AN INVESTIGATION OF ASSESSMENT PRACTICES IN THE MBA DEGREE PROGRAM IN A CHINESE UNIVERSITY

Gryphon Sou

BS(Eng)(CCU) MAdmin (ACU) DMgt(IMC-SCU)

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

University of Technology Sydney
Australia

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student

Production Note:
Signature removed prior to publication.

ACKNOWLEDGMENT

My sincere thanks are due to the senior management of the Asia International Open University (AIOU) at the Macau Special Administrative Region, which has been generous enough to supply me the data for primary and secondary analysis. I wish to take this opportunity to express my gratitude to Prof. Shan-yu Zhuang, ex-Rector of AIOU as well as Prof. Fredrick Kai-ching Tao, the former Pro Vice Rector of AIOU for the support and stimulation they have afforded.

I would also like to acknowledge the cooperation of the internal and external examiners concerned; and the coordination and assistances of the two registrars of AIOU: Dr. Pinqiu Zhou and Ms. Linda Chan. Without their input, this research project could not be completed. Besides, I appreciate the advice from Prof. Geoff Riordan and, Prof. Bob Pithers, and the late Prof. Alison Lee of University of Technology Sydney (UTS). They have given me valuable comments in due course.

Finally, I am indebted to Prof. Tony Holland and Prof. Liam Morgan of UTS for their supervision on this research project. They have invariably shared his ideas and experiences with me and instigated my inspirations. After all, Prof. Holland and Prof. Morgan have undertaken a painstaking review of my final draft of this dissertation and helped me through the internal and external assessment processes.

Ultimately, I should also express thanks to Dr. Jayne Bye of University of Western Sydney and Dr. Stephen Howlett of The University of Sydney. They, as the External Examiners of my dissertation, offered very insightful and enlightening comments to me in the Examiner's Reports.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

The Master of Business Administration (MBA) degree was introduced into mainland China in 1982 and Renmin University of China was one of the first Chinese business schools to start MBA education. In recent years, there has been an increasing demand for the MBA degree programs for Chinese career managers. However during this time in mainland China, international accreditation of MBA degree programs was not in place and there was no evaluation model utilized for quality assurance of those programs especially their assessment practices. Thus the situation merited further research to develop such a model.

This study initially investigated the assessment practices of a 2-year part-time MBA degree program in China. To complete this program, the candidates had to study 17 core subjects plus 3-out-of-6 elective subjects and conduct an action research project leading to a master thesis and a viva. The assessment practices of these subjects, theses and viva were analyzed both from the perspectives of the candidates and the examiners and this formed the basis of the research and evaluative data.

In this research, the Student Learning Experiences (SLE) and Student Learning Outcomes (SLO) of the MBA candidates were collected as well as the perceptions and views of the internal and external examiners. The "quality" of the subject program through the study of the candidates' theses and performance in the viva against the criteria of the Asia International Open University (Macau) and the assessments made by its examiners were also investigated.

A pilot study by means of Pearson and ANOVA Analysis revealed that the students (n = 1,074) of the subject program had favorable Student Learning Experiences on the learning and research processes but less favorable Student Learning Experiences on the appraisal

process. From 2007 to 2008, the author further sampled 331 candidates from 20 higher education institutions within mainland China. Those candidates, who attended the viva, had the Student Learning Experiences that was studied in this research. From the perspectives of the candidates, the subject program provided favorable Student Learning Experiences particularly the action research and the viva voce.

The author explored the candidates' Student Learning Experiences on the research or assessment processes and analyzed their Student Learning Outcomes through the thesis supervisors and examiners. In the sampling period, the author held 23 focus group interviews with all the available examiners and the registrars. To add value to the interpretive part of this study, the author further analyzed the sampled candidates' Student Learning Experiences and Student Learning Outcomes (n = 331) with the aid of Pearson's Analysis, ANOVA, Dunnett's tD Test, S-N-K Test and Tukey HSD Test (n = 67).

Qualitative and quantitative data analysis disclosed that the less favorable Student Learning Experiences of the candidates was partially attributed to the marking variances of the thesis supervisors and the thesis examiners. The former regarded the assessment task of marking the thesis as a norm-referenced assessment while the latter regarded it as a criteria-referenced or an objective-referenced assessment. Findings of the statistical tests on the marking variances showed that the thesis supervisors were more lenient than the examiners in the assessment task.

Focus group interviews revealed that the examiners wished to maintain key characteristic of formative and summative as well as diagnostic assessment in the marking processes. It is difficult for any assessment to have three characteristics. Nevertheless, it might not be impossible for a management education program to have an integrated assessment program that could do the work of these types of assessment.

This study identified nine areas for improvements. They included (1) Scope of Action Research, (2) Compilation of Thesis, (3) Reflection of Candidates, (4) Pre-Notification of Assessments, (5) Elimination of Marking Variances, (6) Change of Assessment Process, (7) Transformation of Students, (8) Benchmarking of Student Learning Outcomes, and (9) Development of Management Learning Styles.

Based on the results of this research, a program evaluation model for quality assurance of an MBA program is proposed. This model is based on the identified needs to assure the quality of a management education program through robust assessment practices that measure whether the program can achieve the desired Student Learning Outcomes. This model also includes quality assurance tools to measure the Student Learning Experiences in the program. With due consideration of Student Learning Outcomes and Student Learning Experiences, the assessment practices should give a more holistic view of the educational quality of the program.

However, such a model should be applied in an integrated manner. Recommendations for its application are detailed in the last chapter of this dissertation. The proposed model contributes to the wide body of knowledge in "program evaluation"; and also enhances the knowledge base of educators who wish to cultivate effective career managers in mainland China. Thus, this research also makes an important contribution to the professional practices of management education.

I. INTRODUCTION

1.1 Topic and Problematic

The topic of this research is "An Investigation of Assessment Practices in the MBA Degree Program in a Chinese University". It investigates the assessment practices of a 2-year part-time Master of Business Administration (MBA) degree program in the mainland of the People's Republic of China[1]. To complete this program, candidates are required to conduct an action research leading to a master thesis and complete a viva voce.

For the purpose of this research, "assessment practices" are operationally defined as the assessment procedures and processes of the subjects that comprise the program. Student assessment in the various subjects usually comprises of conducting an action research project, the completion of a master thesis and the presentation of a viva voce. The action research and so forth thesis-writing takes three to six months while the viva voce lasts for 30 minutes.

The premise of the conceptual framework in this research is that the "quality" of the action research projects, master theses and viva performance of the MBA candidates collectively reflect the "quality" of teaching and learning in the subject program. However, the candidates' performance does not always reflect teaching quality. Nonetheless, there may be a correlation between the teaching activities and the learning outcomes. It is preferable to have a positive correlation. If affirmative, the subject program can be regarded as a "quality" one and meets the expectation of its candidates.

[1] People's Republic of China (PRC) or China consists of mainland China (the Mainland), Hong Kong Special Administrative Region (HKSAR), Macau Special Administrative Region (MSAR), and Taiwan Province (Taiwan).

People expect quality in all sorts of things in their lives. It is nothing but a perception of what they are looking for in a product or service. In the context of the subject program, teaching and learning activities can be regarded as educational services provided to the learners by the program provider while the qualification can be regarded as an educational product to the learners (customers). Therefore, "quality" of the subject program is operationally defined as the desirable student learning outcomes (SLO) and favorable student learning experiences (SLE) within the program.

Higher Education Funding Council for England (1993, para. 11) asserts that "the quality of teaching and learning in a diverse sector can only be understood in the context of an institution's own aims and objectives". Hence, this research focuses on the assessment practices of the subjects that comprise the program and measures them against the aims and objectives set by the educational program provider. The scope of study in this research covers the following key aspects (HEFCE, 1993):

- ✧ Assessment of student performance against the aims and objectives of the educational program provider.
- ✧ Assessment of student learning experiences and student learning outcomes.
- ✧ Assessment by peer review, that is, internal and external examiners vis-à-vis the thesis supervisor.
- ✧ Combination of internal and external processes in the assessment of the thesis and the viva voce.

An historical review (Sou, 2008) reveals that there are various program evaluation models for educational quality assurance. However in the Mainland, there are no tailor-made program evaluation models for the quality assurance of an MBA degree program (Sou, 2005a-d; Sou & Zhou, 2005a-b). The principles of program evaluation rest on an underlying philosophy of quality, which enables an institution to systematically analyze its system(s) for variance, make decisions on fact, consciously define the customers and stakeholders, and actively seek input from both (Cornesky, et al., 1992; Seymour, 1992).

In program evaluation, students become the focus, student learning experiences and student learning outcomes the concern, and assessment the means by which educators gain feedback about what works and what needs to be improved. Previous researchers (Sou 2006a-c; Sou & Zhou, 2005a-e & 2007; Zhou, 2006; Zhou & Sou, 2005a-b) touch on the students' perceived degree of satisfaction and importance of the subject program and its assessment practices. However, they do not throw much light to the correlation of student learning experiences and student learning outcomes. Furthermore, the assessment practices of AIOU were not the foci of the previous research studies.

