

Based on interview data, this section answers several fundamental questions in relation to quality assurance: what are the problems of the existing quality assurance mechanisms perceived through heads' and teachers' perspectives? what drives heads and teachers to invest in teaching? what actually helps heads and teachers to improve teaching along the way? The interviewed heads' and teachers' accounts demonstrate that the existing quality assurance policies fail to address these issues. The answers to these questions will provide some insight into how to improve current quality assurance policies.

6.8.1 Problems with Current Quality Assurance Mechanisms

Both interviewed heads and teachers, mainly lecturers and associate professors, pointed out that the major problem of the current quality assurance mechanisms lay in

the university's teacher appraisal policy. It directed teachers' effort from teaching into research. The repeated complaint about the substantial disparity between the points earned from research and teaching is illustrated by the following quote:

> … the points earned from teaching three courses in one year was less than the points earned from publishing an article… This point system shows us that time should be spent on publishing articles or applying for research projects rather than on teaching.
>
> (Lecturer I, DepSoSc)

Although what Lecturer I described was not necessarily true for every single teacher, it represented a common perception shared by many interviewed teachers.

The substantial disparity in points generated for teaching and research was effectively demotivating teachers from teaching. Lecturer F's (DepHum) quote illustrated a very negative influence:

> You only have to look at the points generated from teaching to feel that teaching is considered worthless at this university, so I never looked at my points. However, I am still affected by this feeling. I really cannot look at my points. If I see them, I feel that they are an insult.
>
> (Lecturer F, DepHum)

It is completely understandable why Lecturer F thought it was an insult for teachers to spend time teaching. The few points received for teaching were in fact telling teachers that their effort expended on teaching was not worth very much. Because of this, even Head J (DepSc) asked 'if teachers have to spend the same amount of effort, why wouldn't they be doing other things (rather than teaching), thereby giving them a better return?'

In addition to the points-based regular appraisal, Head G (DepHum) pointed out that teacher appraisal for promotion also mainly relied on research. The only teaching-related criterion for promotion was a minimum score in the student evaluation of

teachers, which was 70 out of 100 for professor candidates and 75 for associate professor candidates. However, according to the Project 211 University statistics, in the 2006 academic year around 97% of teachers received scores higher than 80 and 99.94% of teachers received scores higher than 70 in the student evaluation of teachers (Project 211 University, 2007k).

Because of the emphasis on research, Head H (DepHum) sensed that many teachers were not willing to teach. He said:

> There are only two kinds of teachers willing to teach now. The first kind are teachers in their forties or fifties. Ten years ago this university did not focus on research, so many teachers only taught. Once they have reached their forties or fifties, it is really hard for them to learn to do research. Teaching is still their main source of income. The second kind are young teachers. They are interested in research and they are educated in the research environment. However, it will take a while before they establish themselves in research. Those young teachers are under an enormous amount of financial pressure, such as the pressure of buying their own flats, so they are willing to undertake more teaching. However, giving more lectures does not necessarily mean high quality. They are different issues. Now those who are not willing to teach are professors, due to the unreasonable university appraisal policy… They could spend less time and earn more money from doing other things than teaching.
>
> (Head H, DepHum)

Head H's quote reflects that teachers were driven by their personal motivation, which is often driven by a need to make a decent living in order to be able to afford to get themselves at least a flat. The university appraisal policy was basically encouraging teachers who were capable of doing research to do it.

Lecturer I (DepSoSc) provided a good quote to illustrate this problem and the influence of the university appraisal policy on young teachers. When she was asked whether she would be demotivated by the policy, Lecturer I answered:

> Of course, it would demotivate me from teaching. In my first year, because I was a new teacher, I liked communicating with students and performing well in the classroom. However, when I looked at some experienced colleagues, I

> found that they were not enthusiastic about teaching. They were doing the bare minimum. They already had pre-prepared slides, so they did not need to prepare lessons. They went into the classroom, showed the slides and then left. I could see that it did not matter whether they had taught well or not or whether they had communicated effectively with the students. It made no difference to a teacher's promotion.
>
> (Lecturer I, DepSoSc)

This extract illustrates the damage caused by the policy's overemphasis on research. First, some experienced teachers had become demotivated, so they gave teaching minimum input. This did not contribute to high quality teaching. Second, Lecturer I's response shows that young teachers who were interested in teaching were driven to concentrate on research. Moreover, the way that more experienced teachers behaved set a bad example for less experienced teachers and reassured them that minimum input was acceptable.

As mentioned above, the Project 211 University had a teaching-path professorship. Would this new form of professorship lend weight to teaching? The interview showed that this new professorship type was considered a secondary kind of professorship. For example, Associate Professor H (DepSoSc) commented:

> Teacher F is not good at doing research, but her teaching is excellent. She became the first professor promoted via this teaching path. It would have led her nowhere if she had had to apply for the normal professorship along the normal lines. She does not even have a doctoral degree.

On the one hand, this extract reflects Associate Professor H's perception that teachers who are not good at research apply for the teaching-path professorship. In fact, this perception was shared by Lecturer I (DepSoSc) and Associate Professor I (DepSoSc). On the other hand, this extract shows that Associate Professor H was glad that Teacher F, who had demonstrated excellence in teaching, was promoted through this new type of professorship. Although this new professorship served as an acknowledgement of teaching, it was also considered a last resort for teachers who

could not gain a regular professorship in the normal way. As a result, teachers, such as Lecturer I, did not want to pursue teaching-path professorship.

6.8.2 What Drives Teachers to Invest in Teaching?

The drivers perceived by the teachers interviewed at the Project 211 University echo the drivers described by their counterparts at the Project 985 University, which were passion for teaching, dedication and a sense of responsibility. Many interviewees are passionate about their subject and simply enjoy teaching it. Their passion for teaching can be represented by the quote from Associate Professor I (DepSoSc): 'I chose this profession and I always like it'. Similar to teachers at the Project 985 University, teachers at the Project 211 University believed that students' reactions were able to enhance their passion, as Lecturer H (DepSc) described as follows:

> What motivated me is students and more particularly the interest they expressed. I can see it from the interaction in the classroom. Without their interest, it would be difficult to make me willing to teach them.
>
> (Lecturer H, DepSc)

This quote shows that students' interest was sometimes very crucial for teachers. It could act as like a catalyst, making teachers deliver their best.

Many interviewees also regarded responsibility as a driver. Associate Professor H (DepSoSc), Professor F (DepSoSc) and Professor G (DepSoSc) expressed the same idea of meeting the expectations associated with the profession of teaching. Professor G said:

> The university offers me the job. Students choose my course. Students choose me to be their tutor. Of course, I cannot let the university and the students down. My conscience does not allow me to disappoint them.
>
> (Professor G, DepSoSc)

This quote illustrates that a sense of responsibility was associated with obligation to others, such as the university or students.

Many interviewees mentioned that students promoted their sense of responsibility. Lecturer G (DepHum) was one of them. She recalled a case that illustrates how big an impact students could have on teachers. In one lecture, she was teaching something very abstract. There was a female student, who was trying her best to open her eyes wide but could not keep her eyes from closing. Lecturer G saw this and she felt very guilty. She thought that it was her fault for making the lecture so boring. She said:

> Throughout my career the pain of students going through an unsuccessful lecture can pass directly to me. This makes me reflect and improve in order to keep this to the minimum.
>
> (Lecturer G, DepHum)

Lecturer G might be an extreme case. She was very sensitive to students' reactions and really cared about how students felt. However, she was certainly not the only one, who felt for their students. Associate Professor E (DepHum), Associate Professor F (DepSc), Associate Professor I (DepSoSc) and Professor G (DepSoSc) had similar feelings.

There was only one teacher, Professor D, who mentioned dedication. When he was asked what motivated him to improve teaching, he answered 'I am not sure why. I just want to do my best. No matter teaching or research, I just want to perform my best'.

Head G (DepHum) and Head H (DepHum) mentioned some motivations that other teachers did not. Head G mentioned the student evaluation of teachers. Head H mentioned his image among students and colleagues.

6.8.3 What Helps Teachers to Improve Teaching?

Agreeing with their counterparts at the Project 985 University, interviewed teachers at the Project 211 University perceived that research, students and colleagues helped them to improve teaching. Regardless of their academic titles, many interviewees believed that research would facilitate teaching. Professor D (DepHum), Associate Professor H (DepSoSc), Professor F (DepSoSc) and Head K (DepSoSc) thought that research would enrich teaching and Lecturer H (DepSc) and Professor G (DepSoSc) considered that having a strong research background was an indispensable quality of teachers.

Associate Professor F (DepSc) and Associate Professor I (DepSc) mentioned help from students. They acknowledged that students' questions or discussions with students provoked their thoughts, therefore helping them to improve their teaching.

Lecturer G (DepSc) and Associate Lecturer E (DepSoSc) emphasised help offered by their colleagues. They described their experiences:

> There was a great teacher at our department. He had 30 years' experience. He took every single lecture very seriously. My habit of writing in a teaching dairy to reflect my teaching is learnt from him. He told me that, when he was in his 20s, at the time he just started his career, his teaching diary had been as heavy as a book.
>
> (Lecturer G, DepHum)

> When I was still a new teacher, my slides were full of words. Students found them difficult to read and even I felt the same. After I observed a teacher who had 20 years' experience, I was inspired, which made me realise that the slides could become concise.
>
> (Associate Professor E, DepHum)

One important factor made their experiences beneficial was teachers' initiative. In other words, the two interviewees took the initiative to learn from their colleagues, which were in contrast to the official mentoring programme.

There was a mentoring programme at the Project 211 University. The university made it a university policy to allocate a mentor to each new teacher. However, the policy was not working effectively, as illustrated by Associate Professor E's experience.

> We (my mentor and I) never contacted each other. Just I know his name and he knows my name. He never asked me to go to his lectures or gave me some advice on my research.
>
> (Associate Professor E, DepHum)

The quote reflects a lack of support to the mentoring programme at the Project 211 University. The mentoring at the Project 211 University is different from its counterpart at the Project 985 University in several ways. At the Project 211 University, new teachers have to teach while being mentored; thus mentoring can be regarded as extra work. Furthermore, mentoring is introduced by university policy rather than initiated by departments. The failure of the mentoring system at the Project 211 University indicates that a policy introduced by a top-down approach is inclined to suffer from a lack of support on the ground, whereas the success of the mentoring arrangement at the two departments at the Project 985 University implies that a policy initiated by the bottom-up approach is likely to be supported by the departmental culture and teachers.

6.9 Chapter Conclusions

There were two features particular to the Project 211 University: a point-based appraisal system and the tension between the lack of respect for and the demand for academic freedom. These two features affected the perceived impact of the quality assurance policies on teachers. The point-based appraisal system reflected the definition of quality as 'perfection and consistency'. Teachers' work was transformed into points and teachers were rewarded based on the points accrued. The point-based appraisal system was designed to improve effectiveness, especially the effectiveness

of research. However, it posed a risk to 'quality as transformation', especially the quality of teaching, for two main reasons: the difficulty in measuring the quality of teaching and the emphasis on research.

When quality was measured by quantitative indicators, quality would be divided into teaching- and research-related performance. The Project 211 University used the number of publications and awards and the value of research grants to measure research performance and the amount of teaching and the results of student evaluation to measure teaching performance. However, the student evaluation of teachers conducted by the Project 211 University was not considered a valid and reliable instrument, as indicated by the interviewed heads' and the teachers' accounts above. Therefore, teaching performance was effectively measured solely by the amount of teaching conducted. This might be able to assure the quantity of teaching but not the quality of teaching. However, even the quantity of teaching cannot be assured at the Project 211 University.

The Project 211 University weighted research over teaching because of *University Institutional Rule Two* (to work hard to improve research performance in order to obtain more resources) and *University Institutional Rule Four* (to work hard to improve research performance in order to become a top university in China). Therefore, the Project 211 University did not set a minimum teaching requirement for regular appraisal and, on the contrary, allocated generous points to research performance. These points were then applied in teachers' salaries. This was a key difference between the Project 211 University and the Project 985 University. Admittedly, the Project 985 University's *University Institutional Rule Two* (to work

hard to improve research performance in order to obtain resources) also encouraged it to improve its research performance. However, the pressure was greater on the Project 211 University due to the Project 211 University's need to improve its reputation through research. Moreover, unlike the Project 985 University, the Project 211 University did not have *consensus* and *time* to reconcile the tension between teaching and research.

In order to balance teaching and research, the Project 211 University offered a teaching-path professorship. However, this professorship was actually balancing disciplinary and pedagogical research, instead of balancing teaching and research. Furthermore, the teaching-path professorship was regarded as secondary because of the lower status attached to teaching in general.

The second feature was the tension between the lack of respect for and the demand for academic freedom. There was a bureaucratic culture at the Project 211 University. The university leadership felt that it could impose its will on teachers because of the concentration of power in the hands of the university leaders. On the other hand, teachers at the Project 211 University had a need for academic freedom. They had relatively high esteem; after all, it was because of their competence that the Project 211 University had earned its reputation. This tension resulted in contempt for quality assurance policies that were not compulsory. In addition, the high esteem also created a tendency to protest unethical behaviour with contempt.

7 City University

The university context and the existing internal quality assurance mechanisms at City University will be introduced first, followed by City University's responses to the UTE and the Quality Project. Department leaders and teachers' perceptions of the impact of these external and internal quality assurance policies will be described, to assess whether these policies have contributed to the improvement of the quality of teaching. Teachers' reflections on the problems associated with these quality assurance policies and factors that contributed to their improvement of teaching will be presented at the end of this chapter.

7.1 The University Context

University institutional rules or the university context are formed under the influence of the *national institutional rules* and the university leadership, which will be described below.

7.1.1 University Institutional Rules Inherited from the National Institutional Rules

City University's *institutional rules* inherited from the *national institutional rules* are: *University Institutional Rule One*: to conform to government policies and requirements; and *University Institutional Rule Two*: to try to solve the problems caused by the dramatic increase of students.

City University has to conform to the policies of the City Council, the provincial government and the State Council, because they have control over the *human resources*, the *legal resources* and the *financial resources.* The President and the CPC Secretary of the University are appointed by the provincial government. The quotas

for student and staff recruitment are allocated by the provincial government based on the total quotas of the province set by the State Council. The City Council imposes further quotas for student recruitment, which are: at least 60% of undergraduates should be from the city where the university is located and in total 90% of undergraduates should come from the same province (Anonymous, 2009). The current City University was created as a result of the City Council's decision to merge all universities and colleges[40] that were regulated by it around ten years ago. The allocation of quotas and the creation of City University are applications of *legal resources*, the power to establish rules for City University. Appointment of university leaders reflects the control over *human resources*.

The financial dependence on the City Council indicates the City Council's possession of *financial resources*. Although City University's income and expenditure are not published, the financing of the new campus may indicate City University's dependence. At the beginning, given that City University did not have enough money to fund the construction of the new campus, the City Council endorsed the university's application for a bank loan and agreed to pay the interest on that loan. One year before the move to the new campus, the President persuaded the City Council to fund the move, so the City Council gave the university ¥200,000,000 CNY (£20,000,000 GBP) (UniCity Yearbook, 2004). Three years later, City University still did not have enough money to pay back the loan. City University turned to the City Council again, so the City Council contributed ¥1,130,000,000 CNY (£113,000,000 GBP), half of the total loan of ¥2,130,000,000 CNY (£213,000,000 GBP) (UniCity Yearbook, 2008).

40 Colleges are higher education institutions providing three-year undergraduate programmes that do not lead to a bachelor's degree.

Although the national policy to increase participation in higher education was launched in 1998, City University did not start increasing the number of undergraduate students until 2005, due to its physical capacity. The university's small campuses in the city centre were unable to meet the demand in terms of the growing number of students. As a result, City University decided to build a bigger campus in the university town, the construction of which was completed in 2005. As shown in Table 20, the number of undergraduates more than doubled in 2005. This dramatic increase is mainly attributed to the increase in the number of undergraduates who studied three-year diploma programmes[41]. Since 2007, City University reduced its enrolment of this type of student and focused on enrolling undergraduates for the bachelor's degree programmes. As summarised in Table 20, the number of undergraduates decreased in 2007 and then started to increase steadily thereafter.

	Number of full-time graduate students	Number of undergraduate students	Number of full-time teachers	Student–teacher ratio
2004	337	17,990	1,156	16:1
2005	503	44,110	1,164	38:1
2006	668	40,904	1,316	32:1
2007	819	23,596	1,315	19:1
2008	999	21,661	1,329	17:1
2009	1,372	29,406	1,328	23:1
2010	1,775	30,591	1,369	24:1

Table 20 The numbers of undergraduate and graduate students and teachers and student–teacher ratio at City University

Extracted from (UniCity Yearbook, 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010)

On the other hand, the problem of a shortage of teachers persisted. The number of teachers remained at a low level from 2004 to 2010. During the peak in student numbers in 2005, the student–teacher ratio reached an astonishing 38:1. Despite some

41 Undergraduates who study three-year diploma programmes can only receive diplomas but not bachelor's degrees; only those who study four-year degree programmes can receive bachelor's degrees.

improvement between 2006 and 2008, the student–teacher ratio began to climb again from 2009.

7.1.2 University Institutional Rules Established by the University Leadership

University Institutional Rule Three, to improve research performance in order to boost the university's reputation, was set by the university's leadership. Before the merger, all the universities and colleges that were regulated by the City Council were teaching-intensive institutions. After the merger, City University decided to establish itself as a university that was both strong in teaching and in research. This goal was stated in the *2003-2006 Plan for the Development of City University*:

> We should establish the idea that teaching is the foundation of our university and research is the key to making our university stronger. We should recognise that teaching is always our focus. We should consider teaching and research equally important. We should improve the quality of teaching and research and enable each to facilitate the development of the other.
>
> (UniCity Yearbook, 2003, 20)

The *2003-2006 Plan* moved City University's focus from teaching only to both teaching and research. The rationale behind this adjustment was that research would facilitate the development of teaching and vice versa.

Several years after the merger, university leaders still considered that City University was lagging behind its counterparts in terms of its research performance and reputation, so the CPC Secretary urged teachers to improve their research performance in order to obtain doctorate degree-awarding powers and the title of Provincial Key Discipline and to increase the size of the graduate student population (UniCity Yearbook, 2004).

In fact, many teachers felt that teaching was no longer the focus of the university. The President was compelled to deliver another speech to restate the relationship between teaching and research:

> Some colleagues thought that teaching was no longer the focus of the university. The university was only focused on research. In fact, teaching is still our focus. However, the university has to take care of research too. The three functions of the university – cultivating talents, conducting research and serving our society – determine that, as long as it is a university, it has to address research…Having both teaching and research as our focus means that the requirements for teachers are higher. Teachers should be more capable and integrate teaching with research.
>
> (UniCity Yearbook, 2004, 60)

The President intended to reemphasise the equal importance of teaching and research but he failed to acknowledge that teachers' perceptions were the result of the university's policies, and he failed to consider adjusting those policies to balance teaching and research. Rather than considering the effect of the university policy, the President assumed that it was the teachers' responsibility to balance teaching and research.

Unlike the Presidents of the Project 985 University and the Project 211 University, the President of City University did not intend to share more power with departments and teachers, nor to promote academic freedom. The President wanted to retain centralised control over the university. This feature is not only reflected in interviewed heads' and teachers' accounts, which will be presented later, but also in the new campus in the university town. Professor L (DepSoSc) commented, 'The new campus looks like a factory, doesn't it?' Unlike the Project 985 University and the Project 211 University's campuses, City University's new campus is not organised by departments, but by functions. Two big office buildings, which host all teachers' and administrative offices, are at one side of the campus. Buildings containing classrooms

are in one area of the campus and buildings containing laboratories are in another area. All buildings and facilities are managed by the university instead of by departments. Due to a lack of physical presence, the traditional atmosphere of departments is absent at City University.

7.2 Internal Quality Assurance Mechanisms

The internal quality assurance mechanisms implemented at City University before the UTE and the Quality Projects were class observation, the student evaluation of teachers and teacher appraisal. Details of these mechanisms will be provided in this section.

7.2.1 Class Observation and the Student Evaluation of Teachers

Class observation by university inspectors, which was launched in 2002, is the first internal quality assurance mechanism that was established by City University. The online student evaluation of teachers and class observation by leaders was initiated in 2004 and 2006, respectively.

The evaluation form for university inspectors is unstructured and that for university leaders is semi-structured. University inspectors are provided with a blank section in which to write their comments and another section in which to provide an overall rating of either ‘excellent’, ‘good’, ‘fair’, ‘acceptable’ or ‘poor’. University leaders have to rate as either ‘excellent’, ‘good’, ‘fair’, ‘acceptable’ or ‘poor’ the teaching attitude, teaching goals, teaching content, teaching methods and teaching outcomes.

The evaluation forms for students and department leaders are more structured. Each evaluation form consists of ten indicators. Each indicator is rated ‘A’, ‘B’, ‘C’, ‘D’ and ‘E’. ‘A’ equals to 10 points; ‘B’ 8.5 points, ‘C’ 7 points, ‘D’ 6 points and ‘E’ 4

points. Out of the full score of 100 points, if a teacher receives more than 85 points, the evaluation result for the teacher will be ‘excellent’; between 75 to 85, it will be ‘good’; between 65 to 75, it will be ‘fair’; between 60 to 65, it will be ‘acceptable’; below 60, it will be ‘poor’. It is compulsory for students to complete the questionnaires. Failure to complete the questionnaires leads to invalidation of students’ exam scores. The indicators for the student evaluation of teachers and class observation by department leaders are summarised in Table 21.

	For students		**For department leaders**
Indicator 1	*Rigorous scholarship*	Indicator 1	*Role model*
	Role model		Not only teach students knowledge, but also cultivate well-rounded individuals
	Caring but strict with students	Indicator 2	*Rigorous scholarship*
Indicator 2	*Prepare lessons well*		*Prepare lessons well*
	Clarity in logic		Familiar with the content
	Precise expression		Extensive knowledge
Indicator 3	Grade student assignments carefully	Indicator 3	Clear concepts
	Answer students' questions promptly and regularly		Correct viewpoints
Indicator 4	*Provide substantive content*		*Clarity in logic*
	Highlight difficult parts or key points	Indicator 4	*Provide substantive content*
Indicator 5	*Student-tailored teaching*		*Provide ample information*
	Relating to practice		*Reflect the latest development in the field*
	Reflect the latest development in the field	Indicator 5	*Student-tailored teaching*
Indicator 6	*Multiple teaching methods*		*Highlight key points*
	Pay attention to communication with students		*Relating to practice*
Indicator 7	*Use appropriate teaching methods*		*Use appropriate teaching methods*
	Provide ample information	Indicator 6	*Multiple teaching methods*
Indicator 8	Promote students' interests		Thought provoking
	Promote active learning		*Promote active learning*
Indicator 9	*Cultivate students' self-learning abilities*	Indicator 7	*Communication between the teacher and students*
	Cultivate students' problem-solving skills		Lively class atmosphere
Indicator 10	*Cultivate innovation*	Indicator 8	*Precise expression*
	Cultivate students' abilities to put theory into practice		Fluent expression
		Indicator 9	Capable of managing and controlling classroom
			Good time management
		Indicator 10	*Cultivate students' self-learning abilities*
			Cultivate students' abilities to analysis problems
			Cultivate students' problem-solving skills
			Cultivate innovation

Table 21 Teaching evaluation forms at City University

Factors highlighted in italics indicate that the same indicator appeared in both the student evaluation of teachers and class observation by department leaders. Suffering from the same problem as applied in the case of the Project 985 University and the Project 211 University's evaluation forms, the indicators listed in Table 21 also incorporate multiple factors within one indicator. Take Indicator One for students, for example: a teacher who can demonstrate rigorous scholarship is not necessarily a role model or caring but strict with students. Furthermore, the majority of indicators overlap, which leads to doubts regarding the validity of the student evaluation – whether these indicators are able to reflect students' perceptions of high quality teaching. Finally, the use of total scores fails to reflect the multi-dimensionality of teaching quality and makes the interpretation of evaluation results difficulty.

7.2.2 Teacher Appraisal

According to City University*'s Allowance Allocation Scheme* (City University, 2002), teachers' salaries are determined solely by the teaching workload that teachers have to complete. The standard teaching workload is 320 units per year. One lecture equates to one unit. Teachers are expected to teach at least eight lectures per week, or equivalent. Supervising placements, experiments and dissertations can be calculated by the units of teaching workload as well. If teachers cannot complete their teaching workload or teachers have taught more than was required, they are paid for the actual amount of teaching that they have conducted.

Although teachers' salaries are not related to research, various research achievements have a significant potential to generate substantial income for teachers, in the form of bonuses according to *City University Research Rewarding Regulations* (City University, 2005). City University offers generous financial rewards for research

achievements, including awards, publication and patents etc. The most substantial reward is ¥500,000 CNY (£50,000 GBP) for winning a first prize in national awards (for both disciplinary and pedagogical research), followed by a ¥200,000 CNY (£20,000 GBP) reward for one publication in *Nature* or *Science*.

In terms of rewards for teaching achievements, City University only offers rewards to winners of Excellent Course Awards. The highest reward is ¥30,000 CNY (£3000 GBP) for winners of national Excellent Course Awards. The gap between the financial values of the rewards indicates the different values of teaching and research at City University.

In relation to promotion, City University does not have the right to grant professorships. Applications for professorships have to be approved by the provincial government. Hence, the promotion requirements made by the provincial government are adopted by City University. According to *The Provincial Requirements for the Application of Professorships and Associate Professorships* (Provincial Personnel Department, 1999), teachers have to fulfil two teaching-related and a number of research-related requirements. The teaching-related requirements relate to the amount of teaching and the home institutions' judgements regarding the applicants' quality of teaching over the previous two or three years. The research-related requirements are the amount of publications, awards and research grants. The minimum requirement for teaching is that applicants should have taught 59 lectures per year; the minimum requirement for research is that applicants should have published five or six disciplinary research articles or books. Disciplinary research books can be substituted by textbooks. Once the minimum requirements have been met, the fulfilment of

teaching-related requirements can be substituted by the fulfilment of research-related requirements. Again, the message delivered by the promotion requirements is that research achievements play a decisive role in promotion whereas teaching is less important.

7.3 The University's Reactions to the UTE and the Quality Project Awards

City University's reactions to the UTE and the Quality Project will be divided into *agenda setting* and *programming*. The analysis will show that the objective of carrying out the UTE changed from assuring the quality of undergraduate teaching to achieving a good result in the UTE during *agenda setting*, which subsequently guided the following *programming* and implementation of the UTE. The Quality Project is also implemented at City University but with less intensity.

7.3.1 Agenda Setting

City University announces its top priorities each year. A list of top priorities from 2003 to 2010 extracted from City University Yearbooks is presented in Table 22. The most frequently mentioned top priorities are: research, graduate education, staff, management, university town and agendas related to the CPC.

Year	Top priorities				
2003	CPC[42]	UTE[43]	Research	Graduate Education	University Town[44]
	Salary[45]	Student[46]	Staff[47]	International Communication[48]	University Culture[49]
2004	University Town	UTE	Research	Graduate Education	Staff
	Management[50]				
2005	CPC	University Culture	UTE	Research	Management
	Staff	University Town			
2006	CPC	University Culture	UTE	University Town	Research
	Graduate Education	Management			
2007	Quality Project	Research	Management	University Culture	
2008	CPC	Quality Project	Research	Management	Staff
	International Communication	University Culture	University Town		
2009	Quality Project	Research	Graduate Education	Management	Staff
	International Communication	University Town	CPC		
2010	Reputation[51]	Management	Research	University Anniversary Celebrations[52]	

Table 22 City University's top priorities

Extracted from (UniCity Yearbook, 2003; 2004; 2005; 2006; 2007; 2008; 2009; 2010)

42 Priorities related to the CPC are about holding meetings to discuss the CPC-related businesses and disseminating policies made by the CPC.

43 Priorities related to the UTE are about making preparations for the UTE reports and inspection.

44 Priorities related to the university town are about the construction and the management of campuses in university towns and tackling problems caused by campuses at different locations.

45 Priorities related to salary are about increasing teachers' salaries based on performance and offering high salaries to staff holding special posts.

46 Priorities related to students are about facilitating student employment, promoting students' ethics and setting standards for student service.

47 Priorities related to staff are about recruiting more high-achieving academics by setting up special posts and providing professional development programmes.

48 Priorities related to international communication are about developing collaborations with overseas institutions and attracting more international students.

49 Priorities related to university culture are about maintaining the order and security of the campuses.

50 Priorities related to management are university policy making and implementation, such as the teacher appraisal policy.

51 Priorities related to reputation are about boosting City University's reputation.

52 Priorities related to university anniversary celebrations are preparations made for the celebrations.

The most frequently identified top priority was research. Research was identified as one of the priorities every year because of *University Institutional Rule Three*: to improve research performance in order to boost the university's reputation. The university leaders constantly urged teachers to improve their research performance in order to win more awards and grants, which were considered to be reputation indicators.

Research was often associated with graduate education. As mentioned above, the CPC Secretary stressed the need to obtain doctorate degree-awarding powers and to increase the size of the graduate student population because he regarded these as important reputation indicators. As a result of unremitting efforts the size of the graduate student population grew fast at City University (see Table 20).

Staff were also repeatedly selected as a top priority, because of the need to improve research performance. University leaders were aware that it was unrealistic to expect existing staff to improve dramatically in a short time. To have an immediate effect, the university therefore resorted to recruiting researchers who had had demonstrated outstanding research performance at other universities. In order to attract the best researchers in China and from overseas, City University proposed to pay ¥150,000,000 CNY (£15,000,000 GBP) for the move and ¥10,000,000 CNY (£1,000,000 GBP) to fund a research project (City University, 2005).

Management was also a top priority due to its association with research. Whenever management was mentioned, it was related to the President's intention to revise the existing teacher appraisal policy in order to motivate teachers to be more productive.

University town was identified as a top priority because of *University Institutional Rule Two*: to try to solve the problems caused by the dramatic increase in student numbers. City University had to deal with various issues caused by the move from the old campuses in the city centre to the new campus in university town in the suburbs of the city.

Agendas related to the CPC were a top priority due to *University Institutional Rule One*: to conform to government policies and requirements. Whenever a national policy was announced, City University issued a corresponding university policy and holds meetings to disseminate the national policy.

The UTE and the Quality Project appeared in the list of top priorities less frequently. This shows that they were less important than the priorities mentioned above. The UTE and the Quality Project were top priorities because of *University Institutional Rule One*: to conform to government policies and requirements. The timing as regards their appearances in the list suggests that the university quality agendas were driven by the national quality assurance policies. The UTE became one of the top priorities from 2003 to 2006, because this was the period for preparing for the UTE inspection. After the State Council launched the Quality Project, the Quality Project replaced the UTE and became a priority.

7.3.2 Programming and Implementation

City University started the preparation three years before the UTE inspection, the longest preparation time among the three case study universities. In 2003, the President's speech '*Seizing the Opportunity of Receiving the Evaluation and Working*

Hard to Achieve a Speedy Development' (UniCity Yearbook, 2003) marked the start of this preparation. The President gave his full endorsement to the UTE because he believed that 'the UTE evaluation scheme is very scientific' (UniCity Yearbook, 2003, 136). By 'scientific', the President meant that '*the UTE scheme* covers a wide range of issues of quality and investigates the quality of universities in depth' (UniCity Yearbook, 2003, 136).

The *objectives* set by the President's speech were improving the quality of undergraduate education and achieving 'excellence' in the UTE. The President regarded the UTE as a challenge, an opportunity and a campaign. The UTE was a challenge because there was a gap between the UTE requirements and the reality at City University. In order to achieve 'excellence', City University had to work hard to meet the UTE requirements, which was considered to involve making improvements in quality. Because of its potential implications the UTE was also regarded as a campaign, as illustrated by the quote below:

> We have to realise that the UTE is not only important for undergraduate teaching, it is also a ticket for the application for the degree-awarding power and the Key Disciplines. It is a test for the university's education sufficiency. Whether we can win this campaign is vital to the university's reputation, brand, survival and development.
>
> (UniCity Yearbook, 2003, 137)

As indicated by this quote, the President considered the UTE result to be a judgement regarding the quality of City University's education. Failure to achieve a good result would destroy the university's reputation. The President added later that achieving 'excellence' could also be used as a bargaining tool to negotiate with local and central government to obtain more resources and support (UniCity Yearbook, 2006). Hence, the President repeatedly stressed that 'the purpose of making preparation is to achieve "excellence" in the UTE' (UniCity Yearbook, 2003, 139).

The President's speech also indicated that all *human resources* would be allocated to the preparation. The President asked all department heads and the CPC Secretaries of departments to assume the responsibilities of organising the preparation for the UTE. The President added, 'Everybody should care about, participate in and invest in the preparation for the UTE and share the responsibilities' (UniCity Yearbook, 2003, 139).

Although one of the *objectives* identified by the President was to improve the quality of undergraduate education, the follow-up actions taken by City University mainly focused on achieving 'excellence'. From 2003 to 2005, City University organised several checks of student examination papers, compiled certain university quality assurance documents, which were in accordance with the UTE requirements, and evaluated departments according to the *2004 UTE Scheme*. In 2006, City University issued the *Rectification Plan for the Preparation for the UTE Inspection and the Work Schedule for Receiving the UTE Inspection* (UniCity Yearbook, 2006), which can be regarded as an official action plan for the UTE.

The *2006 Rectification Plan* demanded that all departments ensure that they would meet all indicators listed in the *2004 UTE Scheme* before the inspection. Significant efforts were made to ensure achievement of the quantitative indicators. For instance, the *2006 Rectification Plan* stated:

> Those who have not had teacher qualifications must obtain the qualifications within one year…If not, they have to be transferred to other positions.
>
> We have to invest more money in buying books in order to meet the UTE requirement of 100 books per student.

> The accounting department should collect and check the data of 'four kinds of funds'. Make sure the data from 2005 can meet the latest requirements of the UTE. Make sure the data shows that 30% of tuition income has been invested in the 'four kinds of funds' and the percentage has increased ever since.
> (UniCity Yearbook, 2006, 54-55)

These quotes show that meeting the UTE requirements became the main purpose, whereas improving the quality of undergraduate education was a by-product of this purpose. The third quote demonstrates that the university instructed its accounting department to manipulate data in order to meet the UTE requirements.

The President acknowledged that, compared to the effort expended on quantitative indicators, the university had not put enough effort into qualitative indicators, such as teachers' conduct and learning culture in the *2004 UTE Scheme*, so he reminded the departments to work on these areas at the final stage (UniCity Yearbook, 2006).

The *2006 Rectification Plan* also influenced the existing internal quality assurance policies, i.e. class observation and teacher appraisal. Class observation by leaders was initiated. The requirement that teachers who scored less than 75 (out of 100) in the student evaluation of teachers would be disqualified from receiving a promotion was added.

In the end, the university report compiled by City University met most of the 'A' criteria set by the *2004 UTE Scheme.* City University, as it had so badly wanted, achieved 'excellence' in the UTE.

Following the launch of the Quality Project, City University issued *City University Undergraduate Teaching Quality and Teaching Reform Implementation Plan*

(UniCity Yearbook, 2007) in 2007, which designated the *objectives* for the Quality Project. These were:

> We endeavour to win one or two national Excellent Course Awards, two to three provincial Excellent Course Awards and 11 to 15 city Excellent Course Awards and we will give 16 to 22 university Excellent Course Awards.
> (UniCity Yearbook, 2007, 102)

City University was enthusiastic about these awards because, firstly, they could serve as reputation indicators; secondly, award-winners could serve as models for other teachers to learn from and therefore the quality of undergraduate teaching would be improved (UniCity Yearbook, 2007). These hypotheses will be tested against teachers' accounts.

The *2007 Implementation Plan* also allocated *human resources* to the Quality Project. It requested that both university and department leaders assume responsibilities as regarded leading the Quality Project, but it did not require that all teachers participate in the Quality Project.

The *2007 Implementation Plan* also added new requirements to the existing internal quality assurance policies. It made teaching of undergraduates compulsory for all teachers at City University. Teachers who did not teach undergraduates were not eligible for promotion. Professors or associate professors' titles would be removed if they had not taught undergraduates in the past two years.

City University did not have detailed *action plans* for the Quality Project. The university left the implementation of the Quality Project to departments.

The number of Quality Project awards received by City University is summarised in Table 23.

	National	*Provincial*	*City*	*University*
Excellent Course Awards	3	9	33	51
Distinguished Teacher Awards	2	5	3	33
Teaching Group Awards	1	NA	4	NA

Table 23 The total number of Quality Project awards that City University had received by 2010 (Data extracted from City University website and the Quality Project website)

Table 23 shows that City University successfully met its *objectives*. In fact, it outperformed as regards the *objectives*.

7.4 Perceived Impact of the UTE

To investigate the impact of the UTE, accounts by the department heads are presented first, followed by teachers' accounts, in order to compare and contrast their accounts. Finally, the impact of the UTE on teachers is summarised.

7.4.1 Department Leaders' Perspectives

The impact of the UTE as perceived by the interviewed department heads relates to both *consequences* and *implementation actions*. One of the *consequences* mentioned by Head N (DepSc) and Head O (DepSc) was improvements in *infrastructure*. Head N stated:

> Before the UTE inspection a large amount of money was invested in our laboratories, so our equipment has improved dramatically.
>
> (Head N, DepSc)

Improvements in *infrastructure* are the *consequence* of the *implementation action* – investing in *infrastructure* – which was requested by *the 2004 UTE Scheme*. Without the UTE, the equipment in the laboratories might not have been improved so rapidly.

Head P (DepSoSc) and Head O (DepSc) mentioned the standardisation of documentation and procedures as a positive *consequence* of the UTE. Their perceptions are illustrated by the following quotes:

> There is an improvement in producing and archiving standardised documents.
>
> Head P (DepSoSc)

> Before the UTE, the teaching procedure was not standardised, so some teachers might omit some of the steps. Given there were many detailed indicators in the *UTE scheme*, now teachers are aware of what they are expected to do in each step.
>
> (Head O, DepSc)

The two quotes reflect what the two heads considered to be the standardisation of documentation and procedures to represent an improvement in quality because they defined quality as 'perfection and consistency'. They believed that, once documents met the same criteria and teaching followed the same steps, quality was assured.

Head L (DepHum) identified *implementation actions* and their *consequence* as being the impact of the UTE. She mentioned both positive and negative impacts, are illustrated by the following quote:

> The UTE was like a gust of wind. We had to work very hard to make up for the shortfall in documents. As any evaluation, there is always some positive impact. Teachers put extra effort into preparing lessons because they were concerned that the UTE evaluators might come to observe their classes… However, preparing documents wasted a huge amount of time. We have to ask the question whether the UTE contributes to the improvement of quality or just to window-dressing.
>
> (Head L, DepHum)

First, Head L acknowledged the positive impact of the UTE – putting extra effort into preparing lessons. This was the *consequence* of the *implementation action* designated by the *action plan*. Then, Head L added another *implementation action*, which was preparing documents that met the UTE requirements. She perceived this as a negative impact of the UTE, because she considered using recently made documents to fake

existing documents unethical conduct. Additionally, the burden added by producing the documents distracted teachers from their daily routines.

7.4.2 Teachers' Perspectives

Teachers identified *implementation actions* and *consequence* as being the impact of the UTE. The *implementation actions* are those actions designated by the university's *action plans*. The interviewed teachers had more complaints than their counterparts did at the other two case study universities. During the UTE inspection, the stress experienced by teachers was beyond that with which teachers were able to cope. Associate Professor L (DepSc) described the situation at that time:

> We were busy preparing documents, more specifically forging documents. I was too exhausted to concentrate on teaching. It was putting the cart before the horse… During lecturing, we had to monitor if there was a breach of discipline all the time, which was completely a distraction from teaching. There were staff patrolling outside the classroom to check the discipline too. Even drinking water in the classroom was not allowed at that time. These were merely a formality and inhumane.
>
> (Associate Professor L, DepSc)

Associate Professor L's reference to forging documents confirms Head L's (DepHum) account. Moreover, the *implementation actions* designated by City University's *action plans* went beyond the unreasonableness and the lack of ethicalness complained of by the teachers at the Project 985 University and the Project 211 University: they reached a level of absurdity, with teachers and students not allowed to drink water in the classroom and send staff deployed to patrol outside the classrooms.

Some absurd requirements persisted after the UTE. Professor I (DepHum) criticised the requirement to ask teachers to submit both a hardcopy and electronic version of the same set of documents, considering this to be a waste of resources and energy. Lecturer L (DepSc) complained about the university policy of storing student

assignments, instead of letting students keep them. Both university requirements were made as precautions of potential evaluations. These absurd requirements reflected overreaction on the part of City University.

At City University, Associate Professor K (DepHum), Associate Professor J (DepHum) and Associate Professor L (DepSc) identified the standardisation of documentation as a positive *consequence* of the UTE, but at the same time they did not think that this contributed to the improvement of their teaching. The contradictory view is illustrated by the following quote:

> The UTE has a huge impact on me. I think that the biggest impact of the UTE is standardisation. The standardisation is achieved through standardisation of documentation, such as curricula, syllabi and examination papers. Checking these documents is a good thing but they only check the format … not the content. This kind of standardisation washes away the essence of higher education… The focus should be on the essence of the course: whether the content is adequate and whether it can provoke students to think.
>
> (Associate Professor J, DepHum)

This contradictory view regarding the standardisation of documentation reflects different definitions of quality. Acknowledgement of the benefit of the standardisation of documentation suggests an understanding of quality as signifying perfection and consistency. On the other hand, the question regarding the essence of education indicates an understanding of quality as signifying transformation.

7.4.3 Summary

How the interviewed heads and teachers perceived the impact of the UTE is summarised in Figure 20 and Figure 21. The impact of the UTE included the *implementation actions* prescribed by the *action plans* and the *consequences* of these *implementation actions*. Among all of the *implementation actions,* investing in infrastructure was an action taken by the university. The rest of them were actions

taken by teachers. The interviewed heads and the teachers mentioned the same *consequences*: the standardisation of documentation, the waste of time and effort and the undermining of ethics. The differences were that the interviewed heads added the standardisation of procedures and the provision of impetus to ensure teaching quality on the one hand; and the interviewed teachers raised the issue of the interference with academic freedom on the other hand.

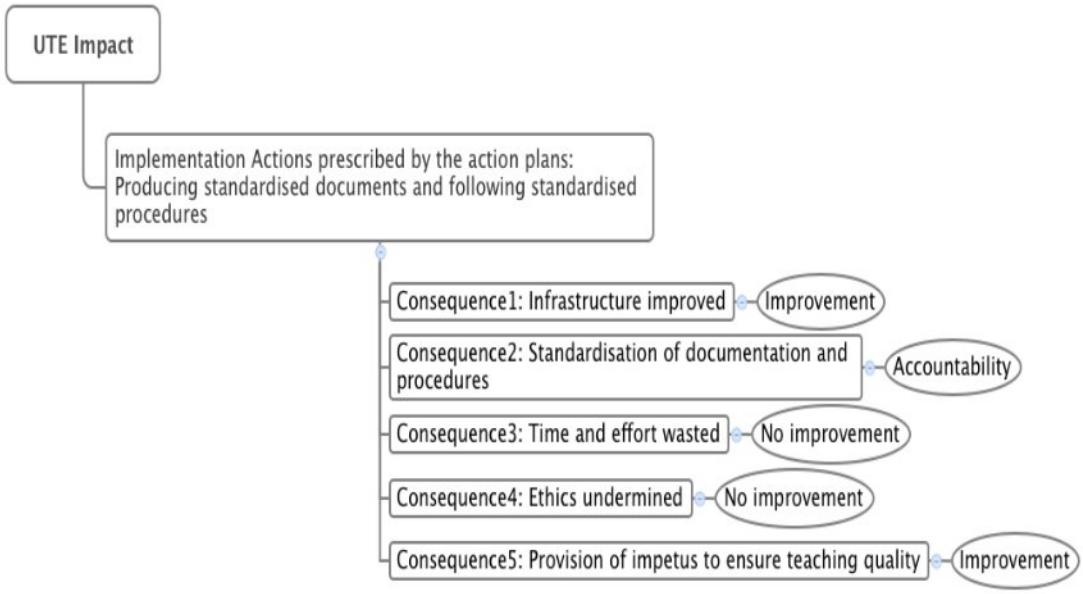


Figure 20 The impact of the UTE as perceived by head interviewees at City University

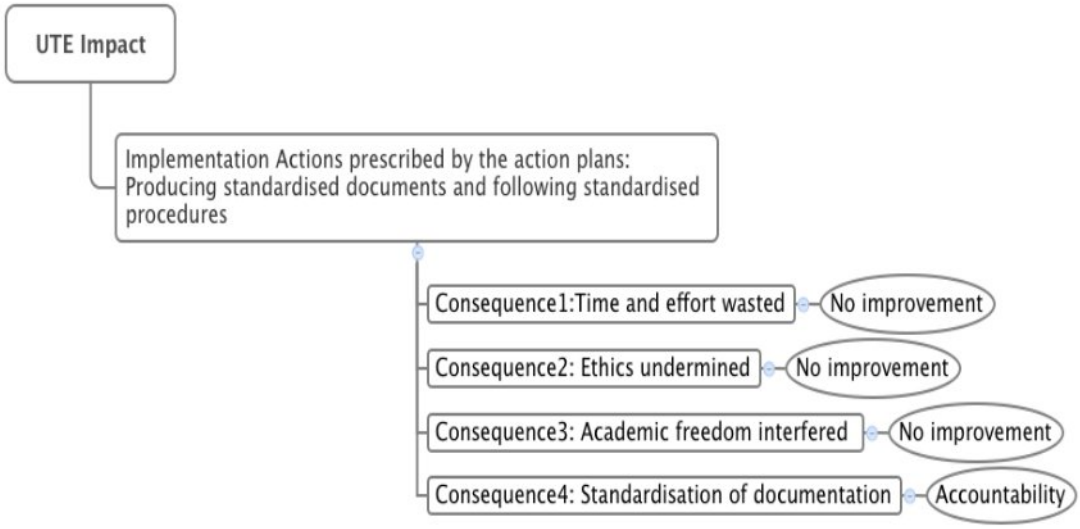


Figure 21 The impact of the UTE as perceived by teacher interviewees at City University

The interviewed heads' and teachers' perceptions differed from those of their counterparts at the other two universities in two ways. Firstly, there were more heads and teachers mentioning unethical conduct at City University. In fact, besides a head at the City University, none of the interviewed heads articulated unethical conduct. As mentioned in the Programming of the UTE (see Section 7.3.2), even the President suggested adopting an irregular approach, i.e. manipulation of data, to meet the UTE indicators. This was very likely attributed to the big gap between what City University had achieved and what the UTE required. Secondly, only at City University were there teachers who perceived the standardisation of documents as a positive *consequence*.

7.5 Perceived Impact of the Quality Project Awards

The awards won by the Humanities Department, the Sciences Department and the Social Sciences Department at City University are listed in Table 24.

		DepHum	DepSc	DepSoSc
Excellent Course Awards	National	0	0	0
	Provincial	0	0	0
	City	0	1	0
	University	0	1	0
Distinguished Teacher Awards	National	0	0	0
	Provincial	0	0	0
	University	0	2	1
Teaching Group Awards	National	0	0	0

Table 24 The number of Quality Project awards won by different departments at City University (Extracted from City University website)

Table 24 shows that the amount of awards won by City University is much smaller than those won by the Project 985 University and the Project 211 University. None of the departments at City University won national or provincial awards.

7.5.1 Department Leaders' Perspectives

Head N and Head O at the Sciences Department argued that the Quality Project awards offered acknowledgement to teachers who had been dedicated to teaching. Both of them mentioned the case of Teacher Q. Head O said:

> At our department, the course taught by Teacher Q has been granted a University Excellent Course Award. Teacher Q spends a lot of time on this course. We all admire him…Teacher Q is also a winner of the [University] Distinguished Teacher Award. This is a very rare case for an associate professor, who does not have a strong research background, to win these awards.
>
> (Head O, DepSc)

This quote shows that Head O was very pleased that Teacher Q received the long overdue acknowledgement of his dedication to teaching. On the other hand, although Head O did not say this explicitly, this quote implies that there was an issue regarding the validity of receiving these awards for teaching. To be able to win these awards for teaching, having only excellent teaching performance was not sufficient. Winners normally were those who had demonstrated excellent research performance. This validity issue was articulated by Head P (DepSoSc):

> Now I think that the Quality Project is not very meaningful. Take the Distinguished Teacher Awards and the Excellent Course Awards for example … Their selection criteria … weigh teachers' research skills and output much more.
>
> (Head P, DepSoSc)

Head P pointed out that although these awards were meant to reward teachers who had demonstrated excellence in teaching, in reality they became awards for research performance.

The Humanities Department was the only department that did not win any Quality Project award, out of all interviewed departments. Hence, Head L's (DepHum) perception was different from that of her counterparts. Head L felt ashamed of not winning any award and perceived the Quality Project awards positively, because they offered financial rewards. She wished her department could win some awards, so as to obtain some funding.

7.5.2 Teachers' Perspectives

Because Teacher Q refused the invitation to participate in this research, the researcher was not able to check if Teacher Q perceived the same impact – offering acknowledgement – as assumed by Head N (DepSc) and Head O (DepSc). The researcher interviewed Professor J (DepSc) – a winner of a City Excellent Course Award. It showed that receiving a City Excellent Course Award was unable to provide Professor J with further motivation. He admitted that he did not have any intention to apply for more awards 'because the teacher appraisal policy emphasises research more, so there is more pressure on research' (Professor J, DepSc).

The majority of teacher interviewees did not apply for the awards because of either the problem of validity or teachers' contempt for official awards. Some interviewed

teachers pointed out the contradiction between the stated aim claimed by the Quality Project awards and the selection criteria, as suggested by Professor I (DepHum):

> Don't think the selection of the winners of the Distinguished Teacher Awards is based on teaching outcome. In the end, it is based on research outcomes.
> (Professor I, DepHum)

This quote reflects the disappointment shared by many interviewed teachers at the City University.

When asked about the impact of the Quality Project awards, many teachers answered that they had no interest in them. Their opinions are illustrated by the following quote.

> The Quality Project awards have little influence on me. I just teach according to my conscience.
> (Lecturer K, DepHum)

Lecturer K believed that teachers taught responsibly because of their professionalism, so he did not need awards to motivate him.

Professor L (DepSoSc) added another reason for the general contempt. He said:

> I think that the Ministry of Education launched the Quality Project in order to make troubles. If it does not launch new initiatives and make new requirements, how can it demonstrate its authority?
> (Professor L, DepSoSc)

Professor L's quote reveals a strong distrust of government policies. Government policies were regarded as a means to maintain control over higher education.

7.5.3 Summary

The impact of the Quality Project awards as perceived by the interviewed heads and teachers at City University is summarised in Figure 22 and Figure 23. The *implementation action* was filing the application and the *interaction* was winning awards. Some of the heads interviewed assumed that winning awards would provide

impetus to improve teaching through acknowledgement. However, the winner's account did not support this assumption. This suggests that the incentive provided by a city award was unable to counterbalance the incentive to research offered by the university's teacher appraisal policy. Some interviewed head and teachers attributed the lack of impact to the problem of validity. They were convinced that the Quality Project awards were designed to reward research performance instead. Some teacher interviewees added contempt for official awards as another reason, either because there was no need for them or because it was a protest against the central government's authority.

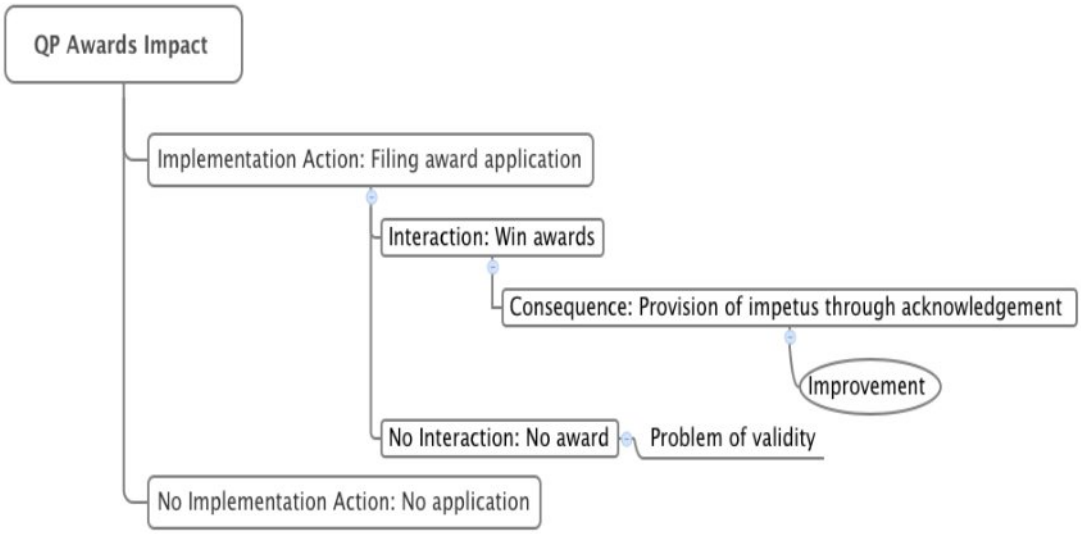


Figure 22 The impact of the Quality Project (PQ) awards as perceived by head interviewees at City University

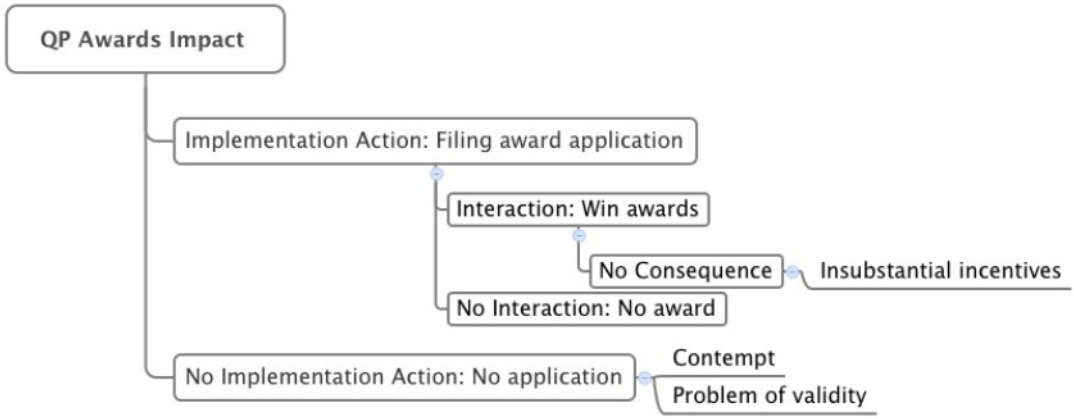


Figure 23 The impact of the Quality Project (PQ) awards as perceived by teacher interviewees at City University

7.6 Perceived Impact of Class Observation

Different types of class observations are introduced by the interviewed heads of the department first. The interviewed heads' and teachers' perceptions are then compared and contrasted to evaluate the impact of class observation.

7.6.1 Department Leaders' Perspectives

Head L (DepHum), Head O (DepSc) and Head P (DepSoSc) explained that there were mainly two types of class observations at City University: class observation carried out by peers and that by university inspectors. Class observation by peers is a university policy and departments undertake the responsibility for organising this observation. There were two forms of class observation by peers, they explained. The first was to observe exemplary teaching, from which other teachers were required to learn. The second was observing new teachers' teaching, which other teachers were required to provide suggestions in response to, to help new teachers to improve.

Head L (DepHum), Head O (DepSc) and Head P (DepSoSc) thought the practice of peer class observation was very helpful because 'teachers can be inspired and they can reflect on their own teaching' (Head P, DepSoSc). However, Head L (DepHum)

felt that many teachers were not enthusiastic about observing their colleagues' classes. Moreover, she added that, even when teachers took part in the observation, 'it would not be very helpful if teachers just left after the observation; it would be more helpful if teachers had a discussion after the observation'.

The heads interviewed at City University pointed out that class observation by university inspectors was used as a means to assure teachers' accountability rather than enhance their practices. The purpose was to monitor teaching and the focus was on finding out problems in teachers' teaching. Because of the focus on accountability, Head L (DepHum) noticed that teachers had a negative attitude towards this form of observation.

However, based on heads' accounts, the researcher found that the practice of class observation by university inspectors was not able to assure accountability for two reasons. First, teachers were not held accountable even when university inspectors identified problems in their teaching. Head L (DepHum) said that teachers were very defensive in response to any criticism, due to the generally negative attitude towards class observation by university inspectors, so normally teachers were not informed about the inspectors' comments unless inspectors identified serious problems. When problems were serious, the interviewed heads dealt with the university inspectors' comments in the following way:

> We cannot show the university inspectors' comments directly to teachers. We have to process the comments. We have to talk to teachers cleverly. Inspectors' comments just serve as reminders.
>
> (Head O, DepSc)

This approach was also adopted by Head L (DepHum) and Head P (DepSoSc) to avoid confrontation. It was teachers' choice to decide whether to act upon the

recommendations or not. Due to the negative attitudes towards the practice of class observation by university inspectors, teachers were less willing to follow the advice of inspectors. Consequently, Head P (DepSoSc) admitted ‘those talks have little impact on teachers’.

Second, class observation by university inspectors was unable to cover all teachers at City University. Head L (DepHum) and Head O (DepSc) revealed that university inspectors tended to observe new teachers, teachers who had submitted applications for promotions or awards, and those who were known to be having problems with teaching. Thus, university inspectors rarely observed classes taught by experienced teachers who were not applying for promotion.

The heads interviewed had different opinions about the reliability of class observation. Head L (DepHum) questioned this reliability because she was aware that ‘different people perceive the quality of the same class differently’. In contrast, Head P (DepSoSc) had confidence in the reliability of class observation. He felt that teachers who received bad comments in class observation were teachers who had low scores in the student evaluations of teachers.

7.6.2 Teachers’ Perspectives

The impact of class observation varied depending on the type of class observation. Class observation by university inspectors had impact on teachers who were frequently observed, whereas class observation by peers organised by departments had impact on those who were observed.

Professor H (DepHum) and Lecturer M (DepSc) admitted that class observation by university inspectors reminded them to deliver high quality teaching. They had to make good preparations for lessons, because, as Lecturer M (DepSc) said, 'if the inspectors come to observe my lesson and I did not prepare it well, it would be very embarrassing'. The reason why Professor H and Lecturer M felt the observation had impact on them was because they belonged to the categories of teachers who were observed by inspectors frequently. Lecturer M was a new teacher and Professor H was among those teachers known to be having problems with teaching. Professor H confessed that once he had a very low score in the student evaluation of teachers, because he did not make his teaching interesting and was very strict with students.

In contrast, teachers who were not frequently observed by the inspectors did not feel any pressure. Professor J (DepSc) said that he had not been observed by university inspectors for almost ten years. As a result, he did not feel any pressure.

In fact, the most frequently mentioned factor preventing class observation by university inspectors from having impact was the issue of reliability. Firstly, the reliability of the observation could be jeopardised by the lack of expertise in the subject areas. The following quotes illustrate this problem:

> Inspectors from different background do not quite understand our subject. As a result, they mainly look at the forms: whether we use multimedia or not, whether our slides look nice or not or whether the class atmosphere is alive.
>
> (Professor I, DepHum)

> The knowledge of the university inspectors in my field is no better than my students' who failed the examinations, so I do not care about them.
>
> (Professor L, DepSoSc)

Professor L's statement is very bold but it makes a valid point. In his opinion, inspectors who did not have expertise in the field were not qualified to judge his

teaching. Professor I pointed out that the lack of expertise hampered the inspectors' abilities to provide comments on the content of teaching. This was the common problem identified by interviewed teachers across the three case study universities.

Secondly, the reliability of class observation could be harmed by the bias introduced by age. Given that most of the university inspectors were retired teachers, Professor K (DepSoSc) reasoned:

> The problem is that retired teachers and young teachers have different understandings of high quality teaching. Retired teachers tend to equate teaching with lecturing, but now we emphasise teacher-student interaction more. When they see teachers allocate half of the time to interaction, retired teachers assume that those teachers are being lazy and have not prepared enough material to teach students.
>
> (Professor K, DepSoSc)

This quote demonstrates the issue regarding reliability caused by different understandings of high quality teaching, which results from different teaching theories having prevailed in different eras.

In order to tackle this problem, City University started to appoint some teachers who were still working as university inspectors. Professor L (DepSoSc) was one of these inspectors. However, another problem that undermined the reliability of class observation then emerged. Professor L felt unease if he did not rate teachers' teaching as 'excellent', because 'teachers would think that I have problem with them personally, rather than regarding it as a professional issue given that it is my duty as a university inspector to rate their teaching' (Professor L, DepSoSc). This reflected a lack of culture of constructive criticism at City University. This factor was the third contributor to the issue of reliability of class observation. Since it was unable to point

out the weaknesses in teaching, class observation by university inspectors failed to facilitate improvement.

Besides the problem of reliability, Professor L's (DepSoSc) experience shows that class observation by university inspectors failed to hold teachers accountable even when their teaching was not acceptable. Professor L (DepSoSc) did give a teacher low rating when he taught badly:

> When he was teaching, the teacher never looked at students nor stood up or wrote something on the blackboard or showed some slides. He did not care about students' reactions. He even taught wrong things to the students. How could I not rate his teaching as 'poor'. However, in the current system it is impossible to fire him. If I were the President of the University, I would definitely fire him.
>
> (Professor L, DepSoSc)

This quote illustrates that, similar to the department heads, the university inspector could not take any more action than giving the teacher a bad rating.