Under quality management, an institution seeks to improve the quality of what it uses, does and delivers. The ultimate goal is to enhance program effectiveness in order to improve student learning (Chaffee & Sherr, 1992; Deming, 1982). For the purpose of this research, "program evaluation" will be operationally referred to as any procedures used to estimate student learning for quality assurance purposes (Brown, et al., 1997). As there is no evaluation model for quality assurance of a management education program in the Mainland, the situation merits further research.

1.2 Research Question

In this research, I aim to study the student learning experiences and student learning outcomes of the candidates through an investigation of the experiences and views of the internal examiners and external examiners of the theses and of the viva voce. During the viva voce, the examiners do not only assess the student learning outcomes. They also have interaction with the candidates who always reflect upon their student learning experiences. Therefore, both the examiners and the candidates of the subject program are untapped data sources for this research.

With the input of the examiners, this research seeks to evaluate the "quality" of the subject program through the study of the candidates' theses and performance in the viva. Two terms introduced by Higher Education Funding Council for England (1993) are adapted to produce the following research questions (para. 11 of Circulars 3/93):

1. Quality of Educational Provision

- ✧ Is this (the research leading to a master thesis) a good[2] learning experience for the students in the educational program?
- ✧ Does this (the research leading to a viva) contribute to a good[3] learning experience for the students in the educational program?

[2] With reference to the Aristotelian notion, this word is established on the premises of "good for what and for whom". "Good" is construed on the basis of the students' reflections upon experience. The reflections lead to differing interpretation of "good" according to the purposes, context and the shifting boundaries of experience. "Assessment of educational quality" in terms of "fitness for purpose" is conceptualized by this belief.

[3] Ditto

2. Fitness for Purpose

- ✧ How well does the educational experience (performance in the master research and the viva) of the students match the claims of the educational program provider?

Higher Education Funding Council for Education (HEFCE) uses the words "quality approved" for the higher institutions which get a minimum score of two or more out of four on each of the six aspects[4] of educational provision. In the light of the available resources, it is not practical to study all of them in a single individual research project. Fortunately, previous research in mainland China (Sou, 2005a-d & 2006a-c; Sou & Zhou, 2005a-e & 2007; Zhou 2006; Zhou & Sou, 2005a-b) has touched on some of those aspects which could become the foundational framework of this research.

To ensure the manageability of this research, focus of the study should be placed on "Assessment of Student Learning Experiences cum Student Learning Outcomes" that has not been well covered in the previous research. That is to say … the correlation of favorable student learning experiences and the desirable student learning outcomes should be examined. Obviously, there may be a disconnection between the students learning experiences and the examiners. Student learning experiences can only be investigated through the candidates of the subject program.

[4] (1) curriculum design, content and organization, (2) learning resources, (3) teaching, learning and assessment, (4) student progression and achievement, (5) student support and guidance, and (6) quality assurance and enhancement.

As mentioned in the preceding paragraphs, previous research projects (Sou 2006a-c; Sou & Zhou, 2005a-e & 2007; Zhou, 2006; Zhou & Sou, 2005a-b) provide foundational information about the student learning experiences of the subject program. In this research, I am looking at how good are the assessment practices and checking whether they match the learning outcomes. There should be a link between the students' view and the examiners' view. To seek a balanced view and thorough findings, I would question the candidates of the subject program on the conclusion at the end of this research. Thus, the link between the students and the examiners can be established.

1.3 Rationale for Research Question

Higher Education Funding Council for England (1993) advocates that it will be more reliable to assess student learning experiences with the input of the professional educators. In the subject program, the internal and external examiners are in the best position to have input to the investigation of the assessment practices as they examine the theses produced by the MBA candidates and conduct a 30-minute viva for the award of MBA degree.

Together, the performance of the MBA candidates in their Master research and viva reflects the student learning outcomes. The internal or external examiners can judge how well the student learning experiences and consequently the student learning outcomes match the aims and objectives of the subject program, that is, cultivation of career managers. The question addressing student learning outcomes is "How do the candidates apply the integrated professional knowledge and skills gained from the subject program in their research projects?".

This research will enable an understanding of "how well" the student learning experiences match the claims of educational program provider. Furthermore, direct measurement of student learning outcomes can tell "in what way" the students' performance matches the claim of the educational program provider in cultivating career managers. Ultimately, readers of this dissertation can learn whether a quality MBA degree program with "Fitness for Purpose" (HEFCE, 1993) has been provided to the students. In this context, "Fitness for purpose" means the examination of the links between particular student learning experiences and specific objectives of the subject program.

A sound assessment of student learning outcomes can facilitate an effective program evaluation for quality assurance. A reliable, valid and justified assessment of student learning outcomes can reflect the quality of an education program. Quality education does not only benefit the students. It also benefits the trainer, the employer of the students, the institution, and the authority. Other than the benefits to stakeholders, this research could contribute to both the wide body of knowledge and the professional practice of education.

II. BACKGROUND AND CONTEXT

2.1 Management Education in Global Context

In this chapter, the general situation of management education in the global context of educational quality assurance will be introduced. There is an international advocate for the MBA degree programs called The Association of MBAs (the Association). The Association believes that it is essential to have a system for independent quality assurance with so many postgraduate management programs on offer worldwide. This concept forms the basis of the accreditation process that was established in the early 1980s (The Association of MBAs, 2004).

The accreditation is international, independent, market-driven, program-specific, developmental and judgmental; and is recognized as the global standard for all MBA programs. The Association measures the characteristics of an institution and its MBA degree programs against a set of criteria established by an International Accreditation Advisory Board. The senior academics and corporate representatives of the International Accreditation Advisory Board constantly monitor the accreditation criteria to keep in pace with the changes in business and management. In short, the accreditation process is rigorous and has become internationally recognized as a benchmark and "best practice".

Obviously, there is a global belief that quality MBA provision can be assured through accreditation (The Association of MBAs, 2004). This belief is reflected by over 130 member institutions in 65 countries accredited by the Association. However, this belief is advocated by some people and on the contrary some people have counter argument about accreditation. In the ensuing paragraphs, there would be discussion about this belief or assertion.

Fundamentally, educational accreditation is a type of voluntary quality assurance process under which services and operations of educational institutions or programs are evaluated by an external body to determine if applicable standards are met. In most countries of the world, the function of educational accreditation is conducted by a government organization, such as a ministry of education (Lenn, 1992). Hence, some countries do not support the existence of a quality assurance process that is independent of government and performed by private membership associations.

Besides, people sometimes question that the accreditation practices are typically meaning that the associations are competent to test and certify third parties, behave ethically and employ suitable quality assurance. In such circumstances, people would further question who else can ensure the quality of the associations that issue credentials or certify third parties?

Thus, the accreditation practices may cause the associations to be formally accredited by other accreditation bodies. As such, the accreditation practices may become rather complicated. Hence, educational accreditation by private and even global membership associations may not necessarily a good choice for educational quality assurance when the accreditation practices turn to be cumbersome. Understandably in mainland China, only two institutions sought accreditation by the Association of MBAs (The Association of MBAs, 2004).

In contrast, the Chinese institutions offering MBA degree programs are accredited by the China Ministry of Education. There may be wide range of explanations for the Chinese institutions not seeking accreditation by the Association. Perhaps, the global standard and accreditation criteria of the Association may not be universally applicable to the MBA provision in mainland China (Sou, 2005e-f & 2005h). After all, the China Ministry of

Education has already played an active role in the accreditation of the higher education programs in mainland China (Sou, 2005h).

When the accreditation criteria of the Association require an intensive study of the purpose, outcomes, curriculum and assessment scheme of an MBA degree program (The Association of MBAs, 2004), they may be regarded as a benchmark in this research. In this research, the purpose, outcomes, curriculum and assessment scheme of the subject program will be investigated. These areas to be investigated match the global context of management educational quality assurance.

After looking at the broad global context of quality assurance in MBA education, it will be informative to look at the management education needs of mainland China and how the MBA degree program situates in the Chinese educational market. It will also be enlightening to contrast how the institutions in mainland China view educational quality assurance through national certification instead of international accreditation.

2.2 Management Education in Mainland China

In the first instance, a historical review traces the root of MBA degree in mainland China and surfaces the prevailing needs of management education. In 1982, Deng Xiao-ping, the late ex-Premier of the PRC was the first to promote the introduction of MBA degree programs in mainland China. In the last two decades, there has been significant development of management education for the career managers in the Mainland.

In recent years, there has been an increasing demand for management education. Some large enterprises and international companies even require their employees at management level to possess MBA degrees. Thus, career managers seek management education such as part-time MBA degree programs to meet their occupational needs. In the Beijing

Youth Daily dated September 15, 2003, it was reported that there would be a pressing demand for hundreds of thousands of MBA graduates in the Mainland's management field. (Beijing Youth Daily, 2003)

The MBA degree program primarily aims to boost the target learners' individual and occupational qualities. It also aims at enhancing the target learners' power of realization, renovation and entrepreneurship. Its goal is to cultivate the target learners who are working in commerce or industry and pursuing key posts of management, that is, career managers. Considering the community and personal needs, it is both important and timely to study the quality of MBA degree programs provided in the Mainland.