In terms of class observation by peers, Lecturer M (DepSc) and Professor J (DepSc) added detail regarding how this was conducted at the Sciences Department. First, teachers who had been asked to attend the class observation observed the first half of the lecture. Then they moved to another room to discuss the observed teaching. As a new teacher, Lecturer M's lesson was observed by all of his colleagues at the Sciences Department. As an experienced teacher, Professor J's lesson was observed by just one or two of his colleagues. The way in which Lecturer M and Professor J were informed of the feedback also varied. The administrator informed Lecturer M about his colleagues' feedback afterwards, while Professor J's colleagues gave him their feedback individually afterwards.

Lecturer M (DepSc) was impressed by the helpfulness of his colleagues' feedback. He said:

> When my colleagues were discussing my teaching, they engaged in a lively discussion. Their comments were more helpful than the university inspectors, because they have expertise in our field.
>
> (Lecturer M, DepSc)

Having many teachers attending the class observation gave rise to a lively discussion, which generated more ideas. Moreover, because Lecturer M's colleagues did not need to inform him of the feedback face to face, they were able to speak out about the weaknesses in his teaching. These two factors made the class observation more beneficial. In Professor J's (DepSc) case, there were only one or two colleagues, so observation did not lead to a lively discussion. Furthermore, it was very likely that Professor J's colleagues were reluctant to point out the weaknesses in his teaching because they informed Professor J of their feedback face to face.

Lecture N (DepSoSc) and Lecturer O (DepSoSc) added that City University encouraged teachers to observe each other's lessons by themselves. They confessed that were not motivated to do so, so they just filled in the observation forms and pretended that they had completed the observation.

Besides class observation by university inspectors and peers, teachers referred to the practice of class observations by department inspectors and leaders. Lecturer M (DepSc) mentioned class observation by department inspectors when he was talking about class observation by peers. He considered that the help offered by the department inspectors was as useful as that offered by peers. Class observation by leaders was only mentioned by Associate Professor J (DepHum). The department head recently observed her class. She said 'the department head acknowledged my

strengths, but I think it would be more helpful if he pointed out the problems with my teaching as well'.

7.6.3 Summary

The interviewed heads' and teachers' perceptions on the impact of the class observation are summarised in Figure 24 and Figure 25. Teachers were absent from the *implementation action*. The observations were conducted by leaders, university inspectors and peers. The *interaction* was receiving the feedback. The heads and teachers interviewed had little in common regarding the impact of the class observation.

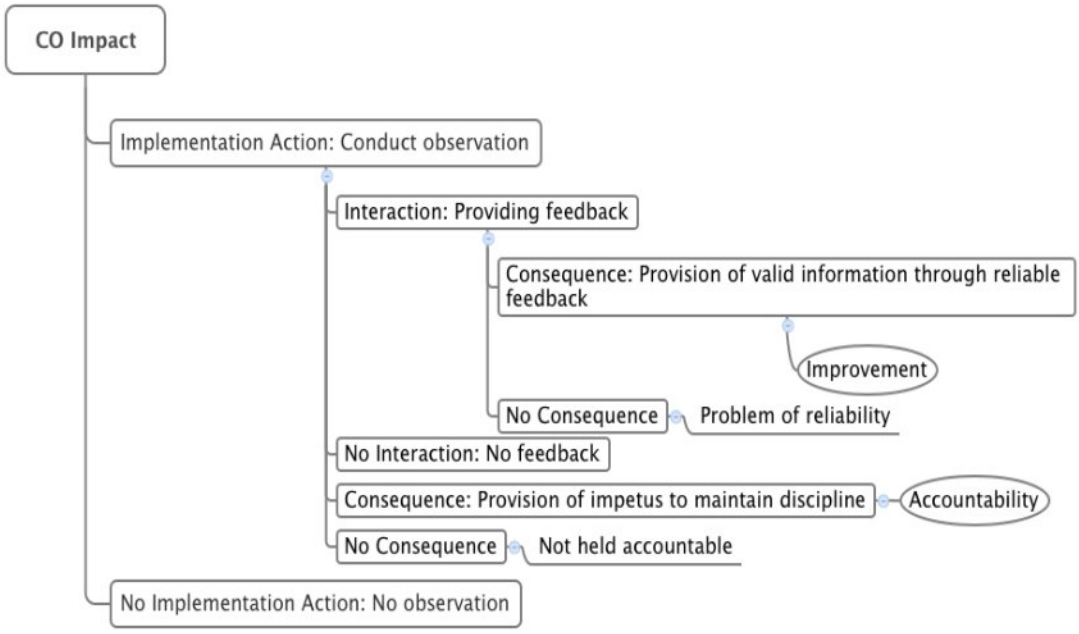


Figure 24 The impact of the class observation (CO) as perceived by head interviewees at City University

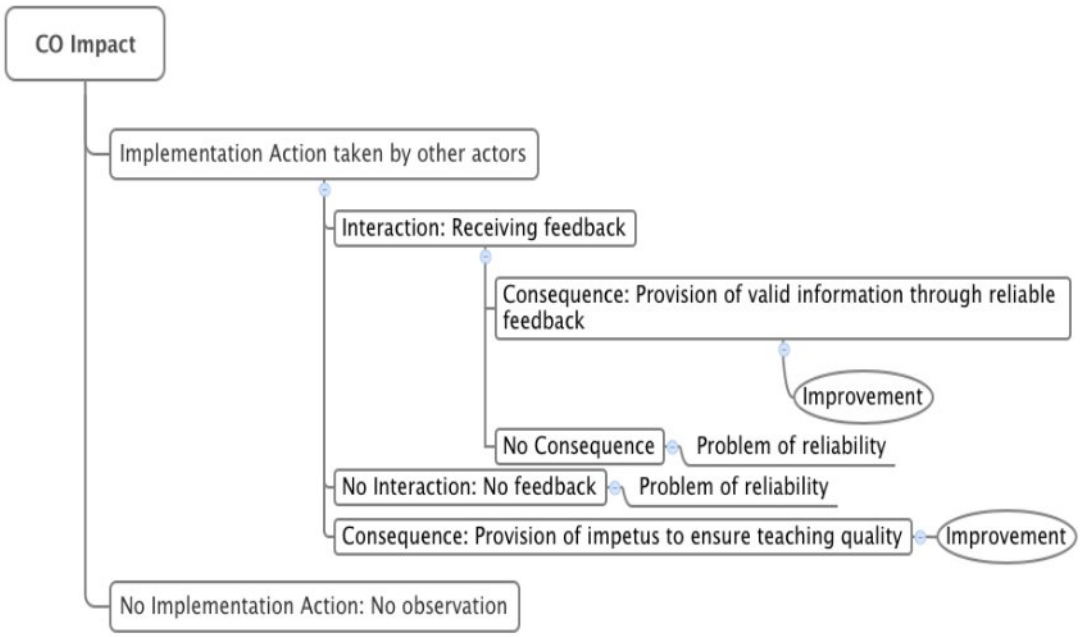


Figure 25 The impact of the class observation (CO) as perceived by teacher interviewees at City University

When there was an *interaction*, although some of the interviewed heads and teachers mentioned the *consequence* of the provision of information through valid feedback, those heads and teachers were referring to different types of class observations. The former were commenting about the *consequence* of class observation by university inspectors; the latter were talking about the *consequence* of class observation by peers. In other words, the interviewed teachers did not think class observation by university inspectors could have this *consequence*. However, some of the interviewed heads and teachers agreed on the reason for the lack of *consequences* – a problem of reliability. The level of reliability could be harmed by a lack of expertise in the subject, the bias caused by age and the reluctance to provide frank feedback. The problem of reliability also prevented teachers from asking for feedback.

In terms of the unorganised class observation by peers, some heads assumed that the observers could obtain new information from reflection. However, some teacher interviewees confessed that they pretended that they completed the observation.

Some interviewed teachers who belonged to the group subject to frequent observations admitted that the class observation provided them with impetus to ensure a high quality in their teaching. On the contrary, when teachers were not frequently observed, they could not identify this *consequence*. Moreover, the heads and a teacher who had experience of being an inspector added that being unable to hold teachers accountable was another factor attributed to the lack of a *consequence*.

7.7 Perceived Impact of the Student Evaluation of Teachers and Teacher Appraisal

As explained in 7.2.2, one of the criteria for promotion is home institutions' judgements regarding the quality of applicants' teaching, which should include the results of student evaluation. However, neither head respondents nor teacher respondents related the student evaluation of teachers to teacher appraisal in the interviews. This response is different from that of the teachers who were interviewed at the Project 985 University and the Project 211 University. The interviewed heads' perceptions of the impact of the student evaluation of teachers are provided first, followed by the perceptions of teachers interviewed.

7.7.1 Department Leaders' Perspectives

The interviewed heads at City University confirmed that teachers could access the results of the student evaluation of teachers online and the ranking was also published. Therefore, they assumed that teachers use these results to improve their teaching.

The majority of the heads interviewed had complete confidence in the reliability of the student evaluation of teachers, whereas only one head had some reservations and one head expressed a lack of confidence. The degree of confidence was determined by his or her perception of the reliability – whether students gave low scores to teachers who were strict with them or not.

At the Humanities Department, Head M trusted student evaluation while Head L had some reservations. Head M's confidence lay in the methodology. He argued that the university removed the highest score and the lowest score before they calculated teachers' scores in the student evaluation. However, this method was neither mentioned in the university documents nor in other interviewed heads' or teachers' accounts. Even if the university removed the highest and the lowest score, this would not mean that all the outliers would be removed from the calculation and thereby the distribution would not be skewed. Hence, the method mentioned by Head M would not ensure the reliability of the results necessarily.

Head L (DepHum) trusted that the majority of students would rate their teachers truthfully but she was aware that a small number of outliers would be able to distort the overall results. She spoke about her personal experience:

> The majority of students act according to their conscience and they are fair. However, a small number of students did not take learning seriously. They skipped their classes and did not do their assignments. Even if I wanted to give them scores in the examinations, it would be impossible, because they just handed in blank examination papers. Because they failed in their examinations, they were not happy. A few very low scores could make the total score very low as well.
>
> (Head L, DepHum)

As a result, Head L said that normally her scores in the student evaluation of teachers were not very high.

At the Sciences Department, Head N was confident about the reliability of the student evaluation of teachers, but Head O was not. Head N was very pleased with his department's performance in student evaluation. He said ' in the fast five years, our department ranked the top among all Sciences Departments' (Head N, DepSc). In contrast, Head O (DepSc) did not trust the reliability of student evaluation, because students were forced to participate in the evaluation. He said:

> Students do not evaluate their teachers voluntarily. It is a task that they have to perform…As a result, students think that it is a burden…Some students just randomly tick the boxes. (Head N, DepSc).

Head N believed that the mandatory nature of the evaluation led to the carelessness about the rating of teachers. Head N also added the problem of students giving low scores to teachers who were strict.

At the Social Sciences Department Head P and Head Q had confidence in the reliability of student evaluation and they assumed that 'teachers who ranked at the bottom will be motivated to improve their teaching' (Head P, DepSoSc). However, no action had been taken against those teachers (Head Q, DepSoSc), although the department was considering transferring teachers who ranked in the bottom 5% to administrative positions (Head P, DepSoSc).

Ironically, it seems City University did not have confidence in its student evaluation of teachers. Head L (DepHum) informed the researcher that at the beginning the decisions regarding City University Excellent Teaching Awards had been mainly determined by teachers' results in the student evaluation of teachers. Later, this was replaced by comments from the class observation by university inspectors.

7.7.2 Teachers' Perspectives

All interviewed teachers at City University had complete or partial confidence in the reliability of the student evaluation of teachers. Some teachers had complete confidence simply because of their belief in student evaluation in general. For instance:

> I think the student evaluation of teachers is like an election. Every citizen will not vote randomly. If teachers teach really well, students will not give them low scores.
>
> (Lecturer M, DepSc)

This quote reflects a complete confidence, almost an excessive faith, in student evaluation. Lecturer M's confidence in student evaluation could be fed by his satisfactory results. He said '90% of students gave me positive feedback. I am very pleased' (Lecturer M, DepSc).

Professor I (DepHum) thought that student evaluation was more reliable than class observation by university inspectors. Hence, student evaluation was trustworthy. However, the researcher would argue that, even if student evaluation were more reliable than class observation by teachers, this would not automatically qualify student evaluation as being a reliable measurement.

Only a small number of teachers had some reservations about the reliability of student evaluation. Lecturer K (DepHum), Associate Professor L (DepSc) and Professor J (DepSc) acknowledged that the student evaluation of teachers was biased towards certain types of teachers. Associate Professor L (DepSc) admitted that 'students often rate me higher than I deserve', because she was not strict with students. Lecturer O (DepSoSc) noticed that students were biased towards teachers who had fun in

teaching even at the cost of the quality of teaching. He often heard the following reflection from students after they graduated:

> At that time (when students were doing undergraduate studies) students considered a teacher who taught really well, because the teacher told them many jokes or used many interesting illustrations. Now when the students look back, they do not think these jokes or examples are the most beneficial knowledge that they have learnt during their undergraduate degrees.
>
> (Lecturer O, DepSoSc)

Regardless of this, Lecturer O (DepSoSc) still considered the results of student evaluation to be reliable because 'students evaluate teachers according to how they feel' (Lecturer O, DepSoSc). At the very least, 'students are able to judge whether teachers take teaching seriously or not' (Lecturer O, DepSoSc).

One reported impact of the student evaluation of teachers was that it led to an intention to please students. This *consequence* was mentioned by Professor H (DepHum). Professor H was very strict with students. 'For those who often skipped classes or showed up late, I just gave them very low scores in their evaluations (Professor H, DepHum).' He also admitted that the style of this teaching had been very 'academic' – in other words, mainly constituting lecturing and content that was difficult to understand. Consequently, Professor H received a very low score in the student evaluation of teachers. Although a bad result could not affect his salary or promotion, Professor H was extremely embarrassed because he was very concerned about his image among his colleagues. After the evaluation, Professor H became less strict with students and made his teaching more interesting and accessible.

7.7.3 Summary

The impact of the student evaluation of teachers as perceived by the interviewed heads and teachers is summarised in Figure 26 and Figure 27.Teachers were absent

from the *implementation action*. Although teachers had access to their results of student evaluation and even the ranking was published, not all teachers would read this information. Therefore, the *interaction* was completed when teachers were aware of their results.

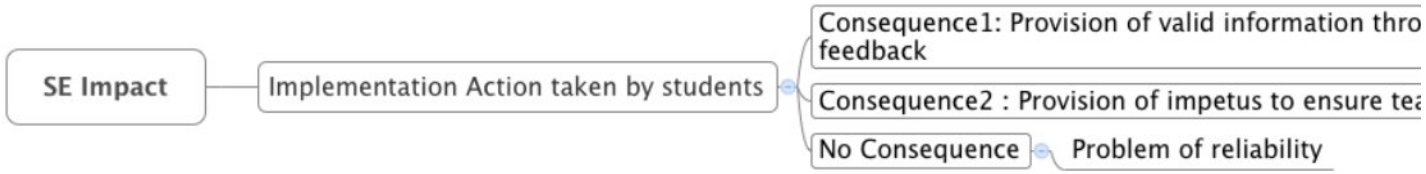


Figure 26 The impact of student evaluation (SE) as perceived by head interviewees at City University

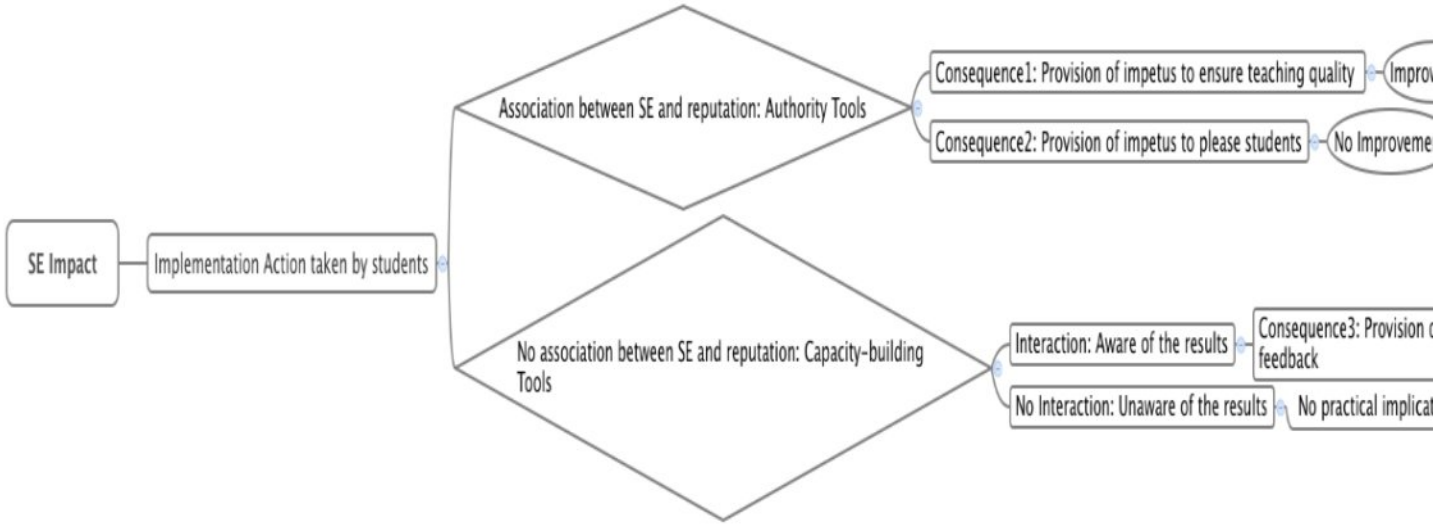


Figure 27 The impact of student evaluation (SE) as perceived by teacher interviewees at City University

The interviewed heads and teachers agreed that the student evaluation of teachers could provide information and impetus to ensure teaching quality to teachers but there was a slight difference between the groups' accounts. Some teachers added the conditions for information and impetus to happen. Information could be provided only when there was an *interaction*; impetus could be provided only when the stakes of student evaluation were raised by the potential damage to teachers' reputation. Some interviewed teachers were not interested in knowing their results. It could be attributed to lack of practical implications. A difference distinguishing teachers' perceptions of the reliability from their counterparts at the other two universities was that all the interviewed teachers and the majority of interviewed heads at City University had (a broad) confidence in the reliability of student evaluation.

7.8 Unaddressed Tensions and Factors Affecting the Assurance of the Quality of Teaching

Based on the interview data, this section answers several fundamental questions in relation to quality assurance: what are the problems of the existing quality assurance mechanisms perceived by heads and teachers' perspectives? what drives heads and teachers to invest in teaching? and what actually helps heads and teachers to improve teaching along the way? The interviewed heads' and teachers' accounts demonstrate that the existing quality assurance policies fail to address these issues. The answers to these questions will provide some insights into how to improve current quality assurance policies.

7.8.1 Problems with Current Quality Assurance Mechanisms

There was a consensus among most of the interviewed teachers at City University that the biggest obstacle to promoting high quality teaching was the existing teacher appraisal policies. As mentioned before, the President of the University claimed that both teaching and research were the focus of the university. However, almost all interviewed heads and teachers had the impression that in practice research was the focus. For instance, Head M (DepHum) argued that:

> Although we claim that teaching is our focus, it is just a claim, because teachers were evaluated mainly by their research performance.
>
> (Head M, DepHum)

Head M pointed out that the importance of research was highlighted by teacher appraisal policies, which affected teachers' incomes and promotion.

According to *City University Allowance Allocation Scheme* (City University, 2002), a large portion of teachers' salaries comes from teaching. Lecturer O (DepSoSc) admitted that more than two thirds of his salary was generated through teaching. In this case, why did teachers still feel that income from teaching was not substantial enough? Head Q (DepSoSc) explained that this was because of the substantial gap between the income generated from teaching and from research and the non-financial benefits brought by research. Head Q elaborated:

> Publishing one research article could receive a ¥10,000 CNY (£1,000 GBP) reward. However, let me count how much a professor could earn from teaching a course [for one term]…The professor could earn around ¥2100 CNY (£210 GBP) but nothing else. If the professor published one article, he or she would also gain a reputation, besides the ¥10,000 CNY's reward. Even the government [values high-achieving professors and] asks them to act as consultants. Publications have more added value.
>
> (Head Q, DepSoSc)

The quote above provides a concrete example of the substantiality of the difference between the benefits (both financial and non-financial) generated from teaching and from research. Therefore, although income from teaching constitutes a big proportion

of teachers' salaries, teachers' salaries are dwarfed by the benefits gained through research.

Whether teachers are promoted or not is also mainly determined by research performance. Associate Professor J (DepHum) pointed out that teaching was not mandatory for promotion. This is consistent with the promotion criteria presented in Section 7.2.2. Teacher Q, the associate professor who won the University Distinguished Teacher Award (who was mentioned by Head O, DepSc, in Section 7.5.1), was referred to as a case illustrating how teachers who were only good in teaching were disadvantaged in promotion. Lecturer M (DepSc) revealed that Teacher Q had applied for a professorship many times but he failed, because he did not have a strong research record. The experience of teachers like Teacher Q reminded other teachers that excellence in teaching would not make them professors.

At City University, there were many teachers who did not have a strong research record, because the university is a teaching-intensive university and teachers were not expected to conduct research before 2000. Many teachers were only interested in teaching or never received systematic training in research. For example, Lecturer L (DepSc), a holder of a Master's degree, admitted that 'I am not a research-oriented teacher'. She often felt 'I have put a lot of effort into teaching but these efforts were not recognised. When I was applying for promotion, I had nothing to present' (Lecturer L, DepSc). This problem was also acknowledged by Head O (DepSc). He said:

> Now many teachers do not have a sense of achievement, because they are appraised by the number of publications and research grants. Teachers who are good in research can feel a sense of achievement but not those teachers who

> are only good in teaching. They have lost their interest in teaching, because how well they teach students does not matter anymore.
>
> (Head O, DepSc)

Without a sense of achievement, teachers lost the motivation for teaching. On the other hand, due to their lack of interest in, or capability as regards, doing research, they were unable to perform well in research either. This was a lose-lose situation for both teachers and the university.

Admittedly the promotion criteria were made by the provincial government. However, City University did not make university policies to counterbalance the emphasis on research. Instead, the university also put more weight on research, as a result of *University Institutional Rule Three*: to improve research performance in order to boost the university's reputation.

In fact, the President often stressed the importance of research in meetings with department leaders. Before the interview, Head Q (DepSoSc) attended a meeting with the President along with heads from other departments. Head Q said that in the meeting the President set some concrete targets for the Social Science Department: publishing a certain number of articles in particular journals. In the meeting the President explained why he had to put more weight on research. The President said:

> The society does not judge our university by how well we educate our students. Take a look at rankings. They only look at research, such as publications and projects. In order to survive, we have to adapt to this.

What the President said in the meeting was different from what he claimed in his speeches. As cited in Section 7.1.2, the President maintained that both teaching and research were the focus of the university. This quote reveals that the focus was on research, due to external influences.

7.8.2 What Drives Teachers to Invest in Teaching?

The interviewed teachers at City University also mentioned dedication, sense of responsibility and passion for teaching as their drivers. These teachers had a particularly strong emphasis on students, compared to their counterparts at the Project 985 University and the Project 211 University.

At City University, some interviewees were motivated by dedication (i.e. had an urge to achieve their personal best) because they wanted to contribute to student learning. For example, when Associate Professor K (DepHum) was asked what drove him to invest in teaching, he answered:

> I just want to do things better…In the end, the ultimate goal … is to develop students' critical thinking.
>
> (Associate Professor K, DepHum)

The simple idea of 'doing things better' reflects dedication. For Associate Professor K, the aim of his dedication was to contribute to student development.

Passion for teaching was also closely connected with students, because teachers' satisfaction or enjoyment came from students. For example:

> I think my biggest satisfaction does not come from how many papers or books I have published. It comes from my students' acknowledgement.
>
> (Professor I, DepHum)

The reason why Professor I was passionate about teaching was not because he was only capable of teaching: it was because he truly enjoyed teaching and needed students' acknowledgement and appreciation. This was where he found satisfaction and enjoyment.

There was also a case of a teacher who did not like teaching in the first place but who changed as a result of the enthusiasm of her students. Associate Professor L (DepSc) was forced to give up her dream of being a doctor because her family could not afford the tuition fees. She studied at a teacher education university because this was free. However, her perception soon changed after she started to teach:

> The first students that I taught were really nice. They treated me as their sister. I began to feel really sorry that I taught so badly, so I tried to improve my teaching.
>
> (Associate Professor L, DepSc)

Associate Professor L's story shows that students can be a crucial factor in developing a teachers' passion for teaching.

The sense of responsibility was often related to students, because it was a feeling of not letting students down. For example, Lecturer K (DepHum) said 'When thinking about why those students come to attend my class, I feel that I have to teach responsibly'. In addition, the sense of responsibility could be promoted by maintaining a good image among colleagues and employers. Professor H's (DepHum) sense of responsibility was promoted by the concern regarding his image among his colleagues by thinking 'if my colleagues can teach well, why can't I do as well as my colleagues?' Lecturer L's (DepSc) sense of responsibility was motivated by her concern of the image of her graduates among employers. She wanted students who graduated from City University to maintain a good reputation, so she was trying her best to teach her students.

7.8.3 What Helps Teachers to Improve Teaching?

In terms of what helped teachers to improve teaching, teachers at City University mentioned research, colleagues, students, and a wide range of knowledge. Research

was mainly mentioned by professors. Professor I (DepHum), Professor L (DepSoSc) and Head L (DepHum) considered that undertaking research inspired and enriched their teaching.

In addition to knowledge in the subject areas, many teachers emphasised a wide range of knowledge was important. For example, Professor I (DepHum) said 'sometimes in my course I talked about astrophysics, mathematics or biology', because he thought this could provoke students' thinking.

The help offered by colleagues was mentioned by Lecturer M (DepSc) and Associate Professor J (DepHum) at the time when they were new teachers. For instance, Lecturer M's experiences were described as follows:

> When I just started, the department allocated a mentor to me. I observed his classes several times, but I thought that his teaching was average and not very helpful. I went by myself to observe some teachers who had a good reputation in teaching … In the beginning I was worried those teachers might not want to help me, because they might worry that I could become a competitor…However, they were very willing to help me and shared many of their materials with me.
>
> (Lecturer M, DepSc)

This quote shows an interesting contrast between the help arranged by the department and the help found by Lecturer M (DeSc). Providing mentors for new teachers should be a very useful mechanism to help new teachers. However, the allocation was made in a very administrative manner, so the mentor allocated to Lecturer M was not very helpful. Lecturer M had to find a mentor by himself. It was the teachers' own initiative that made a difference. Furthermore, this quote also reflects a lack of confidence among teachers. Asking for help from each other should be the most common thing among colleagues, but Lecturer M had some concerns in this regard. It

turned out that his colleagues were very willing to help, so there was the culture of an academic community at City University.

Students were able to help teachers to improve teaching too. Associate Professor J (DepHum) acknowledged that her students helped her in two ways. First, students provided her feedback via direct communication and online student evaluation. Second, when some students were very eager to learn, they posed many questions and demanded more. Through exploring answers to their questions and meeting their demands, Associate Professor J also improved her teaching.

7.9 Chapter Conclusions

City University provided an example of what impact government policies could have when there was a big gap between the official requirements and the reality on the ground. When the government employed an *authority tool*, such as the UTE, it was very effective in generating *implementation actions* and *consequences*, as illustrated above. However, there was a bigger chance that these *implementation actions* would become overreactions or result in unethical conduct. The overreactions reported by the interviewed teachers were: forbidding drinking water in classrooms, patrolling outside classrooms, requesting both hardcopy and electronic versions of the same document and storing student assignments. City University intended to use overreaction in some aspects to compensate for inadequacies in other aspects. Examples of unethical conduct included the President's suggestion to manipulate data in order to meet the UTE requirements and the forging of documents, which was admitted by both heads and teachers. The manipulation of data can be illustrated by the discrepancy in student–teacher ratios reported in the university report submitted for the UTE inspection and in its yearbooks. According to the university report, the

student–teacher ratio remained at 16:1 from 2003 to 2006 (City University, 2006). However, according to data extracted from its yearbooks, the student–teacher ratio reached 38:1 in 2005 and 32:1 in 2006 (see Table 20).

When the government employed *incentive tools*, such as the Quality Project awards, these were ineffective at City University because there was little possibility that the teachers at City University could win awards even with substantial incentives. As illustrated above, City University did not win any national or provincial awards. The award granted by the City Council was not substantial enough to motivate the winner, let alone the awards granted by the university. Therefore, many interviewed teachers developed contempt for official awards.

Because the State Council allocated the *resources* based on universities' research performance, City University reoriented itself and became determined to promote its research performance. However, an improvement in research performance could not be realised solely by a change of university policy. It had to be realised by teachers. However, many teachers did not have the capabilities to or interest in research. As a result, they felt marginalised by both national and university policies. This could result that City University lost its strength in teaching without gaining its strength in research. Teachers had to rely on students to gain a sense of achievement, so there was a prevalent overconfidence in the reliability and validity of the student evaluation of teachers.

8 Conclusions

This chapter will present a synthesis of the impact of quality assurance policies and mechanisms, as perceived by the teachers interviewed at the three case study universities in this research, and highlight the significant findings of the research through comparisons. Firstly, a comparison between the three case study universities in this research will demonstrate that each quality assurance policy or mechanism is largely implemented in the same manner at the three universities, and in general the impact of each policy or mechanism is consistent across the three universities. However, there are a small number of differences in regard to the impact of each policy and mechanism and a three-way comparison of universities will reveal the factors that account for these differences. Secondly, the findings of this research and those of other empirical studies will be compared, with the aim of achieving an analytic generalisation. The comparison will show that the patterns of the impact of one type of quality assurance policy or mechanism identified by this research was similar to the impact identified by previous empirical studies. This support will the analytic generalisability of the Impact Framework proposed by the researcher.

The quality assurance policies and mechanisms discussed in this thesis are: the UTE – a national quality assurance exercise; the Quality Project awards for good teaching; and class observation and the student evaluation of teachers. These quality assurance policies and mechanisms can be defined based upon the classification of policy tools proposed by Schneider and Ingram (1990). The UTE is an *authority tool*; the Quality Project awards are an *incentive and capacity-building tool*; and class observation and the student evaluation of teachers are *authority and capacity-building tools*. The

impact of the UTE, the Quality Project awards, and class observation and student evaluation will be presented in turn. Building upon this, the complete Impact Framework will be proposed, presenting the impact of quality assurance *authority*, *incentive*, *capacity-building* and *learning tools* on teaching quality in the selected universities.

The Impact Framework articulates what the impact of quality assurance is. Impact comprises *implementation actions*, *interactions*, *consequences* and *reactions*. *Implementation actions* are actions taken by the *target groups*, i.e. teachers, or other actors to implement quality assurance policies. *Interactions* are actions that enable the results of the *implementation action* to reach the *target groups*. These *interactions* complete the feedback loop of the *implementation actions*. *Consequences* are the outcomes of *implementation actions* or *interactions*. *Reactions* are actions taken by the *target groups* in response to the *consequences*.

At the end of this chapter, the researcher will reflect on the way the whole project was conducted, in light of the criteria for evaluating qualitative research proposed by Yardley (2000).

8.1 The Impact of National Quality Assurance Exercises

The impact of the UTE comprises *implementation actions*, *consequences* and *reactions*. *Implementation actions* will be discussed first, followed by a discussion of *consequences* and *reactions*. Two comparisons – a comparison between the three case study universities and a comparison between the UTE and quality assurance exercises in other countries – will be woven through the discussion. The impact of quality assurance exercises is presented and compared in Table 25.

The Impact of National Quality Assurance Exerc					
	This research			Other empirical research	
	Uni985	Uni211	UniCity	Chinese literature	International literature
Implementation actions					
Producing standardised documents	√	√	√	√	√
Following standardised procedures	√	√	√	√	√
Consequences					
Time and effort wasted	√	√	√	√	√
Academic freedom interfered with	√	√	√	√	√
Ethics undermined		√	√	√	√
Standardisation of documentation			√	√	
Discipline maintained		√		√	√
Provision of impetus to ensure teaching quality	√				√
Reactions					
Reactions taken to ensure the teaching quality	√				√

Table 25 The impact of national quality assurance exercises

(√ indicates that impact is present)

8.1.1 Implementation Actions

As illustrated in Table 25, the *implementation actions* of quality assurance exercises include producing standardised documents and following standardised procedures. What distinguishes the UTE from its counterparts is the fact that the UTE combined the two actions together – producing a wide range of standardised documents by following standardised procedures. Documents ranged from curricula, syllabi and lesson plans to student examination papers (see Chapters Five, Six and Seven). Producing new documents but also reprocessing existing documents was required. This *implementation action* was consistently reported by the majority of teachers interviewed across the three case study universities of this research and also by a large number of teacher respondents at the four universities in the southeast of China in Lee, Huang and Zhong's (2012) study, as well as at the three higher education institutions in central, north and northwest of China researched in Liu's (2011) study.

The first *implementation action* – producing standardised documents – is also frequently reported in the international literature. For instance, staff with administrative roles were requested to fill quality assurance forms (Cheng, 2008), or to produce course files (Newton, 1999b), module descriptions (Hoecht, 2006), degree programmes (Laughton, 2003) or business plans (Jones and Darshi De Saram, 2005). The format of these documents was standardised but, unlike in the case of the UTE, it was not a requirement that standardised procedures be followed.

The researcher would argue that the production of documents is favoured in quality assurance policies for a variety of reasons. First, the use of documents is very

economical for policy-makers (although the opposite is the case for teachers): documents can provide ample information without requiring policy-makers to mobilise many *resources* (i.e. *time, human* and *financial resources*) to collect that information. Second, documents are easy to preserve: they can be kept for a very long period for consultation. Last, but perhaps most significantly, documents are commonly regarded as legitimate evidence: for instance, well-prepared lesson plans serve as evidence of well-prepared teaching.

In terms of the second *implementation action* – following standardised procedures – this research found that standardised procedures were designed by individual case study universities, rather than the UTE – with the exception of standardised procedures being followed in order to produce documents. These university-designed procedures varied in regard to the UTE indicators that they aimed to address and the amount of *resources* that they consumed. However, they shared a common feature: a focus on producing visible evidence to impress evaluators and the fact that they diverted *resources* away from teaching.

However, this does not mean that standardised procedures do not have the potential to contribute to the improvement of quality. Cases in which standardised procedures have the potential to contribute to the improvement of quality have been reported in other research. For example, Liu (2011) reported the practice of cross-checking student examination papers at one of her case study universities, as a positive impact of the UTE, and Brennan, Frederiks and Shah (1997) reported the application of standardised procedures to record and monitor student work at their case study university, as a positive impact of QAA's audit.

The distinctive feature of the impact of the UTE – the combination of producing standardised documents and following standardised procedures together – is a manifestation of Chinese policy-makers' understanding of quality, and of their approach to quality assurance. Comparing the quality assurance approaches adopted by the UTE in China, accreditation in the US and institutional audits in England, Huang (2007) distinguishes between the different approaches adopted by these policies. Accreditation in the US identifies certain goals and invites universities to use their own ways to demonstrate fulfilment of these goals. Institutional audits in England provide descriptions of areas that universities should address, leaving room for universities to adapt these to their own contexts. The UTE in China employs quantitative indicators and specifies detailed requirements that must be satisfied in order to produce unified results. This approach gives universities little flexibility in regard to tailoring the quality assurance policy to cater to their needs. The three countries' (US, England and China) quality assurance approaches reflect different understandings of quality, as distinguished by Harvey and Green (1993). Accreditation in the US reflects an understanding of 'quality as fitness for purpose' (Harvey and Green, 1993); institutional audits in England reflect an understanding of quality as conformation to standards (Harvey and Green, 1993); and the UTE in China reflects an understanding of quality as conformation to specifications (Harvey and Green, 1993).

8.1.2 Consequences and Reactions

As illustrated in Table 25, this research found that the *consequences* of national quality assurance exercises include the following: time and effort being wasted; interference with academic freedom; the undermining of ethics; the standardisation of

documentation; the maintenance of discipline; and the provision of impetus to ensure teaching quality. The *reaction* to the latter *consequence* (provision of impetus to ensure teaching quality) is spending extra time on lesson preparation.

The first *consequence*, time and effort being wasted, was mentioned by the majority of teachers interviewed across the three case study universities in this research. This *consequence* was the outcome of all *implementation actions* mentioned above. In other words, as time and effort were diverted away from the core activities of teaching towards taking these *implementation actions*, teachers considered that this was a waste of their time and effort.

The *consequence* of time and effort being wasted is also reported as an impact of the UTE by Lee, Huang and Zhong (2012) and is frequently reported to be an impact of other national quality assurance exercises in other countries (Brennan, Frederiks and Shah, 1997; Cheng, 2008; Haapakorpi, 2011; Hoecht, 2006; Jones and Darshi De Saram, 2005; Laughton, 2003; Minelli et al., 2006; Newton, 1999b).

The second *consequence* revealed by this research was interference with academic freedom. Again, this was an outcome of all *implementation actions* mentioned above. These *implementation actions* change teachers' behaviours by requesting them to do things differently. The majority of teachers interviewed considered that academic freedom was interfered with.

The *consequence* of interference with academic freedom is also reported as an impact of the UTE by Liu (2011) and is frequently reported as being an impact of other

national quality assurance exercises in other countries (Askling, 1997; Brennan, Frederiks and Shah, 1997; Jones and Darshi De Saram, 2005; Newton, 1999b). However, in addition to changing existing practices, Askling's (1997) study uncovers that the requirement that there be more transparency (i.e. disclosing existing practices in order to demonstrate accountability) can be regarded as an interference with academic freedom. However, none of the teachers interviewed in this research or in other studies on the impact of the UTE mentioned the issue of transparency.

This research finds that the first *consequence* (a waste of time and effort) and the second *consequence* (interference with academic freedom) are related. The comparison between the three case study universities reveals that the more respect for academic freedom there is, the less time and effort are wasted. As presented in Chapters Five, Six and Seven, the *implementation actions* designed by the three case study universities vary: those designed by the Project 985 University wasted the least amount of teachers' time and effort, and those designed by City University wasted the most. This variation coincides with the degree of respect for academic freedom nurtured by their *university institutional rules*, i.e. it relates to the university context. The Project 985 University respected academic freedom the most, while City University did so the least.

Cheng's (2008) research offers an explanation of the connection between respect for academic freedom and the wasting of time and effort. The favouring among the interviewees of the external examining system over other quality assurance policies and mechanisms indicates that the more respect there is for academic freedom the wider the range of actions and evidence that are accepted as quality indicators, and

thus the less disturbance there will be to academics' work – with the result of less waste of time and effort.

The third *consequence* is the undermining of ethics. This is an outcome of the *implementation action* of reprocessing existing documents in order to meet the UTE requirements. This *consequence* was reported by many teacher interviewees at the Project 211 University and City University, who considered the alteration of existing documents as constituting the forging of documents. Moreover, City University exhibited more unethical conduct than its counterparts – for instance, its president called for data to be manipulated in order to fulfil the UTE requirements on student-teacher ratios and expenditure on teaching. However, this *consequence* was not mentioned by the teachers interviewed at the Project 985 University because documents were reprocessed by administrative staff instead. At the case study universities in Lee, Huang and Zhong's (2012) and Liu's (2011) research, where teachers took part in reprocessing documents, many of the respondents also confirmed this *consequence*.

The researcher would argue that the unethical actions of forging documents or manipulating data are likely to occur when there is a lack of *resources*. After the UTE issued the new requirements regarding the formats of documents, it did not give universities sufficient *time* to accumulate documents naturally. City University lacked *time* and the *financial* and *legal resources* needed to recruit more teachers and generate funds to spend on teaching. Therefore, if a university were provided with sufficient resources, this type of unethical action could be avoided.

The *consequence* of ethics being undermined is recorded in the international literature too (Brennan, Frederiks and Shah, 1997; Cheng, 2008; Hoecht, 2006; Laughton, 2003; Minelli et al., 2006), but the types of actions that cause this *consequence* are different to those reported in the Chinese literature. Whilst the unethical actions reported in the Chinese literature involve altering the presentation of what has been done, the unethical actions recorded in the international literature involve partially presenting what has been done (Cheng, 2008; Hoecht, 2006) and adjusting current activities to address quality indicators, with the aim of achieving good results (Brennan, Frederiks and Shah, 1997; Laughton, 2003; Minelli et al., 2006).

The researcher would argue that those actions recorded in the international literature are likely to be taken when universities are faced with national quality assurance exercises, and thereby the *consequence* of ethics being undermined is difficult to avoid in these cases. It is understandable that universities are motivated to achieve good results in national quality assurance exercises, because good results can bring them *resources* and reputational advantages. When they try to present themselves in the best possible light by highlighting their strengths and underplaying their weaknesses, universities fall into the trap of engaging in the unethical conduct of partially presenting what has been done. When they make preparations to address quality indicators, universities fall into the trap of engaging in the unethical conduct of adjusting current activities to address quality indicators. These actions seem to be logical reactions to high-stakes quality assurance exercises but they are regarded as unethical conduct by some academics, as reported in the international literature.

The first three *consequences* – time and effort being wasted, interference with academic freedom and the undermining of ethics – are negative impacts of quality assurance exercises. These negative impacts fostered negative attitudes towards quality assurance exercises among the majority of teacher interviewees at the three case study universities. These negative attitudes were manifested in the form of irritation in this research, of frustration in Hoecht's (2006) and Newton's (1999b) research, of demotivation in Hoecht's (2006) and Liu's (2011) research, and of stress, anxiety, cynicism, resistance and resentment in Brennan, Frederiks and Shah's (1997) research. As demonstrated in the following sections, this research finds that negative attitudes cause contempt for other quality assurance policies or mechanisms, which demotivates teachers from participating in them.

The fourth *consequence* is the standardisation of documentation. This is an outcome of the *implementation action* of producing standardised documents. This research finds that only teacher interviewees from City University considered this to be a positive *consequence* of the UTE. Similarly, in her study on the impact of the UTE, Liu (2011) observed that almost all respondents from a university college regarded this *consequence* as being positive, whereas only one or two of their counterparts from the other two, more prestigious, case study universities shared this perception. These findings suggest that teachers from the lower end of the spectrum of institutions tend to perceive the standardisation of documentation as a positive *consequence*. Without having the protection that is offered by an established reputation, these teachers may resort to the standardisation of documentation as evidence to demonstrate the fulfilment of accountability.

In the international literature, none of the studies on the impact of national quality assurance exercises mention the standardisation of documentation as a *consequence*. Instead, the focus is on maintaining a consistent national standard. For example, some respondents acknowledged the QAA's contribution to maintaining a consistent standard across the nation (Cheng, 2008; Hoecht, 2006).

The fifth *consequence* – the maintenance of discipline – can be considered as a positive impact. It is an outcome of the *implementation action* of following standardised procedures and *action plans* aimed at disciplining teachers. This positive *consequence* was only mentioned by one teacher respondent at the Project 211 University. None of the teacher interviewees at the Project 985 University or City University mentioned it. This may be attributed to *university institutional rules*, i.e. university contexts. As mentioned in previous chapters, one of the Project 985 University's *university institutional rules* is to promote academic freedom and autonomy. Hence, discipline should be maintained by the academic community, instead of by external measures. At City University, the extreme measures taken by the university to reinforce discipline, such as banning drinking water in classrooms and patrolling outside classrooms, might have fostered a distaste for discipline which prevented the interviewed teachers from mentioning this *consequence.*

Other empirical studies on the impact of the UTE also report the maintenance of discipline as a positive impact (Lee, Huang and Zhong, 2012; Liu, 2011). Another empirical study on the impact of a quality assurance exercise in Hong Kong also reported a similar finding (Jones and Darshi De Saram, 2005). Such impact was not reported in the international literature. The more favourable attitudes towards

discipline found in Chinese and Hongkongese studies may be attributed to the Chinese culture, which values discipline as a virtue.

The sixth *consequence* is the provision of impetus to ensure teaching quality. This is an outcome of the prospect of being observed by evaluators. One teacher interviewee at the Project 985 University in this research acknowledged this impact of the UTE. The *reaction* to this *consequence* was that this interviewee spent extra time on lesson preparation. None of the teachers interviewed at the other two case study universities in this research, nor any respondents in other empirical studies on the impact of the UTE, mentioned this *consequence.* In the international literature, some respondents at a pre-1992 university in Cheng's study (2008) and at the University of Amsterdam in Minelli et al.'s (2006) study confirmed that their national quality assurance exercises provided them with an impetus to ensure teaching quality. Minelli et al. (2006, 117) reported the *reactions* following this *consequence* as being 'a set of improvement actions stimulating changes in teaching practices'.

These three universities (the Project 985 University, the pre-1992 university and the University of Amsterdam) where the provision of impetus to ensure teaching quality was reported as a *consequence* shared some similarities. At these three universities, fewer *implementation actions* were prescribed. As a result, less time and effort were wasted and academic freedom was less interfered with. When the impacts of national quality exercises were less negative, teachers formed less negative attitudes towards them. It seems that if the attitudes that teachers hold are less negative, it is more likely that they will acknowledge the positive role that quality assurance exercises play in ensuring teaching quality.

8.2 The Impact of Awards for Good Teaching

The comprehensive list of intended purposes or outcomes of teaching awards summarised by Dinham and Scott (2002) shows that teaching awards carry high expectations. If these intended purposes or outcomes are interpreted by using the framework of policy tools proposed by Schneider and Ingram (1990), it can be concluded that teaching awards are expected to act as *incentive* and *capacity-building tools*, retrospectively and prospectively. In other words, teaching awards are supposed to offer acknowledgement of previous achievements (i.e. the retrospective function of *incentive tools*) and to incentivise non-award winners to achieve more in the future (i.e. the prospective function of *incentive tools*); as well as to enable reflection on previous achievements (i.e. the retrospective function of *capacity-building tools*) and to enable the sharing of information between award winners and non-award winners (i.e. the prospective function of *capacity-building tools*).

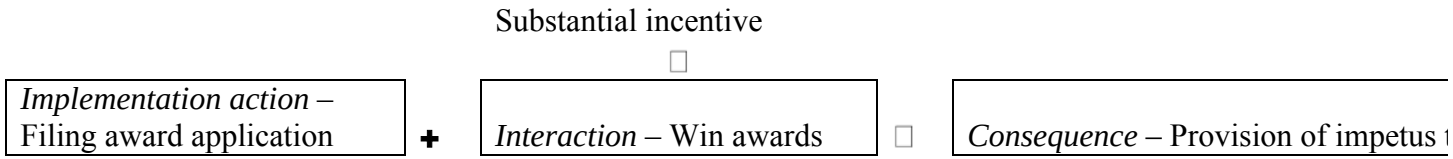
As mentioned in Chapter Two, empirical evidence collected by Dinham and Scott (2002) from the Australian and US education sectors and by Vaden-Kiernan, Michie and Frechtling (1994) from US schools, shows that teaching awards were able to offer acknowledgement of previous achievements, i.e. having the retrospective function of *incentive tools*.

However, whether teaching awards are able to achieve the prospective function of *incentive tools,* or the retrospective and prospective functions of *capacity-building tools,* is open to question. Jacobsen (1989), and Tollefson and Tracy (1983) find that there was no significant difference between award winners' and non-award winners' teaching effectiveness in their case study higher education institutions in the US.

However, this finding can be interpreted in two different ways. One interpretation is that non-award winners were motivated by award winners so they improved their teaching effectiveness. The other interpretation is that the teaching of non-award winners was as effective as that of award winners, but, due to the limited number of awards, non-award winners did not receive awards. The first interpretation implies that teaching awards incentivised non-award winners to achieve more or enabled the sharing of award winners' experience with non-award winners or both, whereas the second does not. Neither of these interpretations is tested empirically. This deficit in the existing research is addressed by the present research.

Table 26 illustrates the impact of the Quality Project awards as *incentive* and *capacity-building tools*. The conditions that enable the impact to occur are also indicated in the tables.

The Impact of Teaching Awards as *Incentive Tools*



The Impact of Teaching Awards as *Capacity-building Tools*

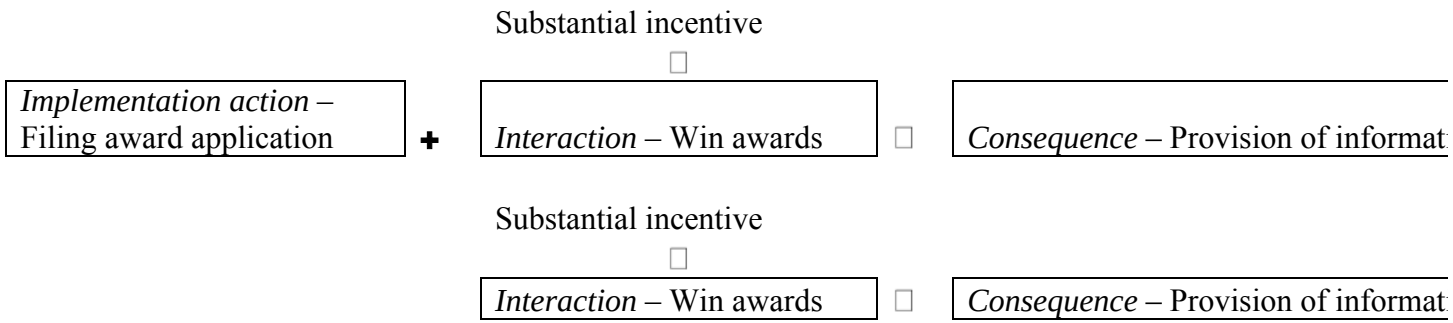


Table 26 The impact of teaching awards as incentive and capacity-building tools

8.2.1 The Impact of Awards as Incentive Tools

This research identifies a general pattern of award winners' and non-award winners' perceptions of the impact of Quality Project awards in the three case study universities. In general, award winners perceived the impact of the awards as being the provision of impetus through acknowledging their previous achievements, whereas non-award winners did not perceive any impact. In other words, as *incentive tools,* the awards realised their retrospective function but not their prospective function. This finding is consistent with Dinham and Scott's (2002) and Vaden-Kiernan, Michie and Frechtling's (1994) conclusions. However, this research uncovers cases where award winners did not perceive this impact. These cases point to a crucial condition that needs to be satisfied for this impact to be realised – award winners have to consider the incentive offered by awards to be substantial, which Dinham and Scott (2002) and Vaden-Kiernan, Michie and Frechtling (1994) did not report as a factor.

Not only does this research reveal the conditions necessary for the impact to be realised, it also expands our understanding of impact. The term 'impact', as used by interviewees in this research, and by Dinham and Scott (2002) and Vaden-Kiernan, Michie and Frechtling (1994), is limited to the *consequence* of awards – which is one area of impact identified by the researcher. As illustrated in Table 26, the impact of teaching awards includes the *implementation action* (filing an award application), the *interaction* (winning awards), and the *consequence* (the provision of impetus through acknowledgement). Furthermore, in addition to substantial incentives, the empirical

evidence collected by this research shows that the *implementation action* and the *interaction* are also necessary conditions for the *consequence* to occur.

Interaction is a necessary condition for the *consequence* to occur because without it the consequence will not take place. An award applicant who took the *implementation action* but did not engage in the *interaction*, testified that the Quality Project awards were not able either to incentivise him to learn from the award winners, or to work harder and apply next time. Instead, he attributed his unsuccessful application to the problem of reliability. This shows that the awards failed to realise the prospective function of *incentive tools*.

Empirical evidence collected by this research enables the researcher to summarise the reasons why the teaching awards failed to have an impact. The reasons include: insubstantial incentives; contempt; the problems of reliability and validity; and lack of awareness.

Insubstantial incentives is the first and foremost reason. It is observed that at the three case study universities the insubstantial incentives offered by the Quality Project awards failed to motivate a teacher to take the *implementation action*, and also prevented some award winners from perceiving the *consequence*.

The incentives offered by the Quality Project awards are insubstantial when they are compared with those offered by research. Incentives can be monetary and honorary. The analysis of national quality assurance policies in Chapter Four showed that the monetary incentives offered by research outweigh those offered by teaching, because

there are more government departments offering monetary incentives to research and the value of those monetary incentives is higher than those offered for teaching. The substantial monetary incentives offered by research in turn reinforce the honorary incentives attached to research.

As presented in Chapters Five, Six and Seven, the analysis of university policies showed that the three case study universities' teacher appraisal policies vary in regard to the amount of incentives for research, on top of those incentives provided by national policies. The Project 985 University did not add extra incentives to encourage research. The Project 211 University's quantitative performance-based appraisal allocated more points to research and allowed teachers to substitute teaching for research. City University offered bonuses for research. The Project 211 University's appraisal policy created the widest the gap between teaching and research incentives among the three universities. This explains why the complaint of insubstantial incentives was voiced by a teacher at the Project 211 University, who had a good prospect of receiving research incentives.

Contempt is the second factor that demotivated many teachers in regard to taking the *implementation action* at the Project 211 University and at City University. Contempt refers to having a prejudice against participation in optional quality assurance policies and mechanisms, such as the Quality Project awards. Contempt stems from the negative attitudes generated by the compulsory implementation of quality assurance policies (examples of these negative attitudes are given in Section 8.1.2). The phenomenon of contempt demonstrates that quality assurance policies or mechanisms can influence one another, and thus a holistic approach is needed for policy-making in

order to avoid different quality assurance policies or mechanisms cancelling out each other's influence.

The case of the Project 985 University, where none of the teachers interviewed expressed contempt for the Quality Project awards, illustrates the elements that minimised negative attitudes. These factors are the same as those that encouraged the teacher at this university to acknowledge the positive *consequence* of the UTE: its *university institutional rule* of promoting academic freedom and autonomy enabled the university to design *implementation actions* in such a way as to reduce the negative *consequences* of the UTE – the waste of time and effort and interference with academic freedom; consequently, less negative *consequences* caused less negative attitudes, which generated less contempt.

The problems of reliability and validity are the third factor that demotivated some teacher interviewees in this research in regard to taking the *implementation action*. Before taking the *implementation action*, teachers would evaluate the reliability and validity of awards in order to judge their chance of winning an award. If the chances were high, they would apply, if not they would not.

In this study, more teachers voiced concerns about validity than about reliability. They complained that the Quality Project awards were awarded based on research performance instead of teaching performance. Most of these teachers were from City University. As illustrated in Tables 13, 18 and 23, teachers from City University received the fewest Quality Project awards among the three case study universities. This fact serves as evidence for Lan's (2006) and Wang's (2010) claim that the

majority of indicators of the Quality Project awards measure applicants' research performance, rather than their teaching performance. This validity issue frustrated teachers from City University more than their counterparts because it was difficult for them to obtain acknowledgement from research. The Quality Project awards raised their hopes of receiving acknowledgement of their teaching performance, but they turned out to be rewards for research performance.

Only one teacher interviewed complained about the reliability of the Quality Project awards. This interviewee called for including front-line teachers and students on the panel of judges deciding on the award winners, which mainly comprised senior managers. The fact that fewer teachers pointed out issues to do with reliability may be attributed to the lack of transparency in the awards process. The process by which the award winners are selected is not publically available, so teachers are less aware of reliability issues.

Lack of awareness is the fourth reason for the lack of impact. This research finds that lack of awareness prevented some teachers interviewed at the Project 985 University from taking *implementation action*, because the university did not publicise the Quality Project awards well enough.

8.2.2 The Impact of Awards as Capacity-building Tools

Teaching awards as *capacity-building tools* are overlooked by existing empirical studies due to the lack of a framework that makes it possible to draw a distinction between *incentive* and *capacity-building tools*. This research is able to address this deficiency.

According to Schneider and Ingram's (1990) framework, *incentive* and *capacity-building tools* are based on different presumptions. *Incentive tools* assume that *target groups* (i.e. teachers) lack the motivation to change their behaviour, so *incentive tools* provide *target groups* with incentives to motivate them. On the other hand, *capacity-building tools* assume that *target groups* have been motivated but they do not have the capacity to make changes, so *capacity-building tools* help *target groups* to develop their capacity. Therefore, the discussion of teaching awards as *capacity-building tools* is based upon the assumption that *target groups* have been motivated by the substantial incentives provided by the awards to improve their capacity.

As presented in Table 26, the impact of the Quality Project awards as *capacity-building tools* includes the *implementation action* (filing an award application), the *interaction* (winning awards,) and the *consequence* (the provision of information through reflection or good examples). This research uncovers that, provided that teachers are motivated by substantial incentives, the *implementation action* and the *interaction* are necessary conditions for the provision of information through reflection *consequence* (i.e. the retrospective function) to occur; only the *interaction* is a necessary condition for the provision of information through good examples *consequence* (the prospective function) to occur. Consequently, it can be concluded that the Quality Project awards were able to fulfil the retrospective and prospective functions of *capacity-building tools.*

This research also uncovers the factors that prevented the Quality Project awards from having an impact as *capacity-building tools*, which are: the problems of validity and reliability; contempt; and lack of awareness. Those teachers who were interviewed

who doubted the validity of the awards did not think they were able to derive useful information from award winners to help them to improve teaching. Those teacher interviewees who had suspicions regarding the reliability of the awards did not consider the award winners as deserving of the awards. Contempt for the Quality Project awards was fuelled by doubts over their validity and reliability, and also by the negative attitudes generated by the UTE. Finally, for those respondents that were not aware of the awards, the awards did not have an impact on them.

8.3 The Impact of Class Observation

By adopting Schneider and Ingram's (1990) framework to analyse the impact of class observation collected by this research and by other empirical research in the international literature, the researcher finds that class observation has the function of *authority and capacity-building tools*. Acting as different policy tools, they create different impacts. The new perspective introduced by policy tools casts a new light on our understanding of the co-existence of the positive and the negative impacts of class observation.

The impact of class observation as an *authority tool* will be discussed first, followed by a discussion of its impact as a *capacity-building tool*. The two comparisons – the comparison between the three case study universities and the comparison between this research and other empirical studies in the international literature – will be woven through the discussion. As presented in Table 27 and Table 28, the impact of class observation includes *implementation actions*, *interactions* and *consequences*.

The Impact of Class Observation as an *Authority Tool*				
	This research			Other empirical research
	Uni985	Uni211	UniCity	International literatur
Implementation actions				
Observing class				√
Organising observation				√
Attending training sessions				√
Writing observation reports				√
Reporting reflection				√
Interaction				
Meeting for discussion				√
Consequence				
Provision of impetus to maintain discipline	√	√		
Provision of impetus to ensure teaching quality			√	
Time and effort wasted				√
Academic freedom interfered with				√
Ethics undermined				√

Table 27 The impact of class observation as an authority tool

(√ indicates that impact is present)

The Impact of Class Observation as a *Capacity-building Tool*				
	This research			Other empirical research
	Uni985	Uni211	UniCity	International literatur
Implementation actions				
Attending training sessions				√
Interaction				
Receiving feedback	√	√	√	
Meeting for discussion	√			√
Consequence				
Provision of valid information through reliable feedback	√	√	√	√

Table 28 The impact of class observation as a capacity-building tool

(√ indicates that impact is present)

8.3.1 The Impact of Class Observation as an Authority Tool

As illustrated in Table 27, this research reveals that the perceived impact of class observation on teachers at the three case study universities involves two positive *consequences* – the provision of impetus to maintain discipline and the provision of impetus to ensure teaching quality – in the absence of teachers' participation in *implementation action* and *interaction.* No other empirical research in the literature has reported this finding. The researcher would argue that this unique finding results from some special features of the *implementation actions* designed by the three case study universities.

The three special features of *implementation actions* of class observation at the three case study universities can be drawn from the department heads' and teachers' accounts presented in Chapters Five, Six and Seven. Firstly, class observation was often conducted by inspectors appointed by a university or department, or the heads of department. Secondly, individual teachers' results were not confidential. The inspectors shared teachers' results with the heads of department, who had the responsibility for discussing the problems revealed in observation with the observed teachers. Thirdly, class observation was conducted more frequently than in the case of other studies in the international literature. The first two features effectively raised the stakes associated with class observation. The first feature meant that the results of class observation were an official judgement on teachers' teaching abilities; the second feature affected teachers' images in the minds of department leaders. The third feature decided which impetus was given – the impetus to maintain discipline or to ensure teaching quality.

This research finds that the frequency of class observation determines the type of *consequence*. When class observation is conducted very frequently, the *consequence* is the provision of the impetus to ensure teaching quality; when class observation is conducted frequently, the *consequence* is the provision of the impetus to maintain discipline. It is observed that at City University, where class observation was conducted very frequently – more often than the other two case study universities – for certain teachers the prospect of being observed became a certainty. Those teachers tended to be inexperienced teachers and teachers who received complaints. When facing this certainty, they were motivated to invest more time and effort to ensure teaching quality in order to secure a good official judgement and image. At the Project 985 University and the Project 211 University, where class observation was conducted less frequently, for some teachers the prospect of being observed was a possibility. This possibility motivated some teachers to discipline themselves, for example by being punctual, in order to avoid the risk of being caught engaging in misconduct.