In the Mainland, MBA degree programs are offered by different higher education institutions. They vary in terms of the curriculum design and program specialization. Some of them have attracted criticism on the basis of their incompatibility with the career managers in the field of business administration (Beijing Youth Daily, 2003; A-Performers Editorial Team A, 2009 & 2010). Such matters might be attributed to the historical background of the higher institutions in the Mainland.

In the past, there were few Schools of Business in the Chinese higher institutions. Instead, the Chinese higher institutions always had Schools of Economics to support the national policy of planned economies (China Education Center, 2011; Gale Group, 1999). Many Schools of Business in higher institutions originated from their Schools of Economics. Academic staffs in these faculties were economic scholars who might have laid much emphasis on economic theories and overlooked the educational quality requirements of an MBA degree program (Beijing Youth Daily, 2003).

In recent years, international universities migrated their MBA degree programs to the Mainland. There have been some changes in the quality perspectives of the existing MBA degree programs offered by the Chinese higher institutions with strong background of economic scholars (China Education Center, 2011; Gale Group, 1999). Some institutions recognize that the quality MBA degree program should bear the global and contemporary perspectives of business administration (Beijing Youth Daily, 2003; A-Performers Editorial Team A, 2009 & 2010).

2.3 MBA of Asia International Open University (Macau)

Asia International Open University (Macau) (currently known as City University of Macau; here-in-after referred to as AIOU) has been offering an MBA degree program since 1988. In collaboration with about 30 Chinese higher institutions, AIOU extends its part-time MBA degree program to mainland China for prospective career managers. Stemming from its experience in management education, AIOU strives to provide quality education to prospective career managers in the Mainland. Hence, AIOU is an appropriate subject of study in evaluation of an MBA degree program for career managers in mainland China.

The part-time MBA degree program of AIOU is a general master's program without specialization. Apart from the general MBA program, some specialty programs such as MBA (Finance), MBA (Marketing), MBA (Logistics) become as just popular amongst the career managers who seek to acquire management education to address specific needs in their roles or industries (A-Performers Editorial Team A, 2009). In contrast, the AIOU's MBA general degree program is characterized by the following features and modes:

- ✧ Faculty – Both local and expatriate faculty staff are involved.
- ✧ Training Materials – Chinese textbooks and references are used with international training aids.
- ✧ Learning Mode – 2-year part-time studies in 17 core subjects[5] and 3-out-of-6 electives[6].
- ✧ Teaching Mode – Action learning (Revans, 1980; Sou, 2000 & 2001a-c) is encouraged.
- ✧ Research Mode – Action research is conducted for the compilation of a master thesis.
- ✧ Appraisal Mode – Viva voce is assessed by an academic panel of internal and external examiners.

Having completed 17 core subjects and 3-out-of-6 electives, candidates are required to complete an action research project leading to a master thesis. Those theses are respectively marked by the responsible supervisor, an internal and an external examiner. Having obtained a pass mark from the supervisor, the candidates have to present a viva voce before the internal and the external examiners for the award. A viva voce for an MBA candidate normally lasts for 30 minutes. Its objectives are: (1) authentication of the MBA thesis, (2) identification of learning outcomes, (3) proliferation of professional knowledge, and (4) determination of final grading.

---

[5] Core Subjects: (1) Business Accounting for Executives; (2) Organization and Management for Business; (3) Financial Management; (4) Quantitative Methods Analysis; (5) Production Management; (6) Information Systems; (7) Economics for Business; (8) Organization Behavior; (9) Human Resources Management; (10) Marketing Management; (11) Cost Accounting; (12) Small Business Management; (13) Strategic Management of Change; (14) Investment; (15) Business Law; (16) Business Strategic Management; and (17) Corporate Recovery.

[6] Electives: (1) Selected Topics in Advanced Finance; (2) Auditing; (3) Marketing Planning & Strategy; (4) Comparative Management; (5) Total Quality Management; or (6) Taxation.

Having passed the viva voce, the MBA candidate would be entitled to the award of AIOU's MBA degree. Simultaneously, the holder of an AIOU's MBA degree can apply for a "Chinese Career Manager Certificate". Chinese Career Manager Certificate is a professional title granted by the National Certification Committee of the Chinese Career Managers, People's Republic of China. It is a widely recognized vocational qualification for the recruitment, professional employment and development of career managers in mainland China.

To this end, it comes to light that the international accreditation of MBA education like the global assurance scheme of The Association of MBAs is not in place in mainland China. Instead, a national certification scheme is in force. AIOU is offering a dual qualifications program of MBA and Chinese Career Manager Certificate to the prospective career managers in the Mainland. Prospective career managers through AIOU's program are working towards the national certification whilst pursuing MBA education.

III. CONCEPTUAL FRAMEWORK AND LITERATURE REVIEW

3.1 Overview of Conceptual Framework

Figure 1 shows the conceptual framework of literature review. When program evaluation is for quality assurance, is it possible to regard effective assessment practices, favorable student learning experiences and desirable student learning outcomes as the evaluative components of a sound program evaluation model? This chapter will address this question by exploring theories relevant to effective evaluation frameworks.

Figure 1: Conceptual Framework of Literature Review

(Source: Sou, 2011)

Based on the above framework, I will first explore the underlying principles of program evaluation and study some principal evaluation approaches (Section 3.2). Amongst those approaches, I will make reference to America which has a long history of program evaluation, particularly in higher education (Section 3.3).

Then, I will try to compare some program evaluation models and look for their evaluative components (Section 3.4). In the search of the evaluative components, I will find out the prerequisites of a sound program evaluation model and identify the appropriate assessment methods suitable for the educational quality assurance of a management education program.

With the application of the appropriate assessment methods (Section 3.5), I will analyze the interrelationship of the evaluative components of a sound program evaluation model (Figure 1). Besides, I would quote the application of an international quality management system that is once used for the assessment practices in the management education of mainland China.

With the reference of the previous studies in mainland China (Section 3.6), I will elaborate the findings of those studies in respect of the proposed evaluative components of a program evaluation model (Figure 1), namely (a) Assessment Practices. (b) Student Learning Outcomes, and (c) Student Learning Experiences. At the end of this Chapter, I will sum up the literature review and report the theoretical preposition (Section 3.7).

3.2 Program Evaluation for Quality Assurance

It comes as no surprise that no specific program evaluation model has been developed in China, due to the short history of management education in the form of MBA since 1980s (Sou, 2006a). Thus, it will be productive to review the established approaches of program

evaluation for this research. After a historical review (Sou 2006a & 2008; Sou & Zhou, 2007), three classical approaches mainly from the late 1800s to 1960s with significance come to light:

1. Accreditation or Professional Judgment Approach
2. Testing Approach
3. Tylerian Approach

3.2.1 Accreditation or Professional Judgment Approach

The first one to be reviewed is the "Accreditation or Professional Judgment Approach". This approach plays a powerful role in educational quality assurance. It enables the public to discriminate between institutions and programs with the aid of professional judgment. However, this approach focuses on the organizational inputs such as resources, facilities or staff instead of the processes or products such as student learning experiences or student learning outcomes.

Madaus and Stufflebeam (1988) comment that professional judgments are notoriously unreliable and provided inadequate feedback. Considering this shortcoming, this approach is incompatible with the research described in this thesis. Above all, the pilot study revealed there are no accreditation bodies specifically established for the MBA degree programs in mainland China.

3.2.2 Testing Approach

The second one to be reviewed is "Testing Approach". This approach has made important contributions to educational quality evaluation. However, testing data are often regarded as the principal or only data source in such evaluations. Naturally, tests can become so important that they eventually have a negative impact on teaching and learning.

Undoubtedly, the Testing Approach could be used to assess performance but it does not contribute to the identification or implementation of remedial strategies.

Obviously, the Testing Approach equates program evaluation to test results. If so, it may be prone to reduced enthusiasm for quality assurance (Sou, 2008) when the rationale for educational change rests upon the test scores of the students. Nonetheless, this approach has a long history in its use in educational evaluation. Tests, either qualitative or quantitative, have been continually used from 1444 to 2000 (Sou, 2007).

It is true that a testing program does not constitute a comprehensive evaluation program (Madaus & Stufflebeam, 1988; Stufflebeam et al., 2000). However, the proper role of testing in educational quality assurance could be an essential part of a comprehensive program evaluation. Thus, it will be constructive to study "testing" under the form of "assessment practices" in this research.

3.2.3 Tylerian Approach

The last evaluation approach to be reviewed is the "Tylerian Approach". Being the father of educational evaluation, Tyler's conceptualization has enormous influence on education in general and education evaluation in particular (Madaus & Stufflebeam, 1988; Stufflebeam, 2000). The Tylerian Approach has a clear-cut advantage over the other two classical approaches. The Tylerian Approach concentrates on student learning outcomes (performance against objectives) instead of organizational inputs, thus it avoids subjectivity of the Accreditation or Professional Judgment Approach (Stufflebeam & Shinkfield, 1984; Stufflebeam et al., 1985).

From an epistemological standpoint, evaluation approaches can be categorized as manifesting subjectivism and objectivism. Subjectivism relies on accumulated experiences

of the evaluator rather than a scientific method. In other words, it depends on the internalized experiences of the evaluator as the basis of understanding. The validity of subjectivist evaluation lies not in data replication and verification, but the relevance of the evaluator's background, qualifications, and keenness of his or her perceptions. (House, 1980, 1983a, 1983b & 2003)

In contrast with subjectivism, objectivism is drawn primarily from the empirical tradition within the social sciences. It requires evaluation information to be scientifically objective. Thus, data collection and analytical techniques produce results that are replicable and verifiable by other reasonable and competent individuals using the same techniques and instruments. In other words, objectivist evaluative approach depends on replicable facts as the basis of knowledge and truth. (House, 1980, 1983a, 1983b & 2003)

With its objectivist evaluation, the Tylerian Approach does not rely heavily on the test scores of the students as the Testing Approach, which has the characteristics of criteria-referenced assessment. Tylerian Approach is an objectives-oriented evaluation approach. Tyler (1942 & 1950) conceives evaluation as the process of determining the extent to which the educational objectives of a school, program, or curriculum are actually attained. Tylerian Approach consists of seven steps in program evaluation:

1. Establishing broad goals or objectives
2. Classifying the goals or objectives
3. Defining objectives in behavioral terms
4. Identifying situations in which achievements of objectives can be demonstrated
5. Selecting or Developing measurement techniques
6. Collecting performance data
7. Comparing performance data with behaviorally stated objectives

Tyler (1942 & 1950) further advocates using general goals to establish purposes, rather than prematurely becoming preoccupied with formulating behavioral objectives, resulting in program focused on arbitrary and restrictive objectives that do not support their underlying purpose or attainment of credible learning outcomes. Over the years, practitioners of program evaluation have further refined Tylerian Approach by adding flexibility and insight to programmatic evaluation within the Tylerian Rationale (Tyler, 1942 & 1950; House, 1980, 1983a, 1983b & 2003):

1. What educational purposes should the school seek to attain?
2. How can learning experiences be selected which are likely to be useful in attaining these objectives?
3. How can learning experiences be organized for effective instruction?
4. How can the effectiveness of learning experiences be evaluated?

Tyler (1950) describes learning as taking place through the action of the student, "It is what he does that he learns, not what the teacher does" (p. 63). In this research, the Tylerian Rationale is employed with the integration of some conventional evaluation approaches that will be explained in the latter part of this chapter. In the following chapters, I also attempt to answer the foregoing questions as follows:

1. Defining appropriate learning objectives of AIOU's MBA degree program.
2. Introducing useful student learning experiences in the subject program.
3. Organizing student learning experiences to maximize their effect.
4. Evaluating the process and revising the areas that are not effective.

Furthermore, the Tylerian Approach requires the evaluator to (a) identify situations in which achievements of objectives can be demonstrated, (b) select or develop measurement techniques, and (c) collect performance data and comparing performance data with behaviorally stated objectives. Therefore, I am required to study the assessment practices of the subject program. Through a systematic study on the assessment practices, I would understand:

- ✧ how AIOU demonstrates the achievements of the learning objectives of the subject program;
- ✧ how AIOU measures the performance of the learners of the subject program; and
- ✧ how AIOU compares the student learning outcomes with the learning objectives.

3.2.4 Educational Quality Assurance through Program Evaluation

The study of these three established approaches (Accreditation or Professional Judgment Approach, Testing Approach and Tylerian Approach) revealed that quality assurance is not a commonly used phrase within the vast lexicon of educational jargon (Madaus & Stufflebeam, 1984; Sou, 2008) whereas program evaluation is often mistakenly viewed as a recent phenomenon (Stufflebeam, 2000). Interestingly, the idea of "educational quality assurance through program evaluation" is being regarded as having a long history in education – over 560 years, firstly in Europe, and then in the United Kingdom and the United States (Sou & Zhou, 2008).

However, program evaluation is still new to China (Sou, 2008) when most universities in mainland China do not have a quality unit which helps the institution assure educational quality. It is noteworthy that two concepts of educational quality assurance have been identified in the historical review (Madaus & Stufflebeam, 1984; Stufflebeam et al., 2000; Sou, 2008) initially in the West and gradually in the East, that is (1) Accountability, and (2)

Program Evaluation. Accountability refers to individuals such as students, teachers or administrators who are held accountable for their actions. Accountability involves the assessment of individual performance which is usually linked with test scores.

Program evaluation refers to the quality of the educational offerings, such as efficiency, effectiveness or achievement of objectives. Program Evaluation involves the commitments of many publics or stakeholders such as students, teachers, administrators, evaluators or policy makers. It aims to provide ongoing feedback to them who would initiate corrective actions or innovative changes for continual improvement of the educational programs. (Shinkfield & Stufflebeam, 1996; Stufflebeam et al., 1985; Kellaghan & Stufflebeam, 2002)

Program Evaluation has a delicate relationship with accountability (Madaus & Stufflebeam, 1984; Stufflebeam, 2000; Sou, 2008). Both play a role in educational quality assurance. It is important to have program evaluation within a broader accountability system to ensure quality educational offerings. The evaluative component of educational quality assurance should refer to the ongoing assessment practices required of the students and proactive study of the assessment practices aimed at continual improvement. Hence, a study of assessment practices could contribute to the educational quality assurance and also inform future decisions.

The accountability function of educational quality assurance tends to be retrospective and the student assessment tends to be "summative" in nature. The "summative" assessment of the student provides test scores that are taken as an indication of the quality of a program. This in turn is put forward as evidence that the institution has performed the requirements which are an accepted part of education provision. Assessments of individuals by testing can in certain circumstances become an administrative device to bully

the educational system into line (Madaus & Stufflebeam, 1988). Nonetheless, the "summative" function of student assessment is indispensable for a comprehensive evaluation program.

Neither accountability nor program evaluation can stand alone for educational quality assurance. Therefore, assessment practices should have a balance of "summative" and "formative" functions in a program evaluation under an accountability system like national certification. In this research, both the "summative" and "formative" functions of the assessment practices (Carless, et al., 2006) should be studied in order to determine whether the subject program is a quality one under the accountability system.

To streamline the flow of this literature review, there will be an elaboration of "summative" and "formative" functions for educational evaluation in section 3.5 viz. "Assessment Practices". Prior to this section however, the review will turn to explore the concepts of "quality assurance of higher education" (section 3.3) and "program evaluation models" (section 3.4).

3.3 Quality Assurance of Higher Education

In this section, there are two main areas to be explored. The first area to be explored is the international experiences pertaining to educational quality assurance. As the MBA degree programs in mainland China originated from programs in the United States, it is appropriate to consider the rich American experiences in the quality assurance of its higher education. In the 20th century, there were ongoing assurance efforts towards the quality of higher education at the state and institutional level in America (Chun, 2002).

3.3.1 Educational Quality Assessment in America

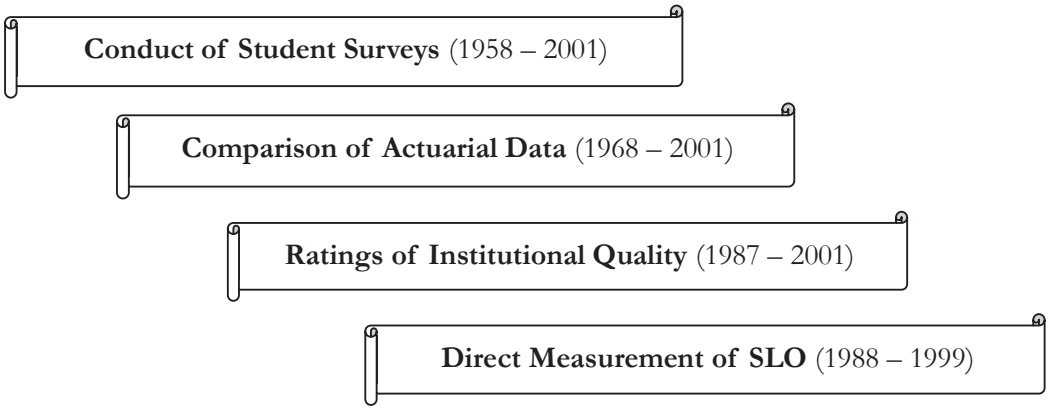
Chun (2002) states that there have been tremendous ongoing efforts at state, institutional and program levels towards quality assurance in American higher education from the 1960s to the millennium. By the end of the 20th century, more than three-quarters of the states in America had some form of higher education policy for educational quality assurance. Nettles et al. (1997) note that little systemic knowledge has been available to measure the extent and scope of publicly mandated "outcomes" assessments.

Peterson et al. (1999a-c) surveyed 1,393 public and private institutions in America and found that 82 percent listed "excellence in undergraduate education" as part of their mission statements. However, 38 percent of those institutions did not conduct studies to link student learning experiences to student learning outcomes. Brown et al. (1997) believe that there are many similarities between the assessment of student learning experiences and educational quality assurance. Rowntree (1987) comments that the truth of an educational system can be discovered through the assessment of student learning experiences and the associated procedures. Along this line of thinking, student learning experiences should have a place in educational quality assurance.