Empirical studies in the international literature report some impacts of class observation that this research fails to capture. The discrepancies are caused by the differences in regard to *implementation actions* and *interaction*. As highlighted by Table 27, at the three case study universities in this research, teachers were not involved in taking *implementation actions or* in *interaction*. In contrast, teachers at the case study universities in other empirical research had to take a number of the *implementation actions*: organising observation (Bell, 2001), attending training sessions (Blackwell and McLean, 1996), and reporting reflection (Bell, 2001). For

those teachers who observed each other's class, they had to undertake extra work: conducting observations and writing reports (Hammersley-Fletcher and Orsmond, 2005; Shortland, 2004). Teachers were actively involved in the *interaction* – meetings with observers to discuss aims and needs before observation, and feedback after observation (Bell, 2001; Blackmore, 2005; Blackwell and McLean, 1996; Hammersley-Fletcher and Orsmond, 2005; Shortland, 2004).

The compulsory nature of class observation as an *authority tool* forced teachers to take part in these *implementation actions* and in the *interaction*. For those teachers who were unwilling to, these *implementation actions* and the *interaction* caused three negative *consequences*, as illustrated in Table 27. Forced to undertake tasks that deviated from their needs, teachers considered that their time and effort were wasted (the first negative *consequence*), and their academic freedom was interfered with (the second negative *consequence*) (Bell, 2001; Blackmore, 2005; Blackwell and McLean, 1996; Cockburn, 2005; Douglas and Douglas, 2006). Consequently, unethical conduct emerged, such as completing observation documents without observing classes (Shortland, 2004), which led to the third negative *consequence*, the undermining of ethics. These three negative *consequences* are the same as those caused by the UTE, another *authority tool*.

8.3.2 The Impact of Class Observation as a Capacity-building Tool

As presented in Table 28, teacher interviewees' accounts collected by this study show that class observation as a *capacity-building tool* includes *implementation actions, interactions* and *consequence*. While *implementation action* are optional, the *interactions* – receiving feedback and meeting for discussion – are necessary

conditions for class observation as a *capacity-building tool* to generate the *consequence* – the provision of valid information through reliable feedback.

Implementation actions are absent from the impact of class observation as a *capacity-building tool* at the three case study universities in this research because these actions were taken by other actors. This proves that *implementation actions* are not a necessary condition for having a *consequence*. However, other empirical studies in the international literature show that there are *implementation actions* that involve teachers taking action, such as attending training sessions (Bell and Cooper, 2013; Lomas and Nicholls, 2005).

As illustrated in Table 28, the *interactions* can take different forms. The most common *interaction* at the three case study universities is receiving feedback without discussion. The form of *interaction* most often reported by other empirical studies in the international literature is meeting with observers to discuss aims and needs before observation and feedback after observation (Bell and Cooper, 2013; Lomas and Nicholls, 2005). The Humanities Department at the Project 985 University is an exception: here, *interaction* took the form of meeting for discussion.

The *consequence* of the provision of valid information through reliable feedback is reported across the three case study universities. This *consequence* is supported and manifested by other empirical studies in the international literature: for instance, offering confirmation or a different perspective and identifying strengths and weaknesses (Bell, 2001; Bell and Cooper, 2013; Blackmore, 2005; Blackwell and

McLean, 1996; Cockburn, 2005; Douglas and Douglas, 2006; Hammersley-Fletcher and Orsmond, 2005; Lomas and Nicholls, 2005; Shortland, 2004).

Again, according to Schneider and Ingram's (1990) framework, the discussion of class observation as a *capacity-building tool* is based upon the assumption that teachers have had the motivation to improve their capacity. Therefore, the foremost reason accounted for the lack of impact when class observation acts as a *capacity-building tool* is that teachers are not motivated to improve their capacity. Moreover, teacher interviewees' accounts collected by this research reveal more reasons for a lack of impact: the problem of reliability and interference with academic freedom.

The problem of reliability is the first reason for a lack of impact. Across the three case study universities, many teacher interviewees' concerns about reliability were related to university inspectors, due to their lack of subject knowledge. The problem of reliability is endorsed by other empirical studies, but there it is referred to in a different sense. Instead of doubt regarding those doing the observation having sufficient subject knowledge, the reliability of peer observers is doubted because of their lack of seniority (Cockburn, 2005) or honesty (Bell, 2001; Blackmore, 2005; Blackwell and McLean, 1996; Douglas and Douglas, 2006).

Interference with academic freedom is the second reason for a lack of impact. There were two cases of resistance to class observation reported at the Project 985 University due to this reason. As stated earlier, this is attributed to *its university institutional rule*: to promote academic freedom and autonomy. In the international

literature, Shortland (2004) also records a case of resistance to class observation due to interference with academic freedom.

Other empirical studies in the international literature add one extra reason for the lack of impact – the problem of validity, in relation to the 'Hawthorn Effect': the presence of observers affects the observed (Blackmore, 2005; Cockburn, 2005; Douglas and Douglas, 2006). Interestingly, none of the teachers interviewed in this study voiced concerns about the validity of class observation.

8.4 The Impact of the Student Evaluation of Teachers

By introducing Schneider and Ingram's (1990) framework into the analysis of the impact of the student evaluation of teachers collected by this research and by other empirical research in the international literature, the researcher finds that student evaluations of teachers function as *authority tools and as capacity-building tools*. When functioning as different policy tools, they generate different impacts. The introduction of policy tools facilitates our understanding of the co-existence of the positive and the negative impacts of student evaluation.

The impact of the student evaluations of teachers found at the three case study universities in the present research is summarised in Table 29 and Table 30. The impact of student evaluations reported by other empirical studies in the international literature is also presented in in Table 29 and Table 30, and illustrates that in general the findings of other empirical studies agree with those of this research.

The Impact of Student Evaluation as an *Authority Tool*				
	This research			Other empirical research
	Uni985	Uni211	UniCity	International literatur
Interaction				
Obtaining of student feedback		√	√	√
Consequence				
Provision of impetus to ensure teaching quality	√	√	√	√
Provision of impetus to please students	√	√	√	√
Reactions				
Changes in teaching activities				√

Table 29 The impact of student evaluations as an authority tool

(√ indicates that impact is present)

The Impact of Student Evaluations as a *Capacity-building Tool*				
	This research			Other empirical research
	Uni985	Uni211	UniCity	International literatu
Interaction				
Obtaining of student feedback		√	√	√
Consequence				
Provision of valid information through reliable feedback		√	√	√
Reactions				
Changes in teaching activities				√

Table 30 The impact of student evaluations as a capacity-building tool

(√ indicates that impact is present)

8.4.1 The Impact of the Student Evaluation of Teachers as an Authority Tool

Based on the empirical data collected by this research, the researcher concludes that the impact of the student evaluation of teachers as an *authority tool* constitutes an *interaction* (obtaining student feedback) and two simultaneous *consequences* (the provision of the impetus to ensure teaching quality and the provision of the impetus to please students), in the absence of any *implementation action*. As illustrated in Table 29, an *interaction* is not a necessary condition for *consequences* to occur. *Implementation actions* are absent because this activity is undertaken by students.

Furthermore, this research reveals the features of that group of teachers who perceive student evaluations as having impact: in other words, those whose behaviour is likely to be influenced by such evaluations. By comparing the perceptions of the teachers interviewed at the three case study universities, this research finds that the groups of teachers who perceived that student evaluations had an impact were those who believed that their results in student evaluations would affect their promotion or those who were afraid that bad results in student evaluations would damage their reputation.

Other empirical research collected evidence of *reactions* following two simultaneous *consequences* – the provision of the impetus to ensure teaching quality and the provision of the impetus to please students, which serves as a validation of the *consequences* reported by this research. Ryan, Anderson and Birchler (1980) find that a very high percentage (80%) of survey respondents reported changes in their behaviour in their case study university, where student evaluations functioned as an *authority tool* (the student evaluations are compulsory and are consulted when making

personnel decisions). The most frequently reported changes were: providing increased course content structure and supporting information that could facilitate learning (a *reaction* to improve teaching quality), and lowering the level of difficulty of course content and exams (a *reaction* to please students).

The co-existence of these two *consequences* can be explained by Nasser and Fresko's (2002) finding that, on the one hand, many teachers believe there is a relationship between the quality of teaching and the results of student evaluations; and on the other hand, they also believe that other factors affect the results. Moreover, Ryan, Anderson and Birchler's (1980) research showed that the majority of their respondents believed their results in student evaluations could be improved; and more respondents (33% more) thought that the results were more likely to be improved by addressing areas that affect student attitudes, rather than solely addressing instructional activities.

8.4.2 The Impact of the Student Evaluation of Teachers as a Capacity-building Tool

When student evaluation acts as a *capacity-building tool*, the present research concludes, based on empirical evidence collected at the three case study universities, that its impact comprises the *interaction* (the obtaining of student feedback), and the *consequence* (the provision of valid information through reliable feedback), see in Table 30.

As illustrated in Table 30, the comparison of the three case study universities reveals that *interaction* is a necessary condition for *consequence* to occur when the student evaluation of teachers acts as a *capacity-building tool,* because the *consequence* was

reported only at the Project 211 University and City University where *interaction* was present, i.e. teachers had access to the results of student evaluations.

Other empirical studies stress the importance of *interaction.* Baxter (1991) and Murray and Newby (1982) reported that a very high percentage of respondents (86% and 78%, respectively) regarded the obtaining of student feedback as the major reason for participating in student evaluations.

Moreover, other empirical research has collected evidence of *reactions* following the *consequence* (the provision of valid information through reliable feedback), which serves as a validation of the *consequence* reported by this research. At Baxter's (1991) case study university, where student evaluations were designed to be a *capacity-building tool* (being voluntary and having the aim to provide confidential feedback to improve teaching and learning), the overwhelming majority of respondents (91%) changed their teaching in some way based on student feedback. At Murray and Newby's (1982) case study university, where student evaluations were more of a *capacity-building tool* but had some features of an *authority tool*, there was still a decent percentage of teachers (33.5%) who utilised the diagnostic feedback provided by student evaluations to improve their teaching.

By extracting data from the accounts of teachers interviewed at the three universities, this research is able to present the factors that account for the lack of impact, besides a lack of motivation – an inherent limit for capacity-building tools. The factors revealed by this research are: the problem of reliability and validity; lack of access; and contempt.

The first factor is the problem of reliability, which was raised by many teachers interviewed at the Project 985 University and at the Project 211 University. These teachers had no confidence in the reliability of student evaluations because they doubted students' abilities to conduct evaluations seriously and objectively, but none of their counterparts at City University lacked this confidence, because student evaluations there were a source of acknowledgement.

As elaborated in Chapter Two: Literature Review, many empirical studies have proved the reliability of student evaluations. However, the findings of this research show that doubts regarding the reliability of student evaluations still persist among many teachers. The researcher argues that two factors account for this persistence. Firstly, doubt regarding reliability can be caused by a lack of knowledge. Franklin and Theall (1989) find that many users of student ratings are not aware of current issues relating to student evaluations in the literature. Secondly, the persistent doubt can also be caused by the difference between an individual case and statistics. In statistics, one single case is not significant. However, when a teacher's results in a student evaluation suffer from grading leniency, workload/difficulty or the Dr. Fox effect, this single case serves as a confirmation of the unreliability of student evaluations for the individual teacher. Despite the persistent doubt regarding reliability, the majority of respondents (71% and 72%, respectively) at Baxter's (1991) case study university considered that students rated teachers in a serious manner and that students' ratings were accurate.

The second factor is the problem of validity, which was pointed out by a few teacher interviewees at the Project 985 University and at the Project 211 University. These teachers claimed that the design of the questionnaire failed to measure the core activities of teaching and learning. However, none of their counterparts at City University mentioned any issue about validity. Similarly, in Murray and Newby's (1982) research, the majority of respondents (74%) thought that the student rating form provided valid information.

Lack of access is the third factor. This was only present at the Project 985 University because it was the only case study university where teachers did not have access to the results of student evaluations.

Contempt is the fourth factor. Contempt for student evaluations was only mentioned by some teachers interviewed at the Project 211 University. In contrast, their counterparts at the Project 985 University were not aware of the results of the student evaluations, and their counterparts at City University had a very positive view towards student evaluations, given that they relied on students to gain acknowledgement. Teachers at the Project 211 University who showed contempt for student evaluations were angry about their colleagues' behaviour in regard to pleasing students and considered that their behaviour undermined their professional ethics. Hence, the contempt is seen to have been caused by a *consequence* of student evaluations as an *authority tool* – the provision of the impetus to please students.

8.5 The Impact Framework

The findings presented above enable the researcher to complete the proposed Impact Framework, articulating the impact of quality assurance *authority, incentive, capacity-building* and *learning tools*. As illustrated above, some impacts emerge consistently across different manifestations of the same quality assurance policy tool. Therefore, these impacts can be identified as the impacts associated with the particular quality assurance policy tool. The Impact Framework extracts these consistent impacts and presents them in Table 31.

The Impact of *Authority Tools*

Implementation actions	□	*Consequences* Time and effort wasted Academic freedom interfered with Ethics undermined
		Consequence Provision of impetus

The Impact of *Incentive Tools*

Interaction	□	*Consequence* Provision of impetus

The Impact of *Capacity-building Tools*

Interaction	□	*Consequence* Provision of information

The Impact of *Learning Tools*

Interactions	□	Consequence Provision of information Provision of impetus

Table 31 Impact Framework

8.5.1 The Impact of Quality Assurance Authority Tools

Quality assurance *authority tools* are compulsory quality assurance policies and mechanisms that have high stakes. The high stakes, often related to *resources*, reputation or career progression, provide very strong motivation for *target groups*, i.e. teachers, to carry out *implementation actions*. Even when *implementation actions* are not taken by *target groups*, they have *consequences* on *target groups*. Compared with the rest of quality assurance policy tools, which require the satisfaction of certain conditions to have impact, quality assurance *authority tools* do not require any condition to be satisfied in order to have impact. Once *authority tools* are employed, they will have impact. Therefore, *authority tools* are the most effective quality assurance policy tools.

The present research studied three manifestations of quality assurance *authority tools*: the UTE, class observation and the student evaluation of teachers. The consistent impacts that emerged from these tools are presented in Table 31, including *implementation actions* and *consequences*.

The content of *implementation actions* varies policy by policy. They are designed by policy-makers aimed at achieving good results. The presence of *implementation actions* makes a difference to the *consequences* that arise. When *implementation actions* are present, the following negative *consequences* arise: a waste of time and effort, interference with academic freedom and the undermining of ethics. When *implementation actions* are absent, the following positive *consequence* arise: the provision of the impetus to ensure teaching quality.

Once *implementation actions* are taken, three negative *consequences* are very likely to occur. Firstly, taking *implementation actions* consumes *resources*, including *time, human resources* and *financial resources*. Secondly, the designated *implementation actions* often intend to change *target groups*' behaviour, given that quality assurance policies are often introduced as interventions. The changes introduced by quality assurance policies are often regarded as a disturbance by *target groups*. When *target groups* are forced to expend *resources* on disturbances, the first and the second negative *consequences* (i.e. a waste of time and effort and interference with academic freedom) emerge. Thirdly, there is a tendency for *target groups* to take actions to present themselves in the best possible light or to make preparation for an inspection in order to achieve good results. This tendency leads to the third negative *consequence*, i.e. the undermining of ethics.

The duration of the two negative *consequences* is longer than that of the positive *consequence*. The provision of the impetus to ensure teaching quality is generated by the prospect of teaching being observed by evaluators directly, so it can only last for the duration of the evaluators' visit. The *consequences* of wasting of time and effort and interference with academic freedom can last after the inspection period if universities adopt as their internal quality assurance policies the *implementation actions* that cause these *consequences*.

Moreover, the three negative *consequences* prevent other quality assurance tools having impact. As pointed out earlier, the three negative *consequences* generate negative attitudes towards quality assurance policy. These negative attitudes cannot affect whether or not *target groups* take *implementation actions* designated by quality

assurance *authority tools*, due to their compulsory nature, but they cause *target groups*' contempt for other types of quality assurance policy tools, where participation is optional. Hence, the effectiveness of quality assurance *authority tools* is achieved at the cost of other quality assurance tools.

8.5.2 The Impact of Quality Assurance Incentive Tools

Quality assurance *incentive tools* are quality assurance policies and mechanisms that use monetary or honorary incentives to motivate *target groups*, i.e. teachers, to improve the quality of teaching. The manifestation of *incentive tools* studied by this research is teaching awards.

As illustrated by Table 31, the impact of quality assurance *incentive tools* comprises *interaction* (the receipt of incentives) and a *consequence* (the provision of impetus through acknowledgement). The *interaction* leads to the *consequence*. As mentioned earlier, there are many expectations attached to teaching awards, but this research demonstrates that awards can only fulfil their retrospective function, which is to offer acknowledgement of previous achievements.

Furthermore, this research reveals that in order to for the *interaction* to generate the *consequence* a condition must be fulfilled. The condition is that the incentive has to be considered by *target groups* to be substantial. As the 'substantiality' of incentives is a relative concept, the incentives offered by competing quality assurance policies and mechanisms are compared by *target groups* when making decisions regarding which one to take part in, given that there is a limit in terms of the time and energy that *target groups* can devote to the policies and mechanisms. This requires policy-

makers to adopt a holistic approach in order to allocate the most substantial incentives to the most desirable goal.

8.5.3 The Impact of Quality Assurance Capacity-building Tools

Quality assurance *capacity-building tools* are quality assurance policies and mechanisms that help *target groups* to build their capacity in order to improve the quality of teaching. Participation in quality assurance *capacity-building tools* is usually optional, but even it is compulsory the stakes associated with *capacity-building tools* are not high, so *target groups* are able to ignore these quality assurance policies and mechanisms.

The manifestations of quality assurance *capacity-building tools* studied by this research are teaching awards, class observation and the student evaluation of teachers. The common impacts shared by these are: an *interaction* (receiving feedback), and a *consequence* (the provision of information) (see Table 31). The *interaction* is a necessary condition for the *consequence* to occur.

Quality assurance *capacity-building tools* face two critical challenges in regard to having an impact: a lack of motivation and the negative influence of quality assurance *authority tools*. As explained before, quality assurance *capacity-building tools* rely on a precondition in order to have impact: the *target groups* must have been motivated to improve their capacity. When investigating the unaddressed tensions and factors of existing quality assurance policies and mechanisms at the three case study universities (see Sections 5.8, 6.8 and 7.8), the researcher uncovered that the interviewed department heads and teachers were motivated by themselves, colleagues or students, and very rarely they were motivated by quality assurance policies and mechanisms.

However, the motivation generated by the academic community is likely to be cancelled out by the contempt caused by quality assurance *authority tools*.

8.5.4 The Impact of Quality Assurance Learning Tools

None of the existing quality assurance exercises reviewed by this research employs *learning tools*. However, the interviewees' reflections on what drove and helped them to improve their teaching throughout their career indicate that learning tools can make a substantial contribution towards assuring the quality of teaching (see Sections 5.8, 6.8 and 7.8). These accounts collected by this research enable the researcher to propose the idea of quality assurance learning tools.

In the light of Schneider and Ingram's (1990) classification, the following definition of quality assurance *learning tools* is proposed: quality assurance *learning tools* do not dictate *target groups*' behaviours. When employing *learning tools*, policy-makers, *target groups* and other actors assume that they do not know how to assure quality, so they have to explore and learn together.

As illustrated by Table 31, the impact of *learning tools* consists of *interactions* and *consequences*. Without predetermined *action plans* made by policy-makers, *learning tools* do not need *implementation actions* to have impact. *Learning tools* require actors' spontaneous participation. Spontaneous participation occurs when actors are interested to learn, where there is a supportive culture and where they have enough *resources*.

The interviewed department heads' and teachers' accounts suggest that having *interactions* between various actors is crucial because *interactions* generate *consequences*.

Interactions in various forms can spontaneously occur when actors communicate. Through *interactions*, *target groups* can receive information to help them to improve (the *consequence* of the provision of information) or the impetus to motivate them to improve (the *consequence* of the provision of impetus).

8.6 Policy Implications

The strengths and weaknesses of quality assurance *authority, incentive, capacity-building* and *learning tools* are summarised in Table 32. This section will discuss the implications of the strengths and weaknesses inherent in different policy tools in relation to policy-making.

	Strengths	**Weaknesses**
Authority tools	Effective in making impact	Consumption of *time, financial and*
	Provision of impetus to ensure teaching quality during inspection	Interference with academic freedom
		Undermining of ethics
		Cause contempt for quality assuranc
Incentive tools	Provision of impetus through offering acknowledgement of previous achievements	Require substantial incentives to hav
		Unable to influence those who do nc incentives
Capacity-building tools	Provision of information to improve teaching	Uncertain whether will have impact
		Unable to provide motivation
		Affected by contempt
		Subject to validity and reliability
Learning tools	Provision of information to improve teaching	Uncertain whether will have impact
	Provision of impetus to improve teaching	Unable to provide motivation
		Affected by contempt
		Subject to validity and reliability

Table 32 The strengths and weaknesses of policy tools

The most important policy implication drawn from this research is that policy-makers should adopt a holistic approach to policy-making by setting priorities and weighing the cost of realising the impact of these priorities. This may sound simple but quite often policy-makers fail to realise it in practice. It is a common practice that the quality of research and the quality of teaching are divided and assured by different bodies in many countries, such as China and the UK. The Chinese case studied by this research illustrates some of the major problems caused by the division of responsibilities between different quality assurance bodies and the lack of coordination between different quality assurance policies issued by the same quality assurance body. They are: different quality assurance policies cancel out each other's influence; *resources* are wasted on promoting activities peripheral to the quality of education; and the negative *consequences* caused by quality assurance policies outweigh the positive ones.

Quality assurance *authority tools* should be allocated the top priority because they are the most effective policy tools. Once they are issued, they will have impact in the form of both *implementation actions* and *consequences*, or *consequences* only. However, quality assurance *authority tools* consume the most *resources*, including from *time*, *human* and *financial resources*. They will divert *resources* to this priority, while sacrificing non- or low- priority tools.

There is a high risk that *resources* will be directed towards producing standardised documents and following standardised procedures, while *resources* for the core activities of teaching will be reduced, due to the nature of documents and the understanding of quality as conformation to specifications, as demonstrated by this

research and other empirical studies in the literature. Admittedly, *authority tools* can lead to a positive consequence – the provision of the impetus to ensure teaching quality – but this *consequence* is not an outcome of the *implementation actions* taken by *target groups* – it results from teachers facing the prospect of being inspected. Once the inspection is finished or is conducted infrequently, the *consequence* will no longer continue to exist. To sustain the *consequence*, more *resources* have to be invested to carry out inspections. Policy-makers have to assess gains and costs, not only in regard to the consumption of *resources* but also the negative *consequences* associated with *authority tools*.

Due to the negative *consequences*, quality assurance *authority tools* are likely to cause contempt for quality assurance policies. This jeopardises other quality assurance tools, especially quality assurance *capacity-building* and *learning tools*, because to have impact these require that *target groups* are motivated. The strength of *capacity-building* and *learning tools* lies in their ability to generate direct impact on promoting the improvement of teaching quality without consuming too many *resources*. These are the exact weaknesses of *authority tools* – consuming a great number of *resources* but promoting activities peripheral to quality improvement and causing many negative *consequences*. Therefore, policy-makers have to weigh whether or not the gain of employing *authority tools* is worth the cost of other policy tools not having impact.

A holistic approach should be applied to the policy-making of quality assurance *incentive tools* too. As analysed above, *target groups* compare the incentives offered by different quality assurance policies and select the ones that offer the most substantial incentives to invest their time and energy. Therefore, policy-makers should

allocate incentives according to the importance of priorities. Moreover, because *incentive tools* can only have impact on those who receive the incentives, policy-makers should avoid focusing on incentivising the same *target groups* repeatedly. Instead, policy-makers should design different incentives for different *target groups,* in order to have wider impact.

The drawback of quality assurance *capacity-building* and *learning tools* is that they do not have a guaranteed to have impact. To have impact, they rely on many conditions being satisfied: motivation, absence of contempt, validity and reliability. However, they have a great potential to contribute to the improvement of quality because they have direct impact on teaching, by providing information and impetus to *target groups*. In order to increase the likelihood that *capacity-building* and *learning tools* will have impact, policy-makers should design quality assurance *authority tools* in such a way as to minimise the negative *consequences* and thereby the contempt for quality assurance policies. Given the importance of *interactions* for *capacity-building* and *learning tools*, policy-makers should create an environment which increases the opportunities for *target groups* and other actors to interact.

8.7 Reflections on this Project

This section provides reflections on this project in light of the criteria for evaluating qualitative research proposed by Yardley (2000), which has been presented in Chapter Three (see Section 3.2). Following the categories used by Yardley, the reflections are divided into: sensitivity to context, commitment and rigour, transparency and coherence, and impact and importance.

8.7.1 Sensitivity to Context

In order to orient my research within the literature on the impact of quality assurance policies, an extensive search for relevant literature was conducted. I started by conducting a key words search in databases such as Education Resources Information Centre (ERIC), Scopus and Web of Science. Disappointingly, relevant search results were very limited. The strategy was then changed to browsing related journals. The journals that regularly publish articles on quality assurance in higher education are: *Quality in Higher Education, Quality Assurance in Education, Assessment & Evaluation in Higher Education, Higher Education, Higher Education Policy* and *Journal of Higher Education Policy and Management*. This strategy yielded fruitful results. Many relevant articles were found. Furthermore, the references contained in these articles led to more relevant publications. Besides journal articles, Copac was used to search theses which have researched similar topics and several relevant theses were identified. Their references also revealed more relevant research. This extensive search enabled the literature review to cover as many publications as possible.

After reviewing these publications, three problems in the existing literature emerged. Firstly, there was a lack of empirical research on the impact of quality assurance policies, particularly concerning Chinese quality assurance policies. For example, although there were several empirical studies of the impact of the UTE, no empirical research on the impact of the Quality Project was found. Secondly, the existing research did not articulate what exactly the impact was. Thirdly, none of the research asked whether different policy tools would create particular patterns of impact. Intrigued by these deficiencies, I decided to dedicate my research to exploring the

answers to these questions and to proposing a model to articulate the impact of different policy tools.

My working experience as a lecturer at a university in China enabled me to gain a good understanding of the sociocultural background in Chinese higher education. More importantly, it made the establishment of trust between the participants and myself easier because they considered me to be a member of the academic community. Establishing trust was what I endeavoured to achieve during data collection. My strategy to achieve it was being frank. I was very frank about the information regarding my research and myself. Moreover, I was very considerate of the participants: during the interviews, I always tried my best to create a relaxing atmosphere. When the participants changed their plans without notifying me beforehand, I always showed my understanding. As promised to my participants, I took extreme care to protect their identities. When some participants asked about other participants, I always politely refused to answer their questions. When reporting data or citing documents published by the case study universities, I replaced the universities, the departments and the participants' real names with generic names.

To familiarise myself with the data, I transcribed all interviews in full. In order to preserve the information as much as possible, the interviews were transcribed and analysed in their original language. Quotes were then translated into English. The principle adopted for translation was: to try to use the same words as long as they made sense in English; if not, to replace them with English expressions having the closest meanings.

8.7.2 Commitment and Rigour

Quality assurance is not a topic that is currently in fashion, but it is of interest to me and I am committed to it. Although some definitions of quality assurance exclude improvements produced by quality assurance, I am convinced that improvement should be an indispensable part of any quality assurance exercise, and in fact, the ultimate goal of such an exercise. Otherwise, it would be a waste of resources to conduct quality assurance. In searching for answers to the question 'How can we make the exercise of quality assurance contribute to improvement?', I started by analysing the national policy of the UTE (Huang, 2007). Naturally, the next step was to investigate how the national policy is implemented and what impact it has. However, I did not proceed in this way. Instead, I chose to receive a systematic training in educational research methodologies. This training gave me the freedom to choose the most appropriate methodology to pursue my research interest. By the time I started this project, several years had passed since the UTE, and the Quality Project had also finished. I thought it would therefore be a good opportunity to explore what impact they had made up to then.

The guiding principle for the research design of the project was maximising its potential. When selecting universities for case studies, I did not select universities that I could access easily. Instead, I chose universities based on the their suitability for the research design. The same criteria applied to the selection of departments for case studies. Without a personal connection, it was difficult to recruit participants and access university documents. However, after I exhausted all the means that I could possibly use, I managed to recruit more participants than the original design and to photocopy the three case study universities' yearbooks. To understand the case study

universities better, I read all these yearbooks and the policy documents published on their websites. Even when some documents could not be accessed outside the university's domain, I found a way to access them through its intranet. Moreover, I tried to stay at their campuses as much as I could, to develop a deeper understanding of the universities.

Although the focus of this research is the impact of quality assurance policies, the scope of this research covers policy-making and policy implementation at institutional and individual levels. By doing so, it locates quality assurance polices in a context and reveals the roots of the problems, which are: the central government defines quality as efficiency and only uses *authority tools* and *incentive tools* for quality assurance purposes without considering their inherent flaws.

Building upon the empirical evidence, this research documents the impact of existing quality assurance policies. In light of the documented impact, I identify patterns regarding the impact and put forward an Impact Framework intending to lift the impact studies to a theoretic level.

8.7.3 Transparency and Coherence

I wanted my research to be as transparent as possible without revealing participants' identities. As elaborated in Chapter Three, I have shared how I conducted this research with readers. When reporting findings, I also wanted to share as much as possible, so there was a tendency to use long quotes and provide too many details. Overuse of long quotes and unnecessary details were two problems frequently pointed out by my supervisor and co-supervisor. After many revisions, I have reduced the number of long quotes and many unnecessary details. However, as readers may have

seen, there are still many long quotes. They are left in because they can give readers an authentic taste of what interviewees said and their emotion and opinions conveyed by their words. I continue to believe that interviewees' original words are more powerful than my paraphrasing of them.

Making clear and coherent arguments was my aim, but this is not easy. Sometimes I had too many thoughts and it was difficult to remain clear; sometimes I thought that I had explained something clearly but I had not. At those moments, I asked for fresh perspectives. My supervisor and co-supervisor were extremely helpful in providing critical comments. My colleagues also offered their help by reading and commenting on my text. After many revisions, I hope readers find the arguments made in this thesis to be clear and coherent, even though I acknowledge that they might not be clear for everybody –we all have different ways of thinking.

As mentioned in Chapter Three, one of the weaknesses of this research relates to generalisability. I have to admit that the claims that I made regarding a department are only based on the accounts of six or seven heads or teachers at that department and the claims regarding a university are only based on the accounts of 18 or 20 heads or teachers at that university. Whether these claims are able to capture every single teacher's opinions requires further investigation. However, due to the constraints as regards resources and the voluntary nature of the research, it is questionable if any research is able to achieve this ambition. Despite this weakness, the contribution of my research is the conceptualisation of a framework articulating the factors underpinning impact and the relationships between these factors. More specifically, the factors are *implementation actions, reactions, interactions* and *consequences*. The

difference that a more representative sample could make would be finding further expressions of these factors. Even if future research developed a new model, this framework could serve as a stepping stone towards the theorisation of impact studies.