3.3.2 Conventional Approaches of Educational Quality Assessment

The second area to be discussed is the conventional approaches of educational quality assurance that have a basis in program evaluation. Since 2000, American researchers in educational quality assurance have continued their efforts to study student learning experiences and so forth the student learning outcomes (SLO) (Banta et al., 1996; Banta & Palomba, 1999). Chun (2002) and Sou (2006a-c & 2008) discover four conventional approaches (Figure 2) of educational quality assurance that is popular from 1958 to 1999.

Figure 2: Popularity Chronology of Educational Quality Assurance Approaches



(Source: Sou, 2008)

3.3.2.1 Conduct of Student Survey

These four conventional approaches have both pros and cons in terms of their efficacy and practicability in monitoring and measuring educational quality. Firstly, "Conduct of Student Surveys" has a long history in its application but it measures quality according to student perceptions of their learning. This approach may be seen by some to be problematic as it relies on students' self-evaluation (Chun, 2002). Researchers (Baird, 1976; Berdie, 1971; Ouiment, et al., 2001; Pace, 1985 & 1990; Pike, 1995; Pohlman & Beggs, 1974; and Turner & Martin, 1984) have questioned the credibility of the students' self-reports on student learning experiences.

In student surveys, Ouiment et al. (2001) consider that a considerable challenge lies in ascertaining whether or not what students report corresponds to what they have actually experienced. Other researchers (Baird, 1976; Berdie, 1971; Kuh, 2001; Pace 1985; Pike 1995; Pohlman & Beggs 1974; Turner & Martin 1984) identify two problems in the "Conduct of Student Surveys". The first one is that some students are unable to supply accurate information about their learning experiences while the second one is that some students are unwilling to supply accurate information (Aaker, Kumar & Day, 1998;

Wentland & Smith, 1993) in the surveys. Either way affects the efficacy of the data and the subsequent analyses.

3.3.2.2 Comparison of Actuarial Data

Secondly, "Comparison of Actuarial Data" is commonly used because of the ease of collection and the patina of scientific objectivity. For instance, actuarial data (Chun, 2002; Sou, 2008) may include:

- ✧ Administration test scores of entering students
- ✧ Selectivity ratio
- ✧ Graduation rates
- ✧ Racial/Ethnic composition of student body
- ✧ Breadth and depth of academic course offerings
- ✧ Highest degree earned by faculty members
- ✧ Levels of external research funding
- ✧ Level of endowment

However, this approach equates quality with discrete, available and easily measurable indicators of quality, such as counts of people and resources (Chun, 2002; Sou, 2008). Reviews of national data system concluded that this approach yielded little information about an institution's educational quality in terms of student learning outcomes (Dey et al., 1997; NPEC, 2000a-b).

3.3.2.3 Ratings of Institutional Quality

Thirdly, "Ratings of Institutional Quality" rely on a formula that combines "Comparison of Actuarial Data" and ratings by informed experts. These rankings have limitations as they provide only an indirect measurement of quality and conflate quality and reputation. This

approach was seen by some as lacking any defensible empirical or theoretical basis for the weightings used to combine the various measures into an overall rating (NORC, 1987). Researchers criticized the ratings on the grounds that they were sensitive to relatively small changes in the weighting scheme (Machung, 1995 & McGuire, 1995).

3.3.2.4 Direct Measurement of Student Learning Outcomes

Fourthly, "Direct Measurement of Student Learning Outcomes" is a comparatively new approach. Several researchers see this method as having the greatest face validity in terms of assuring quality of higher education (Chun, 2001; Sou, 2008). Some researchers used it for educational quality assurance at programmatic and institutional levels (Winter et al., 1981). Some institutions used it for comparing results amongst themselves (Obler et al., 1993).

However, the fourth approach can be problematic if used to standardize a post-learning test for measuring the student learning outcomes within cognate programs offered by different institutions (Cole et al., 1997, NCHEMS, 1994 & 1996; Steels & Lutz, 1995). Perhaps, the main difficulties arise from the huge obstacles whilst making institutional comparison in contexts where the higher institutions cannot agree on what should be measured (Ratcliff et al., 1991). Moreover, it is cost-prohibitive and perhaps the least systematically used of the conventional approaches (Chun, 2002).

Nevertheless, proponents of this approach stresses on the validity of direct measurement of students' ASK (Abilities, Skills, Knowledge) within the context of participation in an educational program (Johnson et al., 1993 & Lenning, 1988). Thus, a program evaluation model on the premise of "Direct Measurement of Student Learning Outcomes" may offer a systematic means of deploying this approach.

3.4 Program Evaluation Models

In this section, different program evaluation models will be explored with the aim of establishing which of these is or are compatible with the study on the subject program. Within the broad category of conventional approaches of educational quality assurance, many evaluation models have emerged in America since 1960. Chronologically, the development of program evaluation was advocated as a professional practice by a number of scholars. To name a few but not all, those scholars in chronological order include:

> *Tyler, 1942b, 1950 & 1966; Campbell & Stanley, 1963; Cronbach, 1963; Stufflebeam, 1966 & 1967; Scriven, 1967, 1991, 1993, 1994a-c; Schuman, 1967; Alkin, 1969; Guba, 1969; Provus, 1969; Stufflebeam et al., 1971; Parlett & Hamilton, 1972; Weiss, 1972; Eisner, 1975; Glass, 1975; Stake, 1975b, 1988 & 1995; Cronbach & Associates, 1980; House, 1980; Patton, 1980, 1982, 1990, 1994 & 1997; Schwandt, 1984; Smith NL, 1987; Smith MF, 1989; Guba & Lincoln, 1991; Cronbach, 1982; Shinkfield & Stufflebeam, 1985; Sanders & Worthen, 1987 & 1997; Cook, Leviton & Shadish, 1991; Freeman & Rossi, 1993; Nevo, 1993; Hatry, Newcomer & Wholey, 1995; Miech, Mosteller & Nave, 1999.*

3.4.1 Foundational Models of Program Evaluation

Evaluation models were typically developed either at the national and institutional level or built at the program level because of their compatibility with specific programs. In the context of this review, 22 foundational Models of Program Evaluation (Figure 3) in four main categories (Sou & Zhou, 2007; Stufflebeam, 2000) have emerged. To ensure readability of this chapter, there will not be in-depth descriptions of those foundational models that emerged mainly between 1960 and 1999. Instead, their strengths and weaknesses will be considered within a description of whether, when, and how they are best applied in program evaluation.

Figure 3: Twenty-two Foundational Program Evaluation Models in Four Categories

Pseudo-Evaluation Model	**Improvement/Accountability-Oriented Evaluation Model**
1. Public Relations-Inspired Studies 2. Politically Controlled Studies	1. Decision/Accountability-Oriented Studies 2. Consumer-Oriented Studies 3. Accreditation/Certification Approach
Quasi-Evaluation Model	**Social Agenda-Directed/Advocacy Evaluation Model**
1. Objective-Based Studies 2. Objective Testing Programs 3. Accountability, particularly Payment by Results Studies 4. Outcomes Evaluation as Value-Added Assessment 5. Performance Testing 6. Experimental Studies 7. Management Information Systems 8. Benefit-Cost Analysis Approach 9. Clarification Hearing 10. Case Study Evaluations 11. Criticism and Connoisseurship 12. Program Theory-Based Evaluation 13. Mixed-Method Studies	1. Client-Centered Studies / Responsive Evaluation 2. Constructivist Evaluation 3. Deliberative Democratic Evaluation 4. Utilization-Focused Evaluation

(Source: Sou & Zhou, 2007)

3.4.1.1 Pseudo-Evaluation Models

Firstly, Pseudo-Evaluation Models like Public Relations-Inspired Studies and Politically Controlled Studies are often adopted for political objectives. They are used by those in power to persuade constituents or to gain and maintain an unfair advantage over opponents. They may look like evaluations but typically lack real rigor and are therefore aptly termed pseudo-evaluations (Stufflebeam, 2000). Evaluators and their clients are sometimes tempted to shade, selectively release, or even falsify findings (Stufflebeam, 2000).

In other words, these models fail to produce and report a valid evaluation of merit and worth to the stakeholders. They may help promote and support injustice, mislead decision making, lower confidence in evaluation services, and discredit the evaluation profession (Brown, et al., 1997; Chun, 2002; Sou, 2008).

3.4.1.2 Quasi-Evaluation Models

Secondly, Quasi-Evaluation Models like Objective-Based Studies, Objective Testing Programs, Outcomes Evaluation as Value-Added Assessment are adopted to address specified questions and answers which may be sufficient to measure a program's merit and worth (Sou & Zhou, 2007; Stufflebeam, 2000). Their focus can be too narrow or their focus may be tangential to questions relating to the merit and worth of a program. Even if they do happen to provide evidence that fully measure a program's merit and worth, they may not necessarily meet the quality prerequisites of program evaluation (Figure 4). These models have legitimate uses since they can still be used to investigate important but narrow questions. However, they cannot be uncritically equated to sound program evaluation (Sou, 2008; Stufflebeam, 2000).

3.4.1.3 Improvement/Accountability-Oriented Models

Thirdly, Improvement/Accountability-Oriented Models like Consumer-Oriented Studies or Accreditation/Certification Approach are expansive (Sou & Zhou, 2007; Stufflebeam, 2000). They seek comprehensiveness through a consideration of all questions and answers needed to measure a program's merit and worth. They primarily employ the needs of a program's stakeholders as the foundational criteria for the evaluation.