8.7.4 Impact and Importance

The impact and importance of this research are evident in its contributions. Firstly, this research addressed the lack of empirical research on the impact of quality assurance policies and mechanisms on teaching in the Chinese context, especially the impact of the Quality Project awards, class observation and the student evaluation of teachers, which is rarely studied empirically. Secondly, by employing Knoepfel et al.'s (2007; 2011) framework to analyse the quality assurance policies at a national and university level, this research revealed the problems in policy-making in Chinese higher education – the failure of national policies to address the problems raised by the public and a shift of focus from assuring quality to achieving good results in university policy-making. Thirdly, by analysing several quality assurance policies, this research was able to observe the influence between quality assurance policies. Fourthly, in the light of Schneider and Ingram's (1990) framework of policy tools, this research articulated what the impacts of quality assurance policy are and provided definitions of quality assurance policy tools. Fifthly, this study proposed an Impact Framework, which reveals patterns of the impact of various quality assurance policy tools. The analytic generalisability of the Impact Framework is supported by other empirical research in the Chinese and international literature. By presenting the strengths and weaknesses associated with different quality assurance tools, the Impact Framework is able to help policy-makers to make informed decisions by pointing out the potential gains and costs. It can also help practitioners to understand the impact of various quality assurance policy tools on them, and to make informed decisions

regarding how to react. The compilation of this Impact Framework represents a step towards bridging the gap between the expected impact of quality assurance policies and the actual impact on the ground.

References

Abrami, P.C., Leventhal, L. and Perry, R.P. (1982) Educational Seduction. *Review of Educational Research,* 52 (3), pp. 446-464.

Al Qahtani, S., Kattan, T., Al Harbi, K. and Seefeldt, M. (2011) Some Thoughts on Educational Peer Evaluation. *South-East Asian Journal of Medical Education,* 5 (1), pp. 47-49.

An, X. (1997) 高等教育质量的界定初探 [Preliminary Definition of Quality of Higher Education]. *Liaoning Higher Education Research,* 2, pp. 9-11.

Anonymous (2012) *中国大学985 工程三期拨款清单 [A List of Project 985 Funding by Universities]*. Sina. [Online]. Available from: <http://blog.sina.com.cn/s/blog_5557acef0101bw77.html> (Accessed Oct/8, 2014).

Anonymous (2009) *城市大学：降分录取是误解 [City University: Lower admission standards are misinterpretation]*. City: City Evening News.

Anonymous (1999) 今天教育扩招可拉动经济增量 70 亿 [The Increase of Participation in Higher Education Can Generate ¥7 Billion Economic Growth]. *Leader Decision Making Information,* 37, pp. 29.

Askling, B. (1997) Quality Monitoring as an Institutional Enterprise. *Quality in Higher Education,* 3 (1), pp. 17-26.

Barandiaran-Galdós, M., Barrenetxea-Ayesta, M., Cardona-Rodríguez, A., Mijangos-Del-Campo, J.J. and Olaskoaga-Larrauri, J. (2012) Attitudes of Spanish University Teaching Staff to Quality in Education. *Journal of Higher Education Policy and Management,* 34 (6), pp. 647-658.

Barnett, R. (1990) *The Idea of Higher Education.* Buckingham: The Society for Research into Higher Education & Open University Press.

Bauer, M. (1996) Quality as Expressed in a Swedish Reform of Higher Education and as Viewed by University Teachers and Leadership. *Tertiary Education & Management,* 2 (1), pp. 76-85.

Bauer, M. and Henkel, M. (1997) Responses of Academe to Quality Reforms in Higher Education: A comparative study of England and Sweden. *Tertiary Education & Management,* 3 (3), pp. 211-228.

Baxter, E.P. (1991) The TEVAL Experience, 1983–88: The impact of a student evaluation of teaching scheme on university teachers. *Studies in Higher Education,* 16 (2), pp. 151-178.

Bell, A. and Mladenovic, R. (2008) The Benefits of Peer Observation of Teaching for Tutor Development. *Higher Education,* 55 (6), pp. 735-752.

Bell, M. (2001) Supported Reflective Practice: A programme of peer observation and feedback for academic teaching development. *International Journal for Academic Development,* 6 (1), pp. 29-39.

Bell, M. and Cooper, P. (2013) Peer Observation of Teaching in University Departments: A framework for implementation. *International Journal for Academic Development,* 18 (1), pp. 60-73.

Blackmore, J.A. (2005) A Critical Evaluation of Peer Review via Teaching Observation within Higher Education. *International Journal of Educational Management,* 19 (3), pp. 218-232.

Blackwell, R. and McLean, M. (1996) Peer Observation of Teaching & Staff Development. *Higher Education Quarterly,* 50 (2), pp. 156-171.

Boyatzis, R.E. (1998) *Transforming Qualitative Information: Thematic analysis and code development.* London: Sage.

Brennan, J., Frederiks, M. and Shah, T. (1997) *Improving the Quality of Education: The impact of quality assessment on institutions.* Bristol: Higher Education Funding Council for England.

Brennan, J. and Shah, T. (2000) *Managing Quality in Higher Education: An international perspective on institutional assessment and change.* Buckingham: OECD, SRHE and Open University Press.

Britt, N. (1982) *Faculty Attitudes about Colleague Evaluation of Teaching.* Doctor of Education. Indiana University.

Brown, R. (2004) *Quality Assurance in Higher Education: The UK experience since 1992.* London: Routledge.

Brunetto, Y. and Farr-Wharton, R. (2005) Academics' Responses to the Implementation of a Quality Agenda. *Quality in Higher Education,* 11 (2), pp. 161-180.

Bryman, A. (2008) *Social Research Methods.* Third ed. Oxford: Oxford University Press.

Cardoso, S., Santiago, R. and Sarrico, C.S. (2012) The Impact of Quality Assessment in Universities: Portuguese students' perceptions. *Journal of Higher Education Policy and Management,* 34 (2), pp. 125-138.

Centra, J.A. (2003) Will Teachers Receive Higher Student Evaluations by Giving Higher Grades and Less Course Work? *Research in Higher Education,* 44 (5), pp. 495-518.

Centra, J.A. and Creech, F.R. (1976) *The Relationship between Student, Teacher, and Course Characteristics and Student Ratings of Teacher Effectiveness.* Princeton: Educational Testing Service.

Chen, Y., Dai, R., Yang, X. and Tian, S. (2004) *高等教育质量保障体系概论 [Introduction to Quality Assurance System of Higher Education].* Beijing: Beijing Normal University Press.

Cheng, M. (2011) 'Transforming the learner' versus 'passing the exam': Understanding the gap between academic and student definitions of quality. *Quality in Higher Education,* 17 (1), pp. 3-17.

Cheng, M. (2010) Audit Cultures and Quality Assurance Mechanisms in England: A study of their perceived impact on the work of academics. *Teaching in Higher Education,* 15 (3), pp. 259-271.

Cheng, M. (2008) *The Perceived Impacts of Quality Audit on the Work of Academics - A case study of a research-intensive Pre-1992 University in England.* PhD. University of Bristol.

City University (2006) *城市大学自评报告 [City University Report].* City: City University.

City University (2005) *科研工作奖励管理条例 [Research Rewarding Regulations].* City: City University.

City University (2002) *教职工教学科研管理奖励暂行规定 [The Allowance Allocation Scheme].* City: City University.

Cockburn, J. (2005) Perspectives and Politics of Classroom Observation. *Research in post-compulsory education,* 10 (3), pp. 373-388.

Cohen, P.A. (1981) Student Ratings of Instruction and Student Achievement: A meta-analysis of multisection validity studies. *Review of Educational Research,* 51 (3), pp. 281-309.

Cohen, P.A. (1980) Effectiveness of Student-rating Feedback for Improving College Instruction: A meta-analysis of findings. *Research in Higher Education,* 13 (4), pp. 321-341.

CPC Central Committee (1985) *中共中央关于经济体制改革的决定 [The CPC Central Committee's Resolution on the Reform of the Economic System].* Beijing: CPC Central Committee.

CPC Central Committee (1984) *中共中央关于教育体制改革的决定 [The CPC Central Committee Decision on the Reform of the Education System].* Beijing: CPC Central Committee.

CPC Central Committee and State Council (1993) *中国教育改革和发展纲要 [China Education Reform and Development Program].* Beijing: CPC Central Committee and State Council.

Di, F. (2000) *扩招让人欢喜让人忧 [The Increase of Participation in Higher Education Makes people happy and worry].* Beijing: China Education Daily.

Dinham, S. and Scott, C. (2002) *Awards for Teaching Excellence: Intentions and Realities.* Annual Meeting of the Australian Association for Research in Education. (Australia, December the Annual Meeting of the Australian Association for Research in Education, 1-5 December, Brisbane.

Douglas, J. and Douglas, A. (2006) Evaluating Teaching Quality. *Quality in Higher education,* 12 (1), pp. 3-13.

Du, J. and Zeng, D. (2007) 高等教育外部质量保障体系闭环系统初探 [Pilot Research on Closed Loop System of External Quality Assurance System in Higher Education]. *Higher Education Development and Evaluation,* 23 (1), pp. 54-59.

Dunstan, K. (2008) Review of Peter Knoepfel, Corinne Larrue, Frederic Varone, and Michael Hill 'Public Policy Analysis. *Journal of Social Policy,* 37 (3), pp. 517-518.

Elton, L. (1992a) Quality Enhancement and Academic Professionalism. *The New Academic,* 1 (2), pp. 3-5.

Elton, L. (1992b) *University Teaching: A professional model for quality and excellence.* 'Quality by Degrees' Conference, 8 June, Aston University.

Feldman, K.A. (1997) Identifying Exemplary Teachers and Teaching: Evidence from student ratings. In *Effective Teaching in Higher Education: Research and Practice.*, eds. R.P. Perry andJ.C. Smart, New York: Agathon, pp. 368-395.

Feldman, K.A. (1976) Grades and College Students' Evaluations of Their Courses and Teachers. *Research in Higher Education,* 4 (1), pp. 69-111.

Feng, J. (2008) *高校：怎样评估更有效 [Higher Education: How to evaluate would be more effective].* Beijing: Guangming Daily.

Filippakou, O. (2009) *The Legitimation of Quality in Higher Education.* PhD. University of London.

Filippakou, O. and Tapper, T. (2008) Quality Assurance and Quality Enhancement in Higher Education: Contested territories? *Higher Education Quarterly,* 62 (1-2), pp. 84-100.

Fisher, A.T., Alder, J.G. and Avasalu, M.W. (1998) Lecturing Performance Appraisal Criteria: Staff and student differences. *Australian Journal of Education,* 42 (2), pp. 153-168.

Franklin, J. and Theall, M. (1989) Who Reads Ratings: Knowledge, Attitude, and Practice of Users of Student Ratings of Instruction. *In: Annual Meeting of the American Education Research Association, 27-31 March 1989, San Francisco.*Boston: Northeastern University, pp. 1-29.

Frederiks, M.M.H., Westerheijden, D.F. and Weusthof, P.J.M. (1994) Effects of Quality Assessment in Dutch Higher Education. *European Journal of Education,* 29 (2), pp. 181-199.

Gao, H. (2000) *高校扩招: 并不轻松的话题 [The Increase of Participation in Higher Education: Not an easy topic].* Shenyang: Liaoning Daily.

Gao, Y., Zhang, P., Chen, H., Lan, L. and Zhang, G. (2006) 本科教学工作水平评估对高校教学工作影响的调查研究 [The Impact of the Evaluation of the Standards of Undergraduate Teaching Work on Teaching at Higher Education Institutions: An investigative study]. *Chinese Education and Society,* 42 (2), pp. 7-29.

Gerbic, P. and Kranenburg, I. (2003) The Impact of External Approval Processes on Programme Development. *Quality in Higher Education,* 9 (2), pp. 169-177.

Gift, S.I. and Bell-Hutchinson, C. (2007) Quality Assurance and the Imperatives for Improved Student Experiences in Higher Education: The case of the university of the West Indies. *Quality in Higher Education,* 13 (2), pp. 145-157.

Goldschmid, M.L. (1978) The Evaluation and Improvement of Teaching in Higher Education. *Higher Education,* 7 (2), pp. 221-245.

Gu, G. (2007) 本科院校教学质量保障体系的构建与运行探索 [On the Establishment and Implementation of the Undergraduate Teaching Quality Assurance]. *Higher Education Exploration,* 1, pp. 136-136.

Gu, M. (ed.)(1998) *教育大辞典 [Education Dictionary]*. Shanghai: Shanghai Education Publishing House.

Guan, H., Cai, Y. and Gong, Y. (2011) 我国高校教学质量监控研究综述 [Review of the Teaching Quality Monitoring]. *Heilongjiang Education,* 2, pp. 27-30.

Guba, E.G. and Lincoln, Y.S. (1989) *Fourth Generation Evaluation.* London: Sage.

Ha, X. (2009) 高等教育质量观的发展综述 [Summary of the Development of the Definition of Quality of Higher Education]. *Intelligence,* 36, pp. 211.

Haapakorpi, A. (2011) Quality Assurance Processes in Finnish Universities: Direct and indirect outcomes and organisational conditions. *Quality in Higher Education,* 17 (1), pp. 69-81.

Hammersley-Fletcher, L. and Orsmond, P. (2005) Reflecting on Reflective Practices within Peer Observation. *Studies in Higher Education,* 30 (2), pp. 213-224.

Han, Y. (2003) *高等教育质量精细分析 [A Fine Analysis of Quality of Higher Education]*. PhD. East China Normal University.

Hanson, J. (1993) Observing classroom teaching in higher education: A case study. *Quality Assurance in Education,* 1 (3), pp. 26-30.

Harvey, L. and Green, D. (1993) Defining Quality. *Assessment and Evaluation in Higher Education,* 18 (1), pp. 9-34.

Harvey, L. and Knight, P.T. (1996) *Transforming Higher Education.* Buckingham: SRHE and Open University Press.

Harvey, L. and Newton, J. (2007) Transforming Qualilty Evaluation: Moving on. In *Quality Assurance in Higher Education.*, eds. D. Westerheijden S. Stensaker and M. Rosa,Dordrecht: Springer, pp. 225-245.

Harvey, L. and Williams, J. (2010) Fifteen Years of Quality in Higher Education (Part Two). *Quality in Higher Education,* 16 (2), pp. 81-113.

He, F. (2003) *研究生扩招能否保证质量 [Can Graduate Education Retain its Quality after the Increase in Enrollment]*. Beijing: Xinhua Daily Telegraph.

Hodson, P. and Thomas, H. (2003) Quality Assurance in Higher Education: Fit for the new millennium or simply year 2000 compliant? *Higher education,* 45 (3), pp. 375-387.

Hoecht, A. (2006) Quality Assurance in UK Higher Education: Issues of trust, control, professional autonomy and accountability. *Higher Education,* 51 (4), pp. 541-563.

Horsburgh, M. (1999) Quality Monitoring in Higher Education: The impact on student learning. *Quality in higher education,* 5 (1), pp. 9-25.

Howard, G.S., Conway, C.G. and Maxwell, S.E. (1985) Construct Validity of Measures of College Teaching Effectiveness. *Journal of Educational Psychology,* 77 (2), pp. 187-196.

Hu, P. (2008) 我国高校教学评估二十年发展历程回顾 [History and Actualities of the Evaluation of Chinese Higher Education]. *Journal of Higher Education Research,* 17, pp. 40-43.

Huang, S. (2007) *Quality Assessment in Higher Education: Tensions between Unity and Diversity of Evaluation Standards in China.* MEd. University of Bristol.

Huang, J. (2002) 高等教育转型期的质量观及其评价 [On the Quality of Higher Education in its Transition]. *Modern University Education,* 3, pp. 12-14.

Jacobs, L.C. (1987) University Faculty and Students' Opinions of Student Ratings. *Indiana Studies in Higher Education,* 55, pp. 1-19.

Jacobsen, R.H. (1989) *The Impact of Faculty Incentive Grants on Teaching Effectiveness.* Annual Meeting of the American Educational Research Associate, 27-31 March, San Francisco.

Ji, B. (2008) *大学评估太多了 [Too Many Evaluations].* Beijing: People's Daily.

Jia, H. (2003) 高等教育质量保障与评价机制建设 [Higher Education Quality Assurance and Evaluation System]. *Higher Education Exploration,* 1, pp. 18-21.

Jones, J. and Darshi De Saram, D. (2005) Academic Staff Views of Quality Systems for Teaching and Learning: A Hong Kong case study. *Quality in Higher Education,* 11 (1), pp. 47-58.

Kang, N. (2000) 论教育决策与制度创新：以 99 高校扩招政策为案例的研究 [On Education Decision Making and Innovation: A case study of 1999 Policy to Increase Participation in Higher Education. *Journal of Higher Education,* 2, pp. 31-38.

Keig, L. and Waggoner, M.D. (1994) *Collaborative Peer Review: The role of faculty in improving college teaching. ASHE-ERIC Higher Education Report No. 2.* Washington, D.C.: George Washington University.

Kell, C. and Annetts, S. (2009) Peer Review of Teaching Embedded Practice or Policy-holding Complacency? *Innovations in Education and Teaching International,* 46 (1), pp. 61-70.

Knoepfel, P., Larrue, C., Varone, F. and Hill, M. (2011) *Public Policy Analysis.* Second ed. Bristol: The Policy Press.

Knoepfel, P., Larrue, C., Varone, F. and Hill, M. (2007) *Public Policy Analysis.* Bristol: The Policy Press.

Kvale, S. (1996) *InterViews: An introduction to qualitative research interviewing.* London: Sage.

Lan, J. (2006) 教学激励政策的反思对中国高等教育国家级教学成果奖的综合考察与分析 [Reflection on Teaching Encouragement Policy: A comprehensive analysis on China Higher Education National Level Teaching Achievement Award]. *Journal of National Academy of Education Administration,* 12, pp. 56-67.

Langfeldt, L., Harvey, L., Huisman, J., Westerheijden, D. and Stensaker, B. (2008) *Evaluation of NOKUT–the norwegian agency for quality assurance in education.* Oslo: Ministry of Education and Research.

Laughton, D. (2003) Why Was the QAA Approach to Teaching Quality Assessment Rejected by Academics in UK HE? *Assessment & Evaluation in Higher Education,* 28 (3), pp. 309-321.

Lee, J.C., Huang, Y.X. and Zhong, B. (2012) Friend or Foe: The impact of undergraduate teaching evaluation in China. *Higher Education Review,* 44 (2), pp. 5-25.

Li, J. (ed.)(2006) *质量管理学 [Quality Management].* Shanghai: East China Normal University Press.

Li, S. and Li, H. (2010) 我国本科教学评估回顾与反思 [Undergraduate Teaching Evaluation in China: A brief review]. *Higher Education Development and Evaluation,* 26 (6), pp. 47-54.

Li, Y., Li, X. and Qu, Q. (2006) 普通高校本科教学工作水平评估调查综述 [Reevaluation of and Prospects for Undergraduate Teaching and Learning Evaluation]. *China University Teaching,* 6, pp. 46-47.

Li, Z. (2010) 构建高校教学质量保障体系理论综述 [On Teaching Quality Assurance in Higher Education]. *Managers,* 12, pp. 85-85.

Liang, W., Zhang, G. and Guo, L. (2006) *骗取公众信任的高校评估风该煞！[Stop the University Evaluation that Betrayed the Public Trust].* Beijing: China Reform News.

Lincoln, Y.S. and Guba, E.G. (1985) *Naturalistic Inquiry.* London: Sage.

Liu, J. (2002) 高等教育的服务质量观 [On the Quality of the Service of Higher Education]. *China Higher Education Research,* 7, pp. 29-30.

Liu, S. (2013) Quality Assessment of Undergraduate Education in China: Impact on different universities. *Higher Education,* 66 (4), pp. 391-407.

Liu, S. (2011) *Impact of the Quality Assessment of Undergraduate Education on University Change in China.* PhD. Institute of Education, University of London.

Lomas, L. and Nicholls, G. (2005) Enhancing Teaching Quality Through Peer Review of Teaching. *Quality in Higher Education,* 11 (2), pp. 137-149.

Lv, N. and Zhang, Z. (2004) *高等教育: 在规模和质量之间怎样寻求平衡 [Higher Education: How to balance size and quality]*. Beijing: China Reform News.

Marsh, H.W. (2007) Students' Evaluations of University Teaching Dimensionality, Reliability, Validity, Potential Biases and Usefulness. In *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective.*, eds. R.P. Perry andJ.C. Smart, New York: Springer, pp. 319-383.

Marsh, H.W. (2001) Distinguishing between Good (Useful) and Bad Workloads on Students' Evaluations of Teaching. *American Educational Research Journal,* 38 (1), pp. 183-212.

Marsh, H.W. (1987) Students' Evaluations of University Teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research,* 11 (3), pp. 253-388.

Marsh, H.W. (1984) Students' Evaluations of University Teaching: Dimensionality, reliability, validity, potential baises, and utility. *Journal of Educational Psychology,* 76 (5), pp. 707-754.

Marsh, H.W. (1983) Multidimensional Ratings of Teaching Effectiveness by Students from Different Academic Settings and Their Relation to Student/Course/Instructor Characteristics. *Journal of Educational Psychology,* 75 (1), pp. 150-166.

Marsh, H.W. (1982) Validity of Students' Evaluations of College Teaching: A multitrait–multimethod analysis. *Journal of Educational Psychology,* 74 (2), pp. 264-279.

Marsh, H.W. (1980) The Influence of Student, Course, and Instructor Characteristics in Evaluations of University Teaching. *American Educational Research Journal,* 17 (2), pp. 219-237.

Marsh, H.W. and Dunkin, M. (1992) Students' Evaluations of University Teaching: A multidimensional perspective. *Higher Education,* 8, pp. 143-223.

Marsh, H.W.and Dunkin, M.J. (1997) Students' Evaluations of University Teaching: A multidimensional perspective. In *Effective Teaching in Higher Education: Research and Practice.*, eds. R.P. Perry andJ.C. Smart, New York: Agathon, pp. 241-320.

Marsh, H.W., Overall, J.U. and Kesler, S.P. (1979) Validity of Student Evaluations of Instructional Effectiveness: A comparison of faculty self-evaluations and evaluations by their students. *Journal of Educational Psychology,* 71 (2), pp. 149-160.

Marsh, H.W. and Overall, J.U. (1980) Validity of Students' Evaluations of Teaching Effectiveness: Cognitive and affective criteria. *Journal of Educational Psychology,* 72 (4), pp. 468-475.

Marsh, H.W. and Roche, L.A. (2000) Effects of Grading Leniency and Low Workload on Students' Evaluations of Teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology,* 92 (1), pp. 202-228.

Marsh, H.W. and Ware, J.E. (1982) Effects of Expressiveness, Content Coverage, and Incentive on Multidimensional Student Rating Scales: New interpretations of the Dr. Fox effect. *Journal of Educational Psychology,* 74 (1), pp. 126-134.

McGhee, P. (2003) *The Academic Quality Handbook: Enhancing higher education in universities and further education colleges.* London: Kogan Page.

McNay, I. (1999) The Paradoxes of Research Assessment and Funding. In *Changing Relationship Between Higher Education and the State.*, eds. M. Henkel andB. Little, London: Jessica Kingley, pp. 191-208.

Minelli, E., Rebora, G., Turri, M. and Huisman, J. (2006) The Impact of Research and Teaching Evaluation in Universities: Comparing an Italian and a Dutch case. *Quality in Higher Education,* 12 (2), pp. 109-124.

Minichiello, V., Aroni, R., Timewell, E. and Alexander, L. (1990) *In-Depth Interviewing: Researching people.* Melbourne: Longman Cheshire.

Ministry of Education (2008) *为经济社会又好又快发展作出贡献 [Contributing to a Rapid economic Growth and a Sound Social Development].* Ministry of Education. [Online]. Available from: <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_1344/201001/xxgk_75827.html> (Accessed Oct/8, 2014).

Ministry of Education (2007) *教育部财政部关于实施高等学校本科教学质量与教学改革工程的意见 [The Launch of the Undergraduate Teaching Quality and Teaching Reform Project].* Beijing: Ministry of Education.

Ministry of Education (2005) *1998 普通高等学校本专科学生数 [Number of Undergraduate Students by Type of Courses in Regular Higher Educational Institutions in 1998].* Ministry of Education. [Online]. Available from: <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_573/200505/7678.html> (Accessed Oct/8, 2014).

Ministry of Education (2004a) *2003-2007 年教育振兴行动计划 [2003-2007 Education Action Plan].* Beijing: Ministry of Education.

Ministry of Education (2004b) *教育部办公厅关于印发《普通高等学校本科教学工作水平评估方案（试行）》的通知 [Circular of the General Office of the Ministry of Education on the Printing and Distribution of the Undergraduate Teaching and Learning Evaluation Plan for Regular Higher Education Institutions].* Beijing: Ministry of Education.

Ministry of Education (2004c) *普通高等学校本科教学工作水平评估方案 [Undergraduate Teaching Evaluation Scheme].* Beijing: Ministry of Education.

Ministry of Education (2003a) *教育部关于做好第一届高等学校教学名师奖获奖教师奖励工作的通知 [Notice on the First Higher Education Distinguished Teacher Award].* Beijing: Ministry of Education.

Ministry of Education (2003b) *教育部关于启动高等学校教学质量与教学改革工程精品课程建设工作的通知 [Notice on the Launch of the Higher Education Teaching Quality Reform Excellent Course Project]*. Beijing: Ministry of Education.

Ministry of Education (2001) *关于加强高等学校本科教学工作提高教学质量的若干意见 [Some Opinions on the Improvement of the Quality of Teaching in Higher Education]*. Beijing: Ministry of Education.

Ministry of Education (1998) *面向 21 世纪教育行动振兴计划 [An Action Plan for Invigorating Education Oriented to the 21st Century]*. Beijing: Ministry of Education.

Ministry of Education (1977) *关于 1977 年高等学校招生工作的意见 [1977 Student Enrolment in Higher Education Institutions]*. Beijing: Ministry of Education.

Moore, S. and Kuol, N. (2005) Students Evaluating Teachers: Exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education,* 10 (1), pp. 57-73.

Morley, L. (2003) *Quality and Power in Higher Education.* Maidenhead: SRHE and Open University Press.

Morley, L. (2002) A Comedy of Manners: Quality and power in higher education. In *Higher Education and Policy Change: Outcomes in turbulent environments.*, ed. P. Trowler,Buckingham: SRHE and Open University Press, pp. 126-141.

Moses, I. (1988) *Academic Staff Evaluation and Development: A university case study.* St Lucia: University of Queensland Press.

Murray, H.G. and Newby, W.G. (1982) Faculty Attitudes toward Evaluation of Teaching at the University of Western Ontario. *Assessment and Evaluation in Higher Education,* 7 (2), pp. 144-151.

Nasser, F. and Fresko, B. (2002) Faculty Views of Student Evaluation of College Teaching. *Assessment & Evaluation in Higher Education,* 27 (2), pp. 187-198.

Newton, J. (2002) Views From Below: Academics coping with quality. *Quality in Higher Education,* 8 (1), pp. 39-61.

Newton, J. (2000) Feeding the Beast or Improving Quality?: Academics' perceptions of quality assurance and quality monitoring. *Quality in Higher Education,* 6 (2), pp. 153-163.

Newton, J. (1999a) An Evaluation of the Impact of External Quality Monitoring on a Higher Education College (1993–98). *Assessment & Evaluation in Higher Education,* 24 (2), pp. 215-235.

Newton, J. (1999b) *The Tension Between Accountability and Quality Improvement in Higher Education.* PhD. University of Wales - Bangor.

Oliveira, C.G., Oliveira, P.C. and Costa, N. (2012) Students' and Teachers' Perspectives About Quality of Engineering Education in Portugal. *European Journal of Engineering Education,* 37 (1), pp. 49-57.

PRC President (1998) *Higher Education Law*. Beijing: Standing Committee of the National People's Congress.

Project 211 University (2007a) *关于印发 211 大学关于实施质量工程深化教学改革全面提高教学质量的决定的通知 [Implement the Quality Project and Deepen the Teaching Reform in Order to Improve the Quality of Teaching]*. City211: Project 211 University.

Project 211 University (2007b) *学校举行本科评建教学成果展示研讨会 [A Meeting to Exhibit the Achievements Made during the Preparation for the UTE]*. City211: Project 211 University Newsletter.

Project 211 University (2007c) *学校召开本科教学评建工作部署会议 [The UTE Work Plan Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2007d) *学校召开院级教学文档循环检查工作总结会议 [Meeting to Summarise the Loop Checks of Documentations]*. City211: Project 211 University Nesletter.

Project 211 University (2007e) *巩固成果加快整改全力以赴做好本科评建创优工作—我校召开本科教学工作水平评估自评自查总结及预评估动员大会 [Consolidate the Achievements, Accelerate Rectification, and Endeavor to Achieve Excellence in the UTE –A meeting to conclude the self-assessment and to mobilise staff for the mock evaluation]*. City211: Project 211 University Newsletter.

Project 211 University (2007f) *我校召开本科教学评建工作校内自查总结及寒假评建工作安排会议 [A Meeting to Conclude the Self-assessment and Arrangement of the Preparation Work During Winter Vocation]*. City211: Project 211 University Newsletter.

Project 211 University (2007g) *我校本科教学工作水平预评估顺利举行 [The Pre-evaluation has been Successfully Conducted]*. City211: Project 211 University Newsletter.

Project 211 University (2007h) *暑假期间校领导带队检查本科教学文档 [University Leaders Checked Documentations During Summer Vocation]*. City211: Project 211 University Newsletter.

Project 211 University (2007i) *校领导带队检查拟作评估专家备选考察点情况 [University Leaders Led the Selection of Departments that will be Recommended to Evaluators]*. City211: Project 211 University Newsletter.

Project 211 University (2007j) *评估专家为我校本科教学工作把脉会诊 [Consultation with Evaluators]*. City211: Project 211 University Newsletter.

Project 211 University (2007k) *领导督导学生评教结果统计 [A Summary of the Results of the Class Observations by Leaders and Inspectors and the Student Evaluation of Teachers]*. City211: Project 211 University.

Project 211 University (2006a) *211 大学专业技术资格评审办法 [The Project 211 University Professional Qualification Review Criteria]*. City211: Project 211 University.

Project 211 University (2006b) *关于211大学教学科研系列教职工量化考核计分标准有关说明 [The Project 211 University Quantitative Performance-based Assessment for Staff in Teaching and Research Posts]*. City211: Project 211 University.