They also examine the technical and economical criteria to measure the program objectives, plans, operations and outcomes. These models are objectivist and assume an underlying reality in seeking definitive, unequivocal answers to the questions keyed to the program evaluation. Usually, multiple qualitative and quantitative methods are used for the cross-checking of their findings.

3.4.1.4 Social Agenda-Directed/Advocacy Models

Fourthly, Social Agenda-Directed/Advocacy Models like Constructivist Evaluation or Utilization-Focused Evaluation are oriented to the perspectives of the stakeholders of a program as well as that of the educational experts (Sou & Zhou, 2007; Stufflebeam, 2000). Their perspectives are mainly built on a constructivist orientation and the use of qualitative methods. These models provide for democratic engagement of stakeholders in obtaining and interpreting findings. They adopt democratic principles of equity and fairness.

To exercise sufficient bias controls within this model, there should also be adequate safeguards such as meta-evaluation strategies grounded in the original evaluation standards (Sou & Zhou, 2007; Stufflebeam, 2000). In this context, meta-evaluation is a means of program evaluation involving all stakeholders of the program concerned, input of experts and utilization of comprehensive methods.

3.4.2 Prerequisites of a Program Evaluation Model

These models prioritize experts' input in characterizing, investigating, and judging programs. Unfortunately, they may focus so heavily on the social agenda that they fail to meet the quality prerequisites of program evaluation (Sou, 2008; Stufflebeam, 2000) (Figure 4). After all, these models also empower stakeholders to make evaluation decisions and this can make program evaluation vulnerable to bias and other misuse.

When program evaluation is used to serve the underprivileged or disenfranchised or uphold issues characterized as involving educational "fairness", evaluators tend to compromise the veracity needs to produce desirable but probably invalid findings (Sou & Zhou, 2007; Stufflebeam, 2000). Toporek et al. (2006) assert that veracity in intentions for many countries has been questionable. They find that there may be historically some distrust of researchers from various disciplines given the exploitative and harmful use of research results or the research process.

Figure 4: Quality Prerequisites of Program Evaluation

(Source: Sou, 2008 & 2011)

3.4.2.1 Development of Various Program Evaluation Methods

Since 1973, evaluators have increasingly realized that program evaluation (Stufflebeam, 2000) should have some perquisites in terms of quality assurance. It was impossible for the evaluators to develop a fully functional methodology meeting all the prerequisites of program evaluation (Figure 4) by the end of 1983. However, there were promising developments of various program evaluation methods from 1970s to 1980s (Sou, 2008).

For instance, those program evaluation methods are called (1) Advocate Teams (Reinhard, 1972), (2) Adversary-Advocacy Teams (Stake & Gkerde, 1974), (3) Goal-Free Evaluation (Evers, 1980; Scriven, 1974), (4) Artistic Evaluation (Eisner, 1975), (5) Naturalistic Evaluation (Guba & Lincoln, 1975), (6) Responsive Evaluation (Stake, 1975), (7) Meta-Analysis (Glass, 1976, Krol, 1978), (8) Needs Assessment (Roth, 1977; Scriven & Roth, 1977; Stufflebeam, 1977; Suarez, 1980), (9) Case Study Method (Stake, 1978), (10) Evaluative Criteria Catalogs and Instruments (Bunda, 1980; Ridings, 1980), (11) Ongoing Curriculum Objectives-Referenced Assessment (Bloom et al., 1981; Chase, 1980), and (12) Panel Investigation (Smith, 1981a & b). Application of various assessment methods leads to the development of more alternative program evaluation models.

3.4.2.2 Alternative Program Evaluation Methods

Apart from the 22 foundational models (Figure 3), there are 17 alternative evaluation models (Kellaghan & Stufflebeam, 2002; Sou, 2008; Stufflebeam et al., 2000) which are characterized by their developers' conceptualization of evaluation and each is provided with guidelines for use in program evaluations. Those alternative evaluation models are named as: (1) Gardening Model, (2) Medicine Model, (3) Factory Model, (4) Legal Model, (5) Naturalist Evaluation Model, (6) Inquiry Oriented Model, (7) Merit Oriented Model, (8) Responsive Model, (9) Evaluation as Persuasion Model, (10) Empowerment Evaluation

Model, (11) Illuminative Model, (12) Expository Storytelling Model, (13) Internal Evaluation Model, (14) External Evaluation Model, (15) Formative/Summative Evaluation Model, (16) Social Science Theory Based Model, and (17) Management Theory Based Model (Madaus & Kellaghan, 2000; Stufflebeam et al., 2000).

Occasionally, it is practicable to mix and match the foundational models (Section 3.4.1.1 to Section 3.4.1.4) and/or alternative models (Section 3.4.2.2) for the purpose of program evaluation (Madaus & Kellaghan, 2000; Sou & Zhou, 2007; Stufflebeam et al., 2000). Mixed application of these models would produce several "hybrid models" to suit the quality assurance needs of different educational programs. Nevertheless, none of these models are tailor-made for the quality assurance of an MBA degree program.

Hence, the prerequisites of program evaluation (Figure 4) can be adopted to develop a fully functional program evaluation model for the MBA degree program that is the focus of this thesis instead of applying the foundational or alternative models. In the data analysis, the prerequisites of program evaluation can be used as yardsticks. Findings of the analysis can tell whether the prototype of a program evaluation model as illustrated in Figure 1 is quality defensible.

3.5 Assessment Practices

Program evaluation can be carried out in relation to specific student assessment practices. Assessment of students in a management education program can take many forms It can be formal or informal in nature, that is, for organizational or personal use. It can be conducted internally by the staff or externally by the outsiders. It can be formative or summative. In such circumstances, what is the most effective form of assessment practices for the subject program?

The answer to this question is determined by the purpose of the student assessment practices. There is a range of assessment methods available to fulfill any purpose of assessment. The match between assessment methods, purposes and learning outcomes is more subtle than it appears. If the primary purpose of the assessment is not only "summative" but also provision of feedback to the students, different criteria and comments may be needed.

3.5.1 Match of Assessment Methods, Purposes and Learning Outcomes

The match of assessment methods, purposes and learning outcomes becomes more subtle when there are different levels of skills generating different levels of outcomes such as class, modular, program or institutional. Student assessment at the program level may be the most subtle task. In management education programs, it is unlikely that there is a one-to-one correspondence (Brown, et al., 1997) between an assessment task and an outcome at the module and program level.

Brown et al. (1997) put forward the view that there is a high probability of over-assessment if indeed there is a one-to-one correspondence. Nonetheless, it is still possible to design assessment tasks that test a set of overlapping outcomes at the class, modular and program levels. Various assessment methods comprising multiple tasks are employed to test overlapping student learning outcomes.

Broadly speaking, there are 13 assessment methods (Brown et al., 1997; Sou, 2008) of student learning outcomes: (1) Multiple-Choice Question, (2) Short Answer Question, (3) Single Essay Question, (4) Essay, (5) Mini-Practical, (6) Report on Practical, (7) Reflective Practice Assignment, (8) Problem, (9) Cases and Open Question, (10) Project, Group Project and Dissertation, (11) Oral, (12) Presentation, and (13) Poster Session (Brown et al., 1997).

3.5.2 Application of Assessment Methods in the MBA Program under Review

The assessment practices of the subject program constitutes about half of these assessment methods. The following paragraphs (Section 3.5.2.1 to Section 3.5.2.7) are adapted from commentaries (Brown et al., 1997, pp. 46-47) on those methods communally used in the subject program.

3.5.2.1 "Problem" as Assessment Task

Firstly, selection of a burning issue in the candidate's job setting for action research is regarded as "Problem" that has potential for measuring the candidate's application, analysis and problem-solving strategies. Complex problem solving exercises and their making schemes can be difficult to design. Marking for grading of easy and similar problems is fast whereas complex and different problems marking can be slow (Brown et al., 1997). Marking for feedback can be slow (Brown et al., 1997). Variation of similar problem marking between examiners is usually low (Brown et al., 1997).

3.5.2.2 "Case and Open Problem" as Assessment Task

Secondly, dynamic management case study in the action research can be regarded as "Cases and Open Problem" which has potential for measuring the candidate's application of knowledge, analysis, problem-solving and evaluative skills. Short cases are relatively easy to design and mark. Design of more complex cases and their marking schemes is more challenging. Marking for grading and feedback are about as fast as "Essay". However, the marking for grading and feedback will become more complicated when the cases or open problems are real-life but not pre-designed.

3.5.2.3 "Project or Group Project" as Assessment Task

Thirdly, action research in a corporate or organizational environment becomes a "Project or Group Project" of the company, corporation or organization in which the candidate is working (Revans, 1980; Sou, 2000 & 2001a-c). When the candidate is engaged in a management position, he or she can choose a real-life management project for his or her action research. The candidate, as a manager, could even form a project group within his or her company to conduct the action research. The action research, being a "Project or Group Project", can be regarded as a good all-round ability testing. Above all, it has great potential for sampling a wide range of practical, analytical and interpretative skills of the candidate.

Eventually, "Project or Group Project" evidences the candidate's wider application of knowledge, understanding and skill to real or simulated situations (Sou, 2008). A candidate may utilize the resources of his or her company to complete the action research project for the subject program. The candidate may form a project team within his or her company which works for a period and tackle a real-life management problem that their company is facing.

In part-time MBA degree programs, the candidates are usually in management positions which enable them to deploy the manpower and resources of their companies to solve a management problem (Beijing Youth Daily, 2003). For instance, a candidate may start his or her action research project on the marketing strategy of the company. The candidate would summon assistance from his or her subordinates in the marketing department, finance department and human resources department. With their input and contribution, the candidate can complete the action research project and compile his or her master thesis.

"Action research projects in real situations" provide a measure of project and time management. A "Group Project submitted by the Candidate as Action Research" can also provide a measure of teamwork skills and leadership. There is feedback potential in the course of conducting the action research and compiling the master thesis. The thesis supervisor can make "formative assessment" on the candidate in due course. Upon completion of the master thesis, the examiners marking the thesis can make "summative assessment" on the candidate. However, there may be variations between examiners because the examiners are required to mark multifarious research projects (Sou, 2008).

3.5.2.4 "Essay" as Assessment Task

Fourthly, the master thesis is an "Essay" which has potential for measuring understanding, synthesis and evaluative skills of the candidates. It is relatively easy to set. Marking for grading using impressionistic marking is relatively fast while marking for feedback can be time-consuming. Impressionistic mark is given on the basis of the examiner's feeling for the candidate's overall performance. Impressionistic marking is the least useful because it provides little "formative" information as the basis for the candidates' action (Rea-Dickins & Germaine, 2003) in the action research projects. Above all, variations between examiners can be high due to the subjectivity. Subjectivity occurs when the examiner is using his or her own opinions and judgment rather than relying on a more objective measure.

The remedy for subjective impressionistic marking may be "Structured Impressionistic Marking" (Quinn, 2000). This method requires a marking specification which serves only as a standard of comparison. The grades used are usually scales such as "excellent", "good", "average", "below average" or "unsatisfactory". The examiner can select examples of answers with reference to the standards for each of the points on the scale.

Having read through the answer quickly, the examiner grades it on the impression of the foregoing scales. The examiner repeats the grading process which can be quite effective for large number of assessment tasks.

3.5.2.5 "Reflective Practice Assignment" as Assessment Task

Fifthly, the action research project is basically an iterative "Reflective Practice Assignment" which can measure the candidate's capacity to analyze and evaluate experiences in the light of theories and research evidence. Conducting an action research project is a critical process that should be used to consider situated experiences and comparing these theories or established practice (Argyris & Schön, 1974, 1978 & 1996; Schön, 1983, 1988 & 1991). The thesis supervisor, gaining a greater understanding into this method of learning, would help the candidates achieve greater learning outcomes. The examiner, gaining a greater understanding into this method of learning, would better assess the candidates learning outcomes in a sensible way.

The focuses of "Reflective Practice Assignment" lie in reflection on action whereby thoughts and feelings about events during the action research are critically analyzed. Schön (1983, 1988 & 1991) provides a differentiation between "reflection in action" and "reflection on action". The former one represents a form of reflection whereby the candidate recalls his or thoughts from when the situation is occurring, for instance: "What should the candidate change to adjust the situation? Why is the candidate doing this?. In Chapter VI, Schön's theories will be used to analyze the qualitative data relating to student learning outcomes.

For the "Reflective Practice Assignment", marking specifications are relatively easy to set down. There is feedback potential from peers, self and the thesis supervisor. Marking for feedback can be slow whereas marking for grading is about the same as Essay (Brown et al., 1997). The thesis supervisor is in the best position to continually assess and provide feedback to the candidates who are conducting the action research as a "Reflective Practice Assignment". A subjective criteria-referenced assessment can usually reduce variability in marking (Sou, 2008).

3.5.2.6 "Oral" as Assessment Task

Sixthly, viva voce is an "Oral" defense of the action research and the master thesis. It tests the candidate's communication skills, understanding, and capacity to think quickly under pressure and their knowledge of procedures. There is feedback potential from the internal and external examiners. Marking for grading can be fast but some standardization of the interview procedure is needed.

3.5.2.7 "Presentation" as Assessment Task

Seventhly, a 10-minute narration during the viva is a "Presentation" which tests the candidate's preparation, understanding, knowledge, and capacity to structure information oral communication skills. There is potential for useful feedback from both the candidate and the examiners. Marking for grading based on simple criteria is fast and potentially reliable (Brown et al., 1997) particularly when there is based on the input of professional educators (HEFCE, 1993) such as internal and external examiners. Above all, it usually becomes the platform for the candidate's reflection upon his or her student learning experiences (Sou, 2008).

3.6 Previous Research in Mainland China

Previous studies (Sou, 2005a-d, 2006a-c & 2008; Sou & Zhou, 2005a-e & 2007; Zhou 2006; Zhou & Sou, 2005a-b) in mainland China reveals that there is a trend towards combined assessment methods to fit the purpose of assessment and measure the student learning outcomes at class, modular and program Levels. As discussed in Section 2.3 of Chapter II, the MBA candidates are required to participate in learning activities that include core subjects and electives. They are also required to conduct a research project at the end of the program and to compile a thesis under the supervision of a professor. Then, the candidates have to present and defend the research findings in a viva voce for the award of the MBA degree.

When different assessment methods are combined within a program, the task of program evaluation naturally becomes more subtle. Admittedly, it is somewhat difficult to get a simple tool for quality assurance in education. How can a program be evaluated systematically through focusing on its assessment practices? In mainland China, a world class quality management system designed by the International Organization of Standardization (ISO) was once used for program evaluation (Sou & Zhou, 2005a). As required by the ISO quality management system, appropriate assessment methods are always applied to evaluate the quality of the services or goods provided.

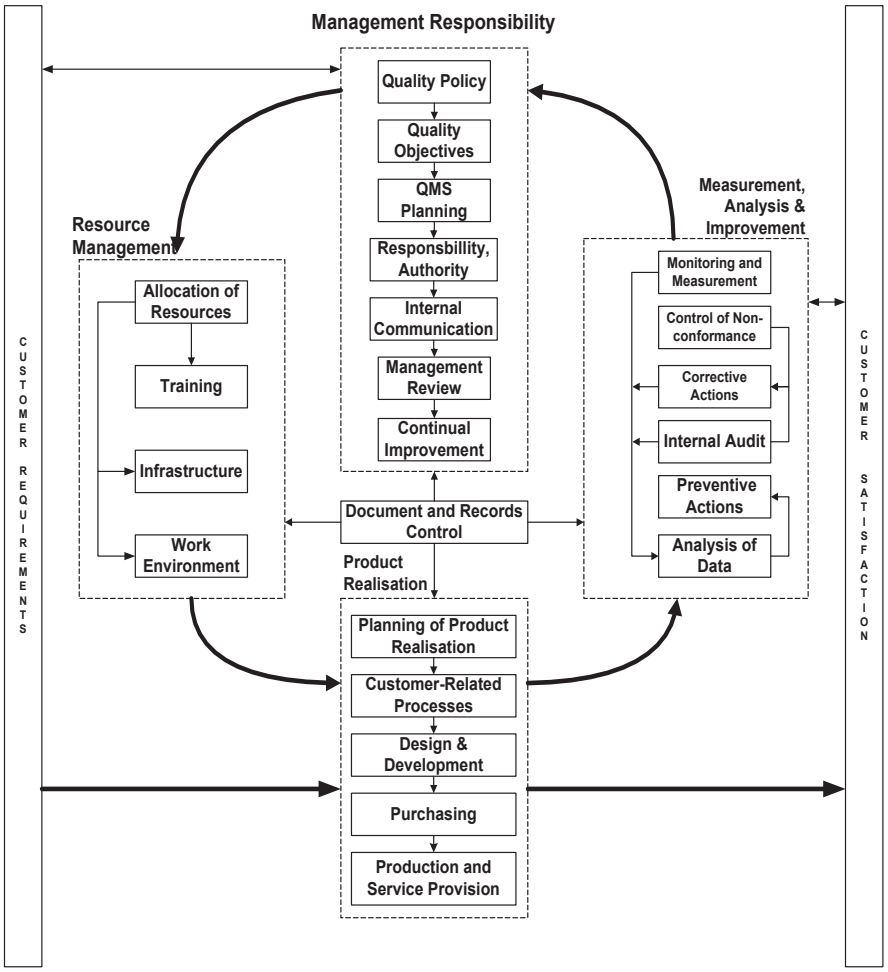
3.6.1 Application of an International Assessment Method in the Subject Program

The International Organization of Standardization (ISO) endeavors to design a generic management system model that is applicable to any organization regardless of size, nature of business and industry sector. ISO 9000 Quality Management System Standard sets out to establish, document and maintain an effective and efficient management system which

could demonstrate an organization's commitment to quality and its ability to satisfy customers' requirements.