Project 211 University (2006c) *学校召开新一轮本科教学迎评工作研讨会 [A Preparation Meeting for the New Round of the UTE]*. City211: Project 211 University Newsletter.

Project 211 University (2006d) *学校召开本科教学工作水平评估工作会议 [The UTE Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2006e) *学校召开本科教学迎评工作检查布置会 [The UTE Work Plan Meeting]*. City211: Project 211 University Nesletter.

Project 211 University (2006f) *我校召开2006年度精品课程建设工作会议 [The Excellent Course Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2006g) *我校召开本科教学工作迎评第二次动员大会 [The Second Mobilization Meeting for the Preparation of the UTE]*. City211: Project 211 University Newsletter.

Provincial Personnel Department (1999) *省高等学校教授研究员资格条件, 省高等学校副教授副研究员资格条件 [The Provincial Requirements for the Application of Professorship and Associate Professorship]*. City: Provincial Personnel Department.

Qi, Y. (2002) 论高等教育大众化时代的质量观 [Discussion on the View of Higher Education's Quality in Era of Masses]. *Teacher Education Research,* 14 (2), pp. 39-44.

Qi, Y. and Chen, Y. (1997) 论教育质量观与素质教育 [On the Definition of the Quality of Education and Essential-Qualities-Oriented Education]. *Journal of the Chinese Society of Education,* 3, pp. 26-29.

Qian, J. and An, X. (2006) 试论高等教育的再生性质量观 [On the Concept of Repeatability Quality in Higher Education]. *Jiangsu Higher Education,* 5, pp. 57-59.

Qu, Q. and Li, X. (2008) 我国首轮本科教学评估效益的调查分析 [Survey and Analysis of the Benefit of the First-round Undergraduate Teaching Evaluation in China]. *Higher Education Development and Evaluation,* 25 (4), pp. 50-55.

Qu, Z. (2008) *坚持内涵发展 着力提升质量—对新阶段我国高等教育发展的几点认识 [Improving Quality Through Internal Development – Some thoughts on the development of higher education]*. Beijing: China Education Daily.

Ratcliff, J.L. (2003) Dynamic and Communicative Aspects of Quality Assurance. *Quality in Higher Education,* 9 (2), pp. 117-131.

Robson, C. (2002) *Real World Research.* Second ed. Oxford: Blackwell.

Ryan, G.W. and Bernard, H.R. (2003) Techniques to Identify Themes. *Field Methods,* 15, pp. 85-109.

Ryan, J.J., Anderson, J.A. and Birchler, A.B. (1980) Student Evaluation: The faculty responds. *Research in Higher Education,* 12 (4), pp. 317-333.

Saarinen, T. (1995) Systematic Higher Education Assessment and Departmental Impacts: Translating the effort to meet the need. *Quality in Higher Education,* 1 (3), pp. 223-234.

Schneider, A. and Ingram, H. (1997) *Policy Design for Democracy.* Lawrence: University Press of Kansas.

Schneider, A. and Ingram, H. (1990) Behavioral Assumptions of Policy Tools. *Journal of Politics,* 52 (2), pp. 510-529.

Seale, C. (2004) *Researching Society and Culture.* London: Sage.

Shah, M. (2012) Ten Years of External Quality Audit in Australia: Evaluating its effectiveness and success. *Assessment & Evaluation in Higher Education,* 37 (6), pp. 761-772.

Shore, C. and Wright, S. (1999) Audit Culture and Anthropology: Neo-liberalism in British higher education. *The Journal of the Royal Anthropological Institute,* 5 (4), pp. 557-575.

Shortland, S. (2004) Peer Observation: A tool for staff development or compliance? *Journal of Further and Higher Education,* 28 (2), pp. 219-228.

Silva, M., Reich, R. and Gallegos, G. (1997) Effects of External Quality Evaluation in Chile: a preliminary study. *Quality in higher education,* 3 (1), pp. 27-35.

Smith, L.M. (1978) An Evolving Logic of Participant Observation, Educational Ethnography, and Other Case Studies. *Review of Research in Education,* 6, pp. 316-377.

Song, Y. (2002) 开展教学状况与质量检查 [Launch of Teaching Quality Investigation]. In *China Education Yearbook.*, ed. China Education Yearbook Editorial Department, Beijing: People's Education Press, pp. 187-188.

Spencer, L., Ritchie, J., Lewis, J. and Dillon, L. (2003) *Quality in Qualitative Evaluation: A framework for assessing research evidence.* London: Cabinet Office.

Spencer, P.A. and Flyr, M.L. (1992) *The Formal Evaluation as an Impetus to Classroom Change: Myth or reality?* Washington, D.C.: ERIC.

Stake, R.E. (2005) Qualitative Case Studies. In *The Sage Handbook of Qualitative Research.*, eds. N.K. Denzin andY.S. Lincoln, Third ed. London: Sage, pp. 443-466.

Stake, R.E. (1995) *The Art of Case Study Research.* London: Sage.

Stake, R.E. (1988) Case Study Methods in Educational Research: Seeking sweet water. In *Complementary methods for research in education.*, ed. L.S. Shulman,Washington, DC: American Educational Research Association, pp. 253-300.

State Council (1981) *中华人民共和国学位条例暂行实施办法 [Degree Regulations of the People's Republic of China].* Beijing: State Council.

State Council Academic Degrees Committee (1981) *国务院学位委员会关于审定学位授予单位的原则和办法 [The Principles and Rules for Granting Degree-awarding Power]*. Beijing: State Council.

State Education Commission (1989) *国家教委、国家物价局、财政部关于普通高等学校收取学杂费和住宿费代规定 [Regulations for Collecting Tuition, Sundry and Boarding Fees in Ordinary Universities]*. Beijing: State Education Commission.

State Education Commission (1985) *关于开展高等工程教育评估研究和试点工作的通知 [Engineering Higher Education Evaluation Research and Trial]*. Beijing: State Education Commission.

Stenhouse, L. (1978) Case Study and Case Records: Towards a contemporary history of education. *British Educational Research Journal,* 4 (2), pp. 21-39.

Sun, H. (2003) *研究生急速扩招质量如何保证 [How can Graduate Education Retain its Quality after the Rapid Increase in Enrollment]*. Xian: Xian Daily.

Sun, Y. and Zhang, Z. (2008) 高等教育教学质量保障体系的分析与构建亲 [On Establishing a Guarantee System for Higher Educational Quality]. *Heilongjiang Education,* 11, pp. 86-87.

Tam, M. (2007) Assessing Quality Experience and Learning Outcomes: Part II: Findings and discussion. *Quality Assurance in Education,* 15 (1), pp. 61-76.

Tan, N. (2005) *如何看待"高等教育质量"—与厦门大学潘懋元教授一席谈 [What is the Quality of Higher Education – A conversation with Professor Pan Maoyuan]*. Beijing: China Education Daily.

Taylor, J.W., Atwood, C.H. and Hutchings, P.A. (2000) Why are Chemists and Other Scientists Afraid of the Peer Review of Teaching? *Journal of Chemical Education,* 77 (2), pp. 239-244.

Teelken, C. and Lomas, L. (2009) "How to Strike the Right Balance Between Quality Assurance and Quality Control in the Perceptions of Individual Lecturers": A comparison of UK and Dutch higher education institutions. *Tertiary Education and Management,* 15 (3), pp. 259-275.

The President of the Project 211 University (2006) *An Interview.* Beijing: Xinhua.

Thomas, S., Chie, Q.T., Abraham, M., Raj, S.J. and Beh, L. (2014) A Qualitative Review of Literature on Peer Review of Teaching in Higher Education An Application of the SWOT Framework. *Review of Educational Research,* 84 (1), pp. 112-159.

Tian, E. (2005) *高等教育质量保证模式研究 [Higher Education Quality Assurance System]*. PhD. Huazhong University of Science and Technology.

Tollefson, N. and Tracy, D.B. (1983) Comparison of Self-reported Teaching Behaviors of Award-winning and Non-award Winning University Faculty. *Perceptual and motor skills,* 56 (1), pp. 39-44.

Trow, M. (1973) *Problems in the Transition from Elite to Mass Higher Education.* Berkeley: Carnegie Commission on Higher Education.

Tuckman, B.W. (1972) *Conducting Educational Research.* New York: Harcourt Brace Jovanovich.

Uni211 Yearbook (2011) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2010) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2009) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2008) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2007) *Project 211 University Yearbook.* City211: Project 211 University.

Uni985 Yearbook (2010) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2009) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2008) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2007) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2006) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2005) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2004) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2003) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2002) *Project 985 University Yearbook.* City985: Project 985 University.

UniCity Yearbook (2010) *City University Yearbook.* City: City University.

UniCity Yearbook (2009) *City University Yearbook.* City: City University.

UniCity Yearbook (2008) *City University Yearbook.* City: City University.

UniCity Yearbook (2007) *City University Yearbook.* City: City University.

UniCity Yearbook (2006) *City University Yearbook.* City: City University.

UniCity Yearbook (2005) *City University Yearbook.* City: City University.

UniCity Yearbook (2004) *City University Yearbook.* City: City University.

UniCity Yearbook (2003) *City University Yearbook.* City: City University.

Vaden-Kiernan, N., Michie, J. and Frechtling, J.A. (1994) *Short-Term Impact Study of the Presidential Awards for Excellence in Science and Mathematics Teaching.* Arlington: National Science Foundation.

Vilgats, B. and Heidmets, M. (2011) The Impact of External Quality Assessment on Universities: the Estonian experience. *Higher Education Policy,* 24 (3), pp. 331-346.

Volkwein, F.J., Lattuca, L.R., Harper, B.J. and Domingo, R.J. (2006) *Getting in Sync: The impact of accreditation on student experiences and learning outcomes.* Annual Meeting of the Association for Institutional Research, 17 May, Chicago.

Vroeijenstijn, A.I. (1995) *Improvement and Accountability: Navigating between Scylla and Charybdis. Guide for External Quality Assessment in Higher Education.* London: Jessica Kingsley.

Wang, J. and Zheng, N. (2008) *提升高等教育质量我们还差什么？[What is Missing for Quality Education?].* Beijing: Science and Technology Daily.

Wang, J. (2008) 高等教育的质量保证与质量提高—与巴奈特教授的对话 [The quality assurance and quality improvement of higher education: Dialog with Professor Barnett]. *Journal of Higher Education,* 29 (12), pp. 23-27.

Wang, N. (2010) *高等学校教学名师工程研究 [Higher Education Distinguished Teacher Project].* Master. South-Central University For Nationalities.

Wang, P. and Zhang, Z. (2008) 地方高校教学质量保障体系的现实分析与模式重构 [Analysis of the Status Quo and a Proposed Model of Quality Assurance in Regional Universities]. *Southwest University of Science and Technology,* 3, pp. 77-80.

Wang, X. (2009) 以质量强化为导向的苏格兰高等教育评估—ELIR 的经验与启示 [Enhancement-led Institutional Review in Scottish Higher Education]. *Journal of Yangzhou University,* 13 (6), pp. 49-53.

Wang, Y.and Zhang, X. (2007) *教学评估岂能变成"集体秀" [How did the Teaching Evaluation Become a Show].* Guangzhou: Nanfang Daily.

Wang, Y. and Liu, B. (2009) *中国教育改革 30 年: 高等教育卷 [30 Years Education Reform: Higher Education].* Beijing: Beijing Normal University Press.

Wang, Z. (2006) *高等教育病在定位模糊 [Higher Education's Illness Comes from Vague Positioning].* Tianjin: Northern Economic Times.

Webb, W.B. and Nolan, C.Y. (1955) Student, Supervisor, and Self-Ratings of Instructional Proficiency. *Journal of Educational Psychology,* 46 (1), pp. 42-46.

Webster, B.J., Chan, W.S.C., Prosser, M.T. and Watkins, D.A. (2009) Undergraduates' Learning Experience and Learning Process: Quantitative evidence from the East. *Higher Education,* 58 (3), pp. 375-386.

Westerheijden, D.F. (1997) A Solid Base for Decisions: Use of the VSNU research evaluations in Dutch universities. *Higher Education,* 33 (4), pp. 397-413.

Williamson, C. J. (2005) *External Quality Assessment in the Context of Managerialism and Higher Education Public Policy.* PhD. University of Strathclyde.

Xi, C. and Huang, X. (2010) 地方高校构建内部教学质量保障体系的主要障碍及其对策 [Major Problems in Teaching Quality Control of Regional Universities in China]. *Journal of Shaanxi University of Technology,* 28 (3), pp. 54-57.

Xie, Y. (2010) *高校教学团队建设研究 [The Research on Teaching Team: Building in university].* Master. Hunan University.

Xu, T. and Fang, B. (2012) 应用型：地方高校人才培养的必然选择 [Application: The inevitable choice of local universities' talents cultivation]. *Journal of Higher Education,* 33 (6), pp. 59-65.

Yan, Z. (2011) 20 世纪 70 年代末以来我国高等教育质量研究的回顾与瞻望 [A Review and Prospect of Higher Education Quality Research in China Since the End of 70's of the 20th Century]. *Journal of Education Institute of Taiyuan University,* 29 (1), pp. 4-9.

Yan, Z. (2006) *中国高等教育宏观决策体制与过程研究 [A Study of the Macro Decision-making System and Process of Higher Education in China].* PhD. Xiamen University.

Yang, D. and Li, M. (2011) "宽进严出"是我国高校发展的必然趋势 ["Lenient Entry , Stringent Exit" as the Inevitable Tendency of the Development of Higher Education in China]. *Modern University Education,* 3, pp. 23-26.

Yang, M. and Zhao, L. (2012) 论普通高校十年扩招中经费投入的特征、问题及对策 [On the Features, Problems and Countermeasures for Funding of Regular Higher Education Institutions(HEIs) during the Last Decade]. *Journal of Zhejiang University,* 42 (5), pp. 165-177.

Yang, Q. (2010) 高等教育质量观:一元与多元 [The View on Quality in Higher Education: One-dimensional or multiple]. *Journal of Changchun University of Technology,* 31 (2), pp. 14-16.

Yardley, L. (2000) Dilemmas in Qualitative Health Research. *Psychology and Health,* 15 (2), pp. 215-228.

Ye, T. (2006) *有些大学上演造假运动闯关本科评估 [Some Universities Relied on Fraud to Pass the Undergraduate Evaluation].* Beijing: China Youth Daily.

Yin, R.K. (2003) *Case Study Research: Design and method.* Third ed. London: Sage.

Yu, K. (2010) *Diversification to a Degree: An exploratory study of students' experience at four higher education institutions in China.* Bern: Peter Lang.

Yuan, Y. (2011) 院校評估如何支持院校質量改進 [How to Support Quality Improvement by Evaluation]. *Higher Education Exploration,* 1, pp. 50-54.

Yuan, Y. (2010) 质量保障：重在保障，而非改进—对本科教学工作水平评估的反思与制度建设构想 [Quality Assurance: More Attention to Assurance than Improvement]. *Higher Education Development and Evaluation,* 26 (1), pp. 1-14.

Zeng, M. and Zhang, J. (2010) 我国首轮本科教学工作水平评估背景综述 [Summary of the Background of China's First Round of the Undergraduate Teaching Assessment]. *Heilongjiang Education,* 12, pp. 26-28.

Zhang, H., Foskett, N., Wang, D. and Qu, M. (2011) Student Satisfaction with Undergraduate Teaching in China – A comparison between research-intensive and other universities. *Higher Education Policy,* 24 (1), pp. 1-24.

Zhang, H. and Xue, Z. (2009) 我国第一轮本科教育评估反馈的有效性分析 [The Impact of the First Round Undergraduate Teaching Evaluation]. *Higher Education Exploration,* 2, pp. 47-53.

Zhang, J. and Zhang, H. (2008a) 我国高等教育质量评估工作现状的反馈—对 54 所高等院校的调查分析 [The Feedback of Chinese Higher Education Quality Evaluation – Investigation to 54 Higher Schools]. *Higher Education Research & Evaluation,* 4, pp. 58-61.

Zhang, J. and Zhang, S. (2008b) 高校教师视角下本科教学评估成效的调查分析 [An Empirical Study on Impact of the Undergraduate Teaching Evaluation from Faculty's Perspective]. *Journal of National Academy of Education Administration,* 6, pp. 30-34.

Zhang, L. (2006) *中国教育存在的问题及发展与改革的政策取向 [Problems and the Direction of the Education Reform in China]*. Beijing: China Economic Times.

Zheng, Y. (2010) 我国高校学院制改革的历史考察与现实思考 [Historical Investigation and Realistic Thinking of Chinese College System Reform]. *Journal of Huaiyin Teachers College,* 32 (6), pp. 822-828.

Zhou, J. (2007) 实施"质量工程"贯彻"2 号文件"全面提高高等教育质量 [Implement the 'Quality Project' and the 'No.2 Document' in Order to Improve the Quality of Higher Education. *China University Teaching,* 3, pp. 4-8.

Zhou, J. (2004) 牢记宗旨、不负使命、办好让人民满意的教育 [Remember Tenet Firmly, not to Fail the Mission, and Work to Ensure that Education Meets with the Satisfaction of the People]. In *China Education Yearbook.*, ed. China Education Yearbook Editorial Department, Beijing: People's Education Press, pp. 16-21.

Zhou, J. (2009) *国家精品课程建设的现状、问题与对策—以教育类国家精品课程为例 [Research on the Status Quo, Problems and Solutions of National Elaborate Course – Analysis of National Educational Elaborate Course]*. Master. Zhejiang Normal University.

Zhou, S. and Zhu, W. (2004) *对我国高等教育发展的理性思考 [A Reflection on the Development of Higher Education]*. Beijing: China Education Daily.

Zhou, X., Zhang, G. and Chen, J. (2010) 高等教育评估制度的沿革与展望 [Evolution and Prospects for Higher Education Evaluation System in China]. *Journal of Architectural Education in Institutions of Higher Learning,* 19 (2), pp. 24.

Appendices

Appendix 1 A List of Reviewed Policy Documents

Policy Documents Regarding National Institutional Rules

CPC Central Committee (1984) *中共中央关于教育体制改革的决定 [The CPC Central Committee Decision on the Reform of the Education System]*. Beijing: CPC Central Committee.

CPC Central Committee (1985) *中共中央关于经济体制改革的决定 [The CPC Central Committee's Resolution on the Reform of the Economic System]*. Beijing: CPC Central Committee.

Ministry of Education (2001) *关于加强高等学校本科教学工作提高教学质量的若干意见 [Some Opinions on the Improvement of the Quality of Teaching in Higher Education]*. Beijing: Ministry of Education.

Ministry of Education (1998) *面向 21 世纪教育行动振兴计划 [An Action Plan for Invigorating Education Oriented to the 21st Century]*. Beijing: Ministry of Education.

Ministry of Education (2004) *2003-2007 年教育振兴行动计划 [2003-2007 Education Action Plan]*. Beijing: Ministry of Education.

PRC President (1998) *Higher Education Law.* Beijing: Standing Committee of the National People's Congress.

State Council (1981) *中华人民共和国学位条例暂行实施办法 [Degree Regulations of the People's Republic of China]*. Beijing: State Council.

State Council Academic Degrees Committee (1981) *国务院学位委员会关于审定学位授予单位的原则和办法 [The Principles and Rules for Granting Degree-awarding Power]*. Beijing: State Council.

State Education Commission (1989) *国家教委、国家物价局、财政部关于普通高等学校收取学杂费和住宿费代规定 [Regulations for Collecting Tuition, Sundry and Boarding Fees in Ordinary Universities]*. Beijing: State Education Commission.

Policy Documents Regarding the UTE

Anonymous (1999) 今天教育扩招可拉动经济增量 70 亿 [The Increase of Participation in Higher Education Can Generate ¥7 Billion Economic Growth]. *Leader Decision Making Information,* 37, pp. 29.

Ministry of Education (1998) *面向 21 世纪教育行动振兴计划 [An Action Plan for Invigorating Education Oriented to the 21st Century]*. Beijing: Ministry of Education.

Ministry of Education (2001) *关于加强高等学校本科教学工作提高教学质量的若干意见 [Some Opinions on the Improvement of the Quality of Teaching in Higher Education]*. Beijing: Ministry of Education.

Ministry of Education (2004) *普通高等学校本科教学工作水平评估方案 [Undergraduate Teaching Evaluation Scheme]*. Beijing: Ministry of Education.

Ministry of Education (2005) *1998 普通高等学校本专科学生数 [Number of Undergraduate Students by Type of Courses in Regular Higher Educational Institutions in 1998]*. Ministry of Education. [Online]. Available from: <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_573/200505/7678.html> (Accessed Oct/8, 2014).

Ministry of Education (2008) *为经济社会又好又快发展作出贡献 [Contributing to a Rapid economic Growth and a Sound Social Development]*. Ministry of Education. [Online]. Available from: <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_1344/201001/xxgk_75827.html> (Accessed Oct/8, 2014).

Zhou, J. (2004) 牢记宗旨、不负使命、办好让人民满意的教育 [Remember Tenet Firmly, not to Fail the Mission, and Work to Ensure that Education Meets with the Satisfaction of the People]. In *China Education Yearbook.*, ed. China Education Yearbook Editorial Department, Beijing: People's Education Press, pp. 16-21.

Policy Documents Regarding the Quality Project

Ministry of Education (2003) 教育部办公厅关于印发国家精品课程建设工作实施办法的通知 [Notice of the General Office of the Ministry of Education on the Issuance of the Excellent Courses Project Implementation Measures]. Beijing: Ministry of Education.

Ministry of Education (2004) *2003-2007 年教育振兴行动计划 [2003-2007 Education Action Plan]*. Beijing: Ministry of Education.

Ministry of Education (2004) 教育部办公厅关于国家精品课程建设工作实施办法补充规定的通知 [Notice of the General Office of the Ministry of Education on the Supplementary Regulations of the Excellent Courses Project]. Beijing: Ministry of Education.

Ministry of Education (2007) 关于进行 2007 年度普通高等教育精品教材评选的通知 [The Notice on the Selection of the 2007 Excellent Textbooks]. Beijing: Ministry of Education.

Ministry of Education (2007) 国家精品课程评审指标 [The Selection Criteria of the 2007 Excellent Courses]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部关于表彰第三届高等学校教学名师奖获奖教师的决定 [The Decision of the Ministry of Education to Reward the Third Distinguished Teacher Award-winning teachers]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部办公厅关于 2007 年度国家精品课程申报工作的通知 [The Notice of the General Office of the Ministry of Education on the application of the 2007 Excellent Courses]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部办公厅关于公布 2007 年度普通高等教育精品教材书目的通知 [The Notice of the General Office of the Ministry of Education on Announcing the Catalogue for the Selection of the 2007 Excellent Textbooks]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部办公厅关于组织开展第三届高等学校教学名师奖评选表彰工作的通知 [The Notice of the General Office of the Ministry of Education Organised the Third Distinguished Teacher Awards]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部财政部关于批准 2007 年度双语教学示范课程建设项目的通知 [Notice on the Ministry of Education and the Ministry of Finance's Approval of 2007 Bilingual Courses]. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部财政部关于批准 2007 年度国家精品课程建设项目的通知 [2007 Bilingual Courses Permitted by the Ministry of Education and the Ministry of Finance]. Beijing: Ministry of Education.

Ministry of Education (2007) *教育部财政部关于实施高等学校本科教学质量与教学改革工程的意见 [The Launch of the Undergraduate Teaching Quality and Teaching Reform Project]*. Beijing: Ministry of Education.

Ministry of Education (2007) 教育部财政部立项建设 2007 年度双语教学示范课程 [2007 Bilingual Courses Permitted by the Ministry of Education and the Ministry of Finance]. Beijing: Ministry of Education.

Ministry of Education (2007) 普通高等教育十一五规划教材申请书 [Application Form for the Eleventh Five-Year National Textbooks]. Beijing: Ministry of Education.

Ministry of Education (2007) 第三届高等学校教学名师奖候选人推荐表 [Application Form for the Third Distinguished Teacher Awards]. Beijing: Ministry of Education.

Ministry of Education (2007) 第三届高等学校教学名师奖评选指标体系 [The Selection Criteria of the Third Distinguished Teacher Awards]. Beijing: Ministry of Education.

Ministry of Education (2006) 教育部关于印发普通高等教育十一五国家级教材规划选题的通知 [Ministry of Education on the Issuance the Topics of the Eleventh Five-Year National Textbooks]. Beijing: Ministry of Education.

Ministry of Education (2006) 教育部办公厅关于加强普通高等教育十一五国家级规划教材管理的通知 [The Notice of the General Office of the Ministry of Education on the Eleventh Five-Year National Textbooks]. Beijing: Ministry of Education.

Zhou, J. (2007) 实施"质量工程"贯彻"2 号文件"全面提高高等教育质量 [Implement the 'Quality Project' and the 'No.2 Document' in Order to Improve the Quality of Higher Education. *China University Teaching,* 3, pp. 4-8.

Policy Documents Regarding the Project 985 University

Anonymous (2012) *中国大学 985 工程三期拨款清单 [A List of Project 985 Funding by Universities].* Sina. [Online]. Available from: <http://blog.sina.com.cn/s/blog_5557acef0101bw77.html> (Accessed Oct/8, 2014).

Uni985 Yearbook (2002) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2003) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2004) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2005) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2006) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2007) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2008) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2009) *Project 985 University Yearbook.* City985: Project 985 University.

Uni985 Yearbook (2010) *Project 985 University Yearbook.* City985: Project 985 University.

Policy Documents Regarding the Project 211 University

Project 211 University (2006) *211 大学专业技术资格评审办法 [The Project 211 University Professional Qualification Review Criteria]*. City211: Project 211 University.

Project 211 University (2006) *关于 211 大学教学科研系列教职工量化考核计分标准有关说明 [The Project 211 University Quantitative Performance-based Assessment for Staff in Teaching and Research Posts]*. City211: Project 211 University.

Project 211 University (2006) *学校召开新一轮本科教学迎评工作研讨会 [A Preparation Meeting for the New Round of the UTE]*. City211: Project 211 University Newsletter.

Project 211 University (2006) *学校召开本科教学工作水平评估工作会议 [The UTE Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2006) *学校召开本科教学迎评工作检查布置会 [The UTE Work Plan Meeting]*. City211: Project 211 University Nesletter.

Project 211 University (2006) *我校召开 2006 年度精品课程建设工作会议 [The Excellent Course Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2006) *我校召开本科教学工作迎评第二次动员大会 [The Second Mobilization Meeting for the Preparation of the UTE]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *关于印发 211 大学关于实施质量工程深化教学改革全面提高教学质量的决定的通知 [Implement the Quality Project and Deepen the Teaching Reform in Order to Improve the Quality of Teaching]*. City211: Project 211 University.

Project 211 University (2007) *学校举行本科评建教学成果展示研讨会 [A Meeting to Exhibit the Achievements Made during the Preparation for the UTE]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *学校召开本科教学评建工作部署会议 [The UTE Work Plan Meeting]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *学校召开院级教学文档循环检查工作总结会议 [Meeting to Summarise the Loop Checks of Documentations]*. City211: Project 211 University Nesletter.

Project 211 University (2007) *巩固成果加快整改全力以赴做好本科评建创优工作—我校召开本科教学工作水平评估自评自查总结及预评估动员大会 [Consolidate the Achievements, Accelerate Rectification, and Endeavor to Achieve Excellence in the UTE –A meeting to conclude the self-assessment and to mobilise staff for the mock evaluation]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *我校召开本科教学评建工作校内自查总结及寒假评建工作安排会议 [A Meeting to Conclude the Self-assessment and Arrangement of the Preparation Work During Winter Vocation]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *我校本科教学工作水平预评估顺利举行 [The Pre-evaluation has been Successfully Conducted]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *暑假期间校领导带队检查本科教学文档 [University Leaders Checked Documentations During Summer Vocation]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *校领导带队检查拟作评估专家备选考察点情况 [University Leaders Led the Selection of Departments that will be Recommended to Evaluators]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *评估专家为我校本科教学工作把脉会诊 [Consultation with Evaluators]*. City211: Project 211 University Newsletter.

Project 211 University (2007) *领导督导学生评教结果统计 [A Summary of the Results of the Class Observations by Leaders and Inspectors and the Student Evaluation of Teachers]*. City211: Project 211 University.

The President of the Project 211 University (2006) *An Interview.* Beijing: Xinhua.

Uni211 Yearbook (2007) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2008) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2009) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2010) *Project 211 University Yearbook.* City211: Project 211 University.

Uni211 Yearbook (2011) *Project 211 University Yearbook.* City211: Project 211 University.

Policy Documents Regarding City University

Anonymous (2009) *城市大学：降分录取是误解 [City University: Lower admission standards are misinterpretation]*. City: City Evening News.

City University (2002) *教职工教学科研管理奖励暂行规定 [The Allowance Allocation Scheme]*. City: City University.

City University (2005) *科研工作奖励管理条例 [Research Rewarding Regulations]*. City: City University.

City University (2006) *城市大学自评报告 [City University Report]*. City: City University.

Provincial Personnel Department (1999) *省高等学校教授研究员资格条件, 省高等学校副教授副研究员资格条件 [The Provincial Requirements for the Application of Professorship and Associate Professorship]*. City: Provincial Personnel Department.

UniCity Yearbook (2003) *City University Yearbook.* City: City University.

UniCity Yearbook (2004) *City University Yearbook.* City: City University.

UniCity Yearbook (2005) *City University Yearbook.* City: City University.

UniCity Yearbook (2006) *City University Yearbook.* City: City University.

UniCity Yearbook (2007) *City University Yearbook.* City: City University.

UniCity Yearbook (2008) *City University Yearbook.* City: City University.

UniCity Yearbook (2009) *City University Yearbook.* City: City University.

UniCity Yearbook (2010) *City University Yearbook.* City: City University.

Appendix 2 The Second Interview Schedule

For the semi-structured interviews in the main study, a list of factors that influence teachers will be provided. Interviewees will be asked to use stars to indicate the degree of influence: 3 stars indicate a big influence; 2 stars indicate some influence; 1 star indicates a small influence; no star indicates no influence. The purpose of asking interviewees to mark items with stars is not to calculate the influence of these factors by the number of stars but to open up a conversation. When interviewees choose to put different numbers of stars to indicate the degree of influence on them, they will feel the need to justify their choice and will be more willing to talk to the researcher and convince the researcher of the reasoning involved in their particular choice. Thus, the researcher is more likely to find out the reasons and the story behind these stars.

The list of factors is:

Students
Classroom interaction*
Social interaction with students*
Informal student feedback
Student evaluation of teachers*
Student course selection*

Colleagues
Classroom observation by colleagues or inspectors
Teaching group meeting
Chatting with colleagues*

Teacher Himself or Herself
Satisfaction*
Professional morale
Self-awareness of image*
Personal or professional development

External Mechanisms
Competitions participated in or awards received by interviewees
Government or commercial projects*
Performance-related pay scheme for teachers*
Undergraduate Teaching Evaluation
Quality Project

Factors with an asterisk have emerged in the pilot study. In other words, interviewees brought out the influences of these factors in the pilot study, without the researcher mentioning them in the interview questions.

After interviewees assign the stars, potential questions for interviews with teachers in the main study are:
Why does the interviewee assign a certain factor with a certain number of stars?

How is the interviewee influenced by the factor or what did the interviewee do in relation to the factor?
Does the interviewee think this factor is helpful and why?
From the interviewee's point of view, how could the policies or the procedures related to the factor be improved?

If the interviewee is a teacher who also has a leadership role, the potential questions are:
Why does the interviewee assign a certain factor with a certain number of stars?
How did this department develop the policies or the procedures related to this factor?
What is the intended purpose for the policies or the procedures related to this factor?
How does this department evaluate the outcomes of the policies or the procedures related to this factor?
What are the outcomes of the policies or the procedures related to this factor?

This interview guide has been piloted with two professors. One of them is a head of department. One interview took 55 minutes; the other took 40 minutes.

Appendix 3 The Third Interview Schedule (used in main study)

Opening

My research is interested in how to ensure the quality of higher education, particularly the quality of teaching. I would like to know throughout the years what have helped you to maintain or improve the quality of teaching?

Before we start the interview, can I remind you again the rights you have as an interviewee? (Present the consent form and read the terms aloud and ask for the signature.)

Questions for teachers

A. Introductory section

Could you tell me how many years you have been working as a teacher?
How many undergraduate classes do you teach per week?
Do you teach graduates as well? How many classes or tutorials do you give per week?

B. Improving the quality of teaching and learning

Thinking about your own practice, what has helped you to improve your teaching and student learning?
What initiatives/programmes aimed at improving quality in higher education have had an influence on your teaching?
Which of the things you mentioned have led to changes in your practice?
What are the changes?

(In which order the following sections will be asked depends on the answers in Section B.)

C. Student engagement

(If the interviewee stresses the influence from students, questions in this section will be asked following the Section B)
You have mentioned the …(repeat the influences from students mentioned by the interviewee when answering Section B).
What changes to your teaching have you made in response to students?
Are there other influences from students that make you want to improve your teaching and in what way?

(If the interviewee did not mention the following issues, the following questions will be asked.)
Student evaluation of teachers
What about student evaluation of teachers? Has it got an influence on your thinking and your practice? Why? How?
What changes to your teaching have you made in response to student evaluation of teachers?

Student course selection
What about student course selection? Has it got an influence on your thinking and your practice? Why? How?
What changes to your teaching have you made in response to student evaluation of teachers?

D. Your engagement with teaching quality
(If the interviewee stresses the influence from himself/herself, questions in this section will be asked following the Section B)
You have mentioned the …(repeat the influences from himself/herself mentioned by the interviewee when answering Section B).
Are there other influences from yourself motivate you to improve your teaching and in what way?

(If the interviewee did not mention the following issue, the following questions will be asked.)
Personal or professional development

Which kind of personal or professional development programme has helped you? How?
What changes to your teaching have you made in response to the personal and professional development programme?

E. Your Colleagues' Engagement

(If the interviewee stresses the influence from the colleagues, questions in this section will be asked following the Section B)
You have mentioned the …(repeat the influences from colleagues mentioned by the interviewee when answering Section B).
What changes to your teaching have you made in response to your colleagues?
Are there other influences from your colleagues' make you want to improve your teaching and in what way?

(If the interviewee did not mention the following issue, the following questions will be asked.)
Classroom observation
What about classroom observation by colleagues or inspectors? Has it got an influence on your thinking and your practice? Why? How?
What changes to your teaching have you made in response to classroom observation?

F. Impact of external mechanisms

(If the interviewee stresses the influence of external mechanisms, questions in this section will be asked following the Section B)
You have mentioned the …(repeat the influences from colleagues mentioned by the interviewee when answering Section B).
What changes to your teaching have you made in response to external mechanisms?
Are there other external mechanisms make you want to improve your teaching and in what way?

(If the interviewee did not mention the following issues, the following questions will be asked.)
Undergraduate Teaching Evaluation

What about the Undergraduate Teaching Evaluation (UTE)? Has it got an influence on your thinking and your practice? Why? How?

What changes to your teaching have you made in response to the UTE?

Quality Project

What about the Quality Project? Has it got an influence on your thinking and your practice? Why? How?

What changes to your teaching have you made in response to the Quality Project?

Government or commercial projects

What about government or commercial projects? Have them got an influence on your thinking and your practice? Why? How?

What changes to your teaching have you made in response to the projects you have conducted?

Competitions and awards

What about competitions and awards? Have them got an influence on your thinking and your practice? Why?

What changes to your teaching have you made in response to the competitions and awards?

Performance-related pay scheme for teachers

What about the current pay scheme? Has it got an influence on your thinking and your practice? Why? How?

What changes to your teaching have you made in response to the pay scheme?

G. Your suggestions

Are there particular types of support you need, which can facilitate the improvement of teaching and learning?

Are there any quality assurance initiatives you would like to see in the future?

Questions for teachers who are in leadership roles

- **Introductory section**

Could you tell me how many years you have been working as a teacher?

How many years you have been in a leadership position?

How many undergraduate classes do you teach per week?

Do you teach graduates as well? How many classes or tutorials do you give per week?

- **Policies in your department/college**

Student evaluation of teachers

How does student evaluation of teachers work in this department/college?

How did this department develop the strategy and the procedures for student evaluation of teachers?

How are the results of student evaluations used? Do teachers use the results to inform their work? If yes, in what ways?

Student course selection

How do students in this department/college select their courses?

How did this department develop strategies and procedures for student course selection?

Are the results of students' course selection processes used for quality assurance purposes? If yes, how?

Do teachers use the results to inform their work? If yes, in what ways?

Personal or professional development

What personal or professional development strategies and procedures for staff members are in place in this department/college?

How and why did this department develop these strategies and procedures?

In what ways do teachers use what they learnt from the personal or professional development programmes to improve their teaching?

Classroom observation

Are there strategies or procedures of classroom observation in place in this department/college? Who is observing? Who is deciding on who is observing whom?

How are the results of classroom observation used? Do teachers use the results to inform their work? If yes, in what ways?

The Undergraduate Teaching Evaluation

What impact has the UTE had at this department?

What were the changes brought by the Undergraduate Teaching Evaluation?

Do teachers use the results of the UTE to inform their work? If yes, in what ways?

The Quality Project

What impact has the Quality Project had at this department?

What are the changes brought by the Quality Project?

Do teachers use the changes to inform their work? If yes, in what ways?

Competitions and awards

What competitions or awards are in place in this department/college?

How are the results of the competitions and the awards used? Do teachers use the results to influence their work? If yes, in what ways?

Performance-related pay scheme for teachers

How does the current pay scheme work in your department/college?

How and why did this department develop the pay scheme?

Are teachers motivated by the pay scheme to improve their teaching? If yes, in what ways?

- **Improving the quality of teaching and learning**

Thinking about your own practice, what has helped you to improve your teaching and student learning?

What initiatives/programmes aimed at improving quality in higher education have had an influence on your teaching?

Which of the things you mentioned have led to changes in your practice?

What are the changes?

- **Your suggestion**

Are there particular helps you need, which can facilitate the improvement of teaching and learning?

Are there any quality assurance initiatives you would like to see in the future?

Closing

Thank you very much for participating this research. I really I appreciate the time you took for this interview.

Appendix 4 A List of Category of Themes

- University institutional rules inherited from the national institutional rules
- University institutional rules established by the President of the University
- University policies on class observation by inspectors
- University policies on class observation by leaders
- University policies on student evaluation of teachers
- University policies on teacher appraisal
- The impacts of the UTE perceived by heads
- The impacts of the UTE perceived by teachers
- The impacts of the Quality Project by heads
- The impacts of the Quality Project by teachers
- The impacts of class observation by university inspectors perceived by heads
- The impacts of class observation by university inspectors perceived by teachers
- The impacts of class observation by department inspectors perceived by heads
- The impacts of class observation by department inspectors perceived by teachers
- The impacts of class observation by leaders perceived by heads
- The impacts of class observation by leaders perceived by teachers
- The impacts of student evaluation perceived by heads
- The impacts of student evaluation perceived by teachers
- The impacts of teacher appraisal perceived by heads
- The impacts of teacher appraisal by teachers
- Problems with current quality assurance mechanisms
- Factors that drives teachers to invest in teaching
- Factors that help teachers to improve teaching

Appendix 5 CUREC Form

University of Oxford
CENTRAL UNIVERSITY RESEARCH ETHICS COMMITTEE (CUREC)
CUREC/1A Checklist for the Social Sciences and Humanities

The University of Oxford places a high value on the knowledge, expertise, and integrity of its members and their ability to conduct research to high standards of scholarship and ethics. The research ethics clearance procedures have been established to ensure that the University is meeting its obligations as a responsible institution. They start from the presumption that all members of the University will take their responsibilities and obligations seriously and will ensure that their research on human subjects is conducted according to the established principles and good practice in their fields and in accordance, where appropriate, with legal requirements. Since the requirements of research ethics review will vary from field to field and from project to project, the University accepts that different guidelines and procedures will be appropriate. Please check the CUREC website to ensure that you have the correct form for your project.

This form does not cover research governance, satisfactory methodology, or the health and safety of employees and students. As principal investigator, it is your responsibility to ensure that requirements in these areas are met. Please carry out a risk assessment of the project, in consultation with all researchers involved, using the checklist and CUREC's other documentation.

The use of an asterisk in this form indicates a phrase defined in the glossary. The glossary and further information on the University's research ethics procedures are available from the CUREC website:

www.admin.ox.ac.uk/curec

This form is designed largely for research that falls within the Divisions of Social Sciences and Humanities and which does not involve a high-level of risk to the subjects. Elite interviews, field work and oral history are included in the CUREC process. Please take a moment to read through it and if you have any questions or doubts as to whether it is the appropriate form, please review Section A or consult the CUREC website.

Note on anonymised data: If you are using previously collected anonymised data about people which neither you nor anyone else involved in your study can trace back to the individuals who provided them (e.g. census data, administrative data, secondary analysis), you do not need to obtain ethical approval for your study. Please refer to the definition of *personal data in the glossary and FAQ no. 6 for further guidance.

Note on audit: If you are conducting research on behalf of or at the request of a service provider that matches the definition of *audit in the glossary, you do not need to obtain ethical approval for your study.

SECTION A

	Yes	**No**
1) Are you using research methodologies commonly used in biomedical or behavioural laboratory sciences?		√
2) Is there a significant risk that the research will induce anxiety, stress or other harmful psychological states in participants that might persist beyond the duration of any test or interview in which they are participating?		√
3) Will the research involve human participants recruited by means of their status as present or past NHS patients or their relatives or carers or present or past NHS staff?		√
4) Does the research involve *human participants aged 16 and over who do not have *capacity to consent for themselves? See the Mental Capacity Act 2005		√
5) Is the study to be funded by the US National Institutes of Health or another US federal funding agency?		√

Office use only: IDREC Ref. No. _____
Date of confirmation that checklist accepted on behalf of IDREC: // //

If you answered 'yes', please stop work on this checklist and
for questions 1 and 2, complete CUREC/1 instead (available from www.admin.ox.ac.uk/curec);
for questions 3 and 4, submit your proposal to the appropriate NHS ethics committee (see www.nres.npsa.nhs.uk and www.admin.ox.ac.uk/rso/clinical for further information);
for question 5, or **if you answered 'yes' to questions 1, 2 or 4 and your research will take place outside the EU**, submit your proposal to OXTREC, which uses separate documentation (www.tropicalmedicine.ox.ac.uk/oxtrec).

If you have answered 'no' to all questions in Section A, please complete Sections B-E. This form and any supporting materials should be typewritten.

SECTION B

Principal investigator/ supervisor/student researcher (title and name):	***Ms Shan Huang***
Name of supervisor (STUDENT RESEARCH PROJECTS ONLY):	***Dr Hubert Ertl***
Department or institute:	***Department of Education, University of Oxford***
Address for correspondence (if different):	
Email and phone contact:	***Shan.Huang@education.ox.ac.uk 07809537636***
Title of research project:	***Improving Quality Assurance in Chinese Higher Education: The interaction between teaching staff and other elements of the quality framework***
Brief description of research methods and goals plus description of the nature of participants (including the criteria for inclusion/exclusion, method of recruitment), explanation of how professional guidelines and/or CUREC protocol(s) will be applied (if relevant) and expected use to which the results/data will be put. Please describe how you will obtain informed consent. Approx 400 words.	

My research intends to explore the interaction between teaching staff and other elements of the quality framework in the context of internal quality assurance in universities in China. The quality framework refers to the dynamic relationships between the elements within the university context, such as leadership, institutional policy, institutional procedures, resources, teaching, student learning outcomes, student learning experience, and staff achievement.
To gain insights into what is currently happening in Chinese universities from the perspective of teaching staff, my project will adopt semi-structured interviews as the main research method, combined with documentary analysis. Documents from the public domain capture the national framework for quality assurance in higher education and current practice of internal quality assurance in universities. The participants in my research will be teaching staff in universities. To gather opinions from teaching staff, interviewees will be selected from lecturers, professors, and heads of departments. Interviewees will be purposely selected from different disciplines and a small number of universities.
Initial contact will be made through emails, calls or face-to-face coversations, in which the purposes and methods of the research will be explained. The form for formal, written, informed consent will be given to the participants either at the initial contact or before the interviews. Participants have the right to withdraw at any time from the research. Participants will be debriefed after the interview. To protect interviewees' privacy, their identities will remain anonymous. The data will be treated confidentially, the identities of interviewees will not be revealed at any stage in the research process. My research does not involve deception or harm to participants.
The data obtained from interviews will help me to understand the quality framework in universities by providing answers on what are the elements of quality framework in universities and what are the interactions between the elements from the perspective of teaching staff.

List actual or probable location(s) where project will be conducted, if known:	Universities in China
Anticipated duration of project:	18 months
Anticipated start date:	1 / Feb / 2011
Anticipated end date:	31 / July / 2012
Name and status (e.g. 3rd year undergraduate; post-doctoral research assistant) of others taking part in the project:	None
Please indicate what training on research ethics the researchers involved with this study have received, e.g. the title of the online or in-person course, and date completed (online training available at www.admin.ox.ac.uk/rso/integrity/#Training):	Training received in the context of Educational Research Methods courses required for the completion of the MSc in Educational Research Method and DPhil. Some of the courses involved lectures, discussions and readings on research ethics.

SECTION C
Methods to be used in the study (**tick** as many as apply: this information will help the committee understand the nature of your research and may be used for audit).

	Please tick
Interview	√
Questionnaire	

Analysis of existing records	
Participant performs verbal/paper and pencil/computer based task	
Measurement/recording of motor behaviour	
Audio recording of participant	√
Video recording or phtotography of participant	
Physiological recording from participant	
Participant observation	
Covert observation	
Systematic observation	
Observation of specific organisational practices	
Other (please specify) ***Document analysis***	√

SECTION D

Have you read one or more of the following professional guidelines and do you undertake to use the principles listed there as a guide for your own work? Please note that this is not intended to be an exhaustive list. Links to the guidelines listed below are included on the CUREC website.

	Please tick
British Society of Criminology: Code of Ethics for Researchers in the Field of Criminology [britsoccrim.org/ethical.htm]	
British Educational Research Association Ethical Guidelines for Educational Research [www.bera.ac.uk/publications/guidelines/]	√
Academy of Management's Professional Code of Ethics [www.aomonline.org]	
Association of American Geographers Statement on Professional Ethics [www.aag.org/Info/ethics.htm]	
Oral History Society of the UK Ethical Guidelines [www.oralhistory.org.uk/ethics/index.php]	
American Political Science Association (APSA) Guide to Professional Ethics in Political Science (Section H) [www.apsanet.org/content_9350.cfm]	
British Psychological Society Code of Ethics and Conduct [www.bps.org.uk/the-society/code-of-conduct/code-of-conduct_home.cfm]	
Ethics Guidelines of the Association of Social Anthropologists of the UK and Commonwealth [www.theasa.org/ethics/guidelines.htm]	
Social Research Association: Ethical Guidelines [www.the-sra.org.uk/ethical.htm]	
Statement of Principles of Ethical Research Practice from the Socio-Legal Studies Association [www.slsa.ac.uk]	
Statement of Ethical Practice for the British Sociological Association [www.britsoc.co.uk/equality/Statement+Ethical+Practice.htm]	
Other professional guidelines (please specify):	

SECTION E

Please put a tick in the yes/no column as appropriate to indicate your response.

	Yes	**No**
1) Will you obtain informed consent according to good practice in your discipline before participation?	√	
2) Will you ensure that *personal data collected directly from participants or via a *third party is held and processed in accordance with the provisions of the Data Protection Act?	√	
3) Does the research involve as participants *people whose ability to give free and informed consent is in question? (This includes those under 18 and vulnerable adults.)		√
4) As a consequence of taking part in the research, will participants be at serious risk of rendering themselves liable to criminal prosecution (e.g. by providing information on drug abuse or child abuse)?		√
5) Does the research involve the *deception of participants, as part of the investigation/experiment?		√

If any of your answers above are in a shaded box, please indicate whether those aspects of your project are fully covered by the following.		
6) Research protocol(s) which has/ve received IDREC/CUREC approval?	**Yes**	**No**
If yes, please give protocol number(s):		
7) Professional guidelines that you will be following, as noted under Section D?	**Yes**	**No**

If any of your answers in Section E are in a shaded box and are not covered by a protocol or by professional guidelines, please complete CUREC/2, available to download from the CUREC website. Then submit both this form (you need not complete section F) and the CUREC/2 to the Social Sciences and Humanities IDREC.

If all your answers in Section E are in the unshaded boxes or your answers in shaded boxes are covered by a protocol or professional guidelines, complete Section F and submit this form and any accompanying documents to the Social Sciences and Humanities IDREC (see notes and address below).

SECTION F

Complete this section only if you do not need to submit form CUREC/2.

I understand my responsibilities as principal researcher/supervisor/student researcher as outlined in the CUREC glossary and guidance on the CUREC website.

I declare that the answers above accurately describe my research as presently designed and that I will submit a new checklist should the design of my research change in a way which would alter any of the above responses so as to require completion of CUREC/2 (involving full scrutiny by an IDREC). I will inform the relevant IDREC if I cease to be the principal researcher on this project and supply the name and contact details of my successor if appropriate.
Signed by principal researcher/supervisor/student researcher:…… ……
Date:…26 Jan 2011……
Print name (block capitals)……SHAN…HUANG……
Signed by supervisor:……(for student projects)
Date:……
Print name (block capitals)……HUBERT…ERTL……

I understand the questions and answers that have been entered above describing the research, and I will ensure that my practice in this research complies with these answers, subject to any modifications made by the principal researcher properly authorised by the CUREC system.

Signed by associate/other researcher: …….
Print name (block capitals)……
Date ……

I have read the research project application named above. On the basis of the information available to me, I:

(i) consider the principal researcher/supervisor/student researcher to be aware of her/his ethical responsibilities in regard to this research;

(ii) consider that any ethical issues raised have been satisfactorily resolved or are covered by relevant professional guidelines and/or CUREC approved protocols, and that it is appropriate for the research to proceed without further formal ethical scrutiny at this stage (noting the principal researcher's obligation to report should the design of the research change in a way which would alter any of the above responses so as to require completion of a CUREC/2 full application);

(iii) am satisfied that the proposed project has been/will be subject to appropriate *peer review and is likely to contribute something useful to existing knowledge and/or to the education and training of the researcher(s) and that it is in the *public interest.

(iv) [FOR DEPARTMENTS/FACULTIES WITH A DEPARTMENTAL RESEARCH ETHICS COMMITTEE (DREC) OR EQUIVALENT BODY - PLEASE DELETE IF NOT APPLICABLE] confirm that this checklist (and associated research outline) has been reviewed by the Department's Research Ethics Committee (DREC)/equivalent body, and attach the associated report from that body.

Signed:……..
(Head of department or nominee e.g Chair of DREC, Director of Graduate Studies for postgraduate student projects)
Print name (block capitals)……
Date:……

Please send an electronic copy and a signed paper copy of this completed checklist, together with any supporting documentation, to the following addresses, keeping a copy for yourself:

Secretary of the Social Sciences and Humanities IDREC Email: ethics@socsci.ox.ac.uk
University of Oxford Social Sciences Division
Hayes House, George Street
Oxford, OX1 2BQ

Forms may be sent by email (without signature), where both the note of submission from the researcher and the note of endorsement from the supervisor/Head of Department are sent from a University of Oxford email address.

IDRECs and/or CUREC will review a sample of completed checklists and may ask for further details of any project.

FINAL CHECK

To prevent delay please check each of the following before submitting the application.

Have you completed Sections A-E? √

Have you defined all technical terms and abbreviations used? √

Have you included any supporting documentation, including – as appropriate – questionnaires and participant information, consent forms/form or note of procedure for recording oral consent, advertisements and surveys to be used?

Are all pages (including appendices and attachments) numbered? √

Are all relevant declarations in Section F complete and any necessary authorisations obtained (by email or by signing the form)? √

Appendix 6 Participant Information Sheet and Consent Form

Participant Information Sheet

I am a doctoral student at the Department of Education, University of Oxford, conducting research on quality assurance in Chinese higher education to explore factors that motivate university teachers for further improvement of their academic work. My research is based on interviews. Interviewees will be selected from teachers from chosen case universities. In the interview, I will ask you what are the factors generated by students, colleagues, yourself or external measures have an influence on you.

As you are also taking the leadership role in this department/ college, I would like to know what are the policies in place in this department/college to facilitate teachers' further improvement of their academic work. [53]

The research will be written up as a doctoral thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis may be published as a book or articles for academic journals.

To protect interviewees' rights (including privacy), the University of Oxford requires researchers to adopt a set of strict ethical procedures under the supervision of the University of Oxford Central University Research Ethics Committee. Following these procedures, your identity will remain anonymous. Every single interviewee's name, department, and universities will not be revealed at any stage in the research process. In my doctoral thesis or future publications, the names of the universities, the names of the departments, and the names of the interviewees will be replaced by alphabetic letters (such as A, B, C, D) and/or numbers (such as 1, 2, 3, 4). The data, including notes and audio records, will be treated confidentially and stored securely with only me having access to the data.

With your permission, I will audiotape our interview. This will allow me to transcribe our conversation in order to make the most of your insight. I hope you can allow me to audio record our interview. In addition, as part of the university's ethical procedures, the researcher as well as the interviewee is required to sign a consent form. On the one hand, the consent shows that the interviewee agrees to participate the research; on the other hand, the consent is researcher's promise to protect the interviewee's rights, such as privacy.

This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.

---

[53] This paragraph is the only different between the versions for teachers are taking leadership role and those who are not.

Consent Form

I have read the participant information sheet.

I have had the opportunity to ask questions about the study and have received satisfactory answers to questions, and any additional details requested.

I understand that I may withdraw from the study without penalty at any time by advising the researchers of this decision.

I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.

I understand who will have access to personal data provided and how the data will be stored.

I understand that personal data will be treated confidentially and will be de-identified in order to protect participants' identity.

I understand how to raise concerns about the study.

I agree to participate in this study.

Name (Teacher):

Date:

Signature:

Name (Researcher):

Date:

Signature:

For more information contact:

Shan Huang
Doctoral Student
University of Oxford
Department of Education
15 Norham Gardens
Oxford OX2 6PY
UK
Contact number in UK: 0044 (0)780 9537 636
Contact number in China: 0086 13987116881
Email: Shan.Huang@education.ox.ac.uk

Dr Hubert Ertl
Lecturer in Higher Education
Director of MSc Education in Higher Education
Senior Research Fellow of the ESRC-funded Centre on Skills, Knowledge and Occupational Performance (SKOPE)
Fellow and Tutor for Admissions of Linacre College

University of Oxford
Department of Education
15 Norham Gardens
Oxford OX2 6PY
UK
Contact number: 0044 (0)1865 274044
Email: Hubert.Ertl@education.ox.ac.uk

Appendix 7 The *2004 UTE Scheme*

In the *2004 UTE Scheme*, there are 8 tier-one indicators, 19 tier-two indicators and 45 main observation points (MOP). Among the 19 tier-two indicators, 11 indicators, which are in bold font, are key indicators. For each main observation points, there are criteria for Level A and Level C in the evaluation scheme, but there are no criteria for Level B and Level D. If it is lower than Level A but higher Level C, it is Level B; if it is lower than Level C, then it is Level D. Evaluators have to judge by themselves. The indicators in the *2004 UTE Scheme* are:

Tier 1 Index	Tier 2 Index	Main observation points (MOP)	Weights
1 Guiding Principles of School Operation	1.1 Positioning of School	111 Positioning and planning of school	1.0
	1.2 Thoughts on School Operation	121 Thoughts and concepts of education	0.5
		122 Central position of teaching	0.5
2 Faculty Team	**2.1 Size and Structure of Faculty**	211 Student-faculty ratio	0.3
		212 Overall state of structure and development trend	0.4
		213 Percentage of full-time faculty with master's and Ph.D.degrees	0.3
	2.2 Principal Lecturers	221 Qualifications of principal lecturers	0.3
		222 Teaching by professors and associate professors	0.3
		223 Level of teaching	0.4
3 Teaching Conditions and Utilization	**3.1 Basic Teaching Facilities**	311 State of school building	0.2
		312 State of labs and practicum bases	0.2
		313 State of library	0.2
		314 State of school network building	0.2
		315 Stadium and sports facilities	0.2
	3.2 Teaching Fund	321 Percentage of four kinds of funds among tuition income	0.6
		322 Growth of four kinds of funds per student	0.4
4 Program Construction and Teaching Reform	**4.1 Program**	411 Disciplinary structure and layout	0.5
		412 Cultivation plan	0.5
	4.2 Curriculum	421 Teaching content and curriculum reform	0.3
		422 Textbook construction and selection	0.3
		423 Reforms in teaching method and approach	0.3
		424 Bilingual teaching	0.1
	4.3 Practicum	431 Practicum and practical training	0.4
		432 Contents and system of practical teaching	0.3
		433 Synthetical and designing experiment	0.2
		434 Opening of labs	0.1
5 Teaching Management	5.1 Management Team	511 Structure and quality	0.6

		512 Results in research and practice on teaching management	0.4
	5.2 Quality Control	521 Establishment and implementation of teaching rules and regulations	0.3
		522 Quality standards at major stages of teaching	0.3
		523 Monitoring of teaching quality	0.4
6 Academic Atmosphere	6.1 Teacher Manners	611 Professional ethics and commitment of teachers	1.0
	6.2 Learning Culture	621 Compliance of students to school regulations	0.3
		622 Construction of learning culture, measures and effect of motivating students' initiative	0.3
		623 Extracurricular activities in science, technology, and culture	0.4
7 Learning Outcome	**7.1 Basic Theory and Skills**	711 Students' actual level of basic theory and basic skills	0.7
		712 Students' innovative spirit and practical ability	0.3
	7.2 Undergraduat e Thesis or Design	721 Nature, difficulty, significance of the topic selection, and integrated training	0.5
		722 Quality of thesis or design	0.5
	7.3 Cultivation of Ideological and	731 Ideological and moral quality, and cultural and psychological quality	1.0
	7.4 Physical Education	741 Physical education	1.0
	7.5 Social Reputation	751 Source of students	0.6
		752 Social evaluation	0.4
	7.6 Employment	761 Employment situation	1.0
Feature Program			

Table 33 UTE indicators

(Extracted from an English translation of the original policy document – Circular of the General Office of the Ministry of Education on the Printing and Distribution of the Undergraduate Teaching and Learning Evaluation Plan for Regular Higher Education Institutions (Trial Implementation) published in Chinese Education and Society, 2009, Volume 42 Issue 1.)

There is weight allocated for each main observation point in the evaluation scheme but they play no role in producing the results. The results of the UTE are classified into: excellent, good, acceptable and unacceptable and they are based on the number of A, B, C and D:

Excellent: $A \geq 15$, $C \leq 3$ ($A \geq 9$, $C \leq 1$ for key indicators), $D = 0$, with distinct features

Good: $A + B \geq 15$ ($A + B \geq 9$, $D = 0$ for key indicators), $D \leq 1$, with distinct features

Acceptable: $D \leq 3$ ($D \leq 1$ for key indicators)

Appendix 8 The List of the *Action Plans* of the Quality Project Awards

The action plans of the Excellent Courses Awards, the Distinguished Teacher Awards, the Teaching Groups Awards, the Bilingual Courses Awards, and the Ten Thousand New Textbooks Project are listed in Table 34.

The Excellent Courses Awards	2003	2003 Notice of the General Office of the Ministry of Education on the Issuance of the Excellent Courses Project Implementation Measures
	2004	2004 Notice of the General Office of the Ministry of Education on the Supplementary Regulations of the Excellent Courses Project
	2007	The Notice of the General Office of the Ministry of Education on the application of the 2007 Excellent Courses
	2007	The Selection Criteria of the 2007 Excellent Courses
	2007	2007 Application Form for the Excellent Courses
	2007	2007 Bilingual Courses Permitted by the Ministry of Education and the Ministry of Finance
The Teaching Group Awards	2007	The Notice on Evaluating and Selecting the 2007 National Teaching Groups
	2008	The Notice of the Ministry of Education and the Ministry of Finance to Set Up 2008 National Teaching Groups
The Distinguished Teacher Awards	2007	The Notice of the General Office of the Ministry of Education Organised the Third Distinguished Teacher Awards
	2007	The Selection Criteria of the Third Distinguished Teacher Awards
	2007	Application Form for the Third Distinguished Teacher Awards
	2007	The Decision of the Ministry of Education to Reward the Third Distinguished Teacher Award-winning teachers 2007
The Bilingual Courses Awards	2007	Notice on the Launch of the 2007 Bilingual Course Project
	2007	Notice on the Ministry of Education and the Ministry of Finance's Approval of 2007 Bilingual Courses
	2007	2007 Bilingual Courses Application Form
	2008	2007 Bilingual Courses Permitted by the Ministry of Education and the Ministry of Finance
The Ten Thousand New Textbooks Project	2006	The Notice of the General Office of the Ministry of Education on the Eleventh Five-Year National Textbooks
	2006	Ministry of Education on the Issuance the

		Topics of the Eleventh Five-Year National Textbooks
	2007	Application Form for the Eleventh Five-Year National Textbooks
	2008	The Notice of the General Office of the Ministry of Education on the Supplementary Regulations of the Eleventh Five-Year National Textbooks
	2007	The Notice of the General Office of the Ministry of Education on Announcing the Catalogue for the Selection of the 2007 Excellent Textbooks
	2007	The Notice on the Selection of the 2007 Excellent Textbooks

Table 34 Action Plans of the Quality Project awards