Through time, the ISO 9000 standard has been revised to reflect the best practices and changing market expectation of progressive organizations and businesses. ISO 9001:2000 focuses on ultimate result and effectiveness. It was introduced with a lot of hype and excitement to the quality management circle worldwide. From 2000, the compatibility of this quality management tool, the "Quality Management Process Model" (Figure 5), was established with different industries including the education industry.

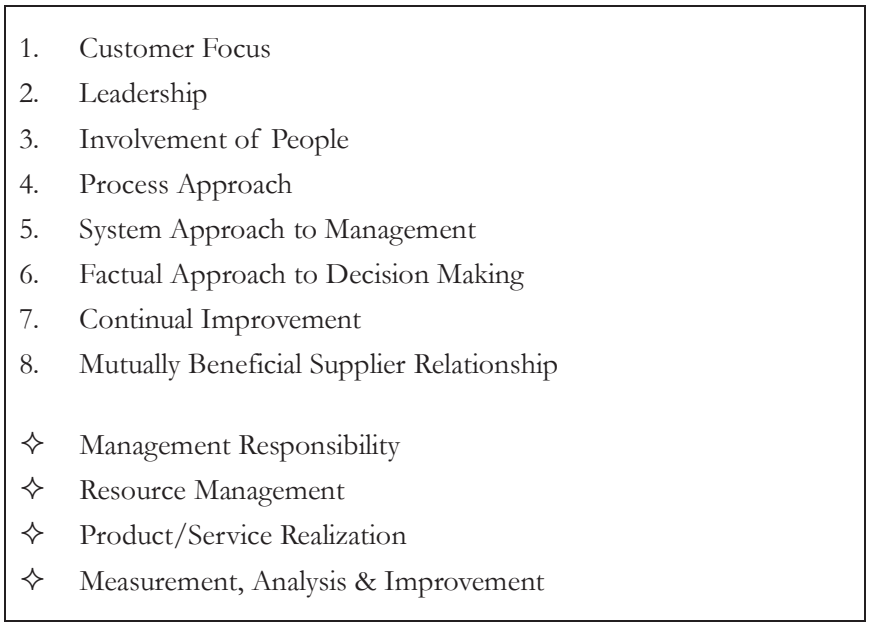
Figure 5: Quality Management Process Model under ISO 9001:2000



(Source: Sou & Zhou, 2005a; Zhou & Sou, 2005a)

ISO 9001:2000 is based on a quality management model of inter-related processes, propelled by the Plan-Do-Check-Act (PDAC) improvement cycle and driven by the 8 fundamental quality management principles with 4 key elements (Figure 6). Acquiring and retaining a competitive edge in business requires modern organizations to constantly seeking improvement opportunities for advancement.

Figure 6: Quality Management Process Model under ISO 9001:2000



(Source: Sou & Zhou, 2005a; Zhou & Sou, 2005a)

Economic changes and structural shifts in the many business environments have created formidable challenges for modern organizations, stimulating them to re-evaluate and improve their management practices. Educational institutions, like other modern organizations, have been inspired to adopt a more well-rounded and holistic approach in driving institutional improvement initiatives, placing greater focus on quality management and social expectation (Sou & Zhou, 2005a; Zhou & Sou, 2005a-b).

From 2004 to 2007, some quality management practitioners including me embarked on a series of quantitative and qualitative research (Sou, 2005a-d & 2006a-c; Sou & Zhou, 2005a-e & 2007; Zhou 2006; Zhou & Sou, 2005a-b). Based on the ISO 9001:2000 quality management system, they developed a feasible prototype of a program evaluation model for higher education in mainland China (International Organization for Standardization, 2000; Sou & Zhou, 2008). However, they also found that the model needed to be refined to reflect both the student learning experiences and the student learning outcomes (Sou & Zhou, 2008).

Previous research on a feasible program evaluation model involved the utilization of some conventional approaches such as Conduct of Student Surveys (Section 3.3.2.1) and Comparison of Actuarial Data (Section 3.3.2.2). Specimens of the tailor-made instruments (questionnaires used for the "Conduct of Student Surveys") in the previous research ((Sou, 2005c) are available for public reference in the "Sciencepaper Online" (a website administered by the Center for Science and Technology Development, Ministry of Education, People's Republic of China.

As introduced in section 3.1, there may be three evaluative components namely (a) Assessment Practices, (b) Student Learning Outcomes, and (c) Student Learning Experiences in a program evaluation program. Literature review surfaces the theories pertaining to these components in Section 3.2 to Section 3.5. Previous research in mainland China pertaining to the component (c) may add value to the literature review. In the following section, reference will be made to the findings of the previous studies on "Student Learning Experiences" in mainland China.

3.6.2 Literature Review in parallel with Findings of Previous Research

Through "Conduct of Student Survey" and "Comparison of Actuarial Data", affluent quantitative data pertaining to student learning experiences in previous research in mainland China could be the foundation of this research. Findings of the previous research can save the ground-work of studying some aspects of the student learning experiences. Those findings include the students' perceived degree of importance/satisfaction in the teaching, learning, research and appraisal processes of the subject program. This research will yield more qualitative and quantitative data about student learning outcomes thus filling the research gap between assessment of student learning and the evaluation of educational program. They would lead to a holistic view on the quality assurance of management education program.

Bickerstaffe (2000) believes that most MBA degree programs explicitly aim to improve effectiveness on the job and are usually designed in a similar way to a physical fitness regime. Teaching staff introduce a broad range of management functions in which they drill the learners, enabling them to respond swiftly and almost automatically. Case studies, action learning, and action research, which analyze real-life corporate problems, are typical components of an MBA. Such pedagogy is regarded as the tradition of a generalist but quality MBA degree program and its associated assessment practices are crucial to the quality assurance of the program (Bickerstaffe, 2000).

With reference to the quality assurance framework of the Higher Education Funding Council for Education (HEFCE), educational quality sometimes points to the links between the notions of "fitness for purpose" and "fitness of award". "Fitness for purpose" examines the links between particular students' experience and specific objectives of the educational program. In contrast, "fitness of award" concerns the links between

the judgment made against general criteria, usually national standards and the running of the educational program. In relative terms, "fitness of award" is independent of particular educational program and the students.

Obviously, there is an inherent conflict in educational quality assurance in terms of "fitness for purpose" and "fitness of award". The conflict is related to the issue of relativities and absolutes with reference to the interpretation of "good". As mentioned in Chapter I, "good" is construed on the students' reflection upon experiences which lead to differing interpretation of "good for what and for whom". According to the purposes, context and the shifting boundaries of experiences, "assessment of educational quality" in terms of "fitness for purpose" is conceptualized by this differing interpretation. However, appropriate assessment practices can mitigate the conflict of these 2 terms because "assessment" is at the heart of the student experiences (Brown and Knight, 1994).

3.6.3 Definition of "Assessment of Student Learning Experiences"

In higher education, "Assessment of Student Learning Experiences" defines what the students regard as important, how they spend their time and how they come to see themselves as students and then as graduates (Brown et al., 1997). Previous studies in mainland China have provided background information about the student learning experiences from the perspectives of the candidates of the subject program. This research will be built on the previous research and provide additional information about student learning experiences and student learning outcomes assessed by professional educators (HEFCE, 1993; Sou, 2008).

3.6.3.1 Meaning of "Assessment"

Brown et al. (1997) state that the term "assessment" is derived from ad sedere – to sit down beside. In the context of learning, "assessment" is primarily concerned with providing guidance and feedback to the learners which in turn promotes learning (Figure 7). Furthermore, student assessment contributes to quality education whilst judging achievement and maintaining standards (Carless, et al., 2006).

Figure 7: Main Purposes of Student Assessment

(Source: Sou, 2011)

Other scholars (Brown et al., 1997; Rowntree, 1987) hold that there are many similarities between the assessment of student learning experiences and educational quality assurance. They contend that the truth of an educational system can be discovered through the assessment of student learning experiences and the associated procedures. Therefore in the context of program evaluation, "Assessment of Student Learning Experiences" (fitness for purpose) can add value to "Assessment of Student Learning Outcomes" (fitness of award) and contribute collectively to educational quality assurance.

3.6.3.2 Nature of Assessment Practices

On the other hand, assessment practices of the subject program differentiate the educational quality in different ways depending on whether the assessment is "formative", "summative" or "diagnostic". "Formative" and "Summative" assessments are easily confused. The focus of the former is provision of feedback to the students during a training program. The students then would have opportunities to improve. Obviously, "formative assessment" does not contribute to the student marks earned in a training program. However, it does influence the student learning experiences as well as the student learning outcomes.

In educational practice, few training programs contain purely formative assessments. In higher education particularly management education, assessment tasks are usually both "formative" and "summative" (Bickerstaffe, 2000). "Assessment of Student Learning Outcomes" is intended as feedback to the students as well as contributing to the final profile of marks. Sometimes, the relationship between the type, the function and so forth the outcome of the assessment tasks are somewhat blurred.

Taking the subject program as an example, action research tasks are compulsorily assigned to the students during or at the end of the subject program. Such tasks enable the

examiners to assess the student learning outcomes upon the completion of the subject program. When a research task is assigned by the thesis supervisor to a student during the action research process and is used to provide feedback to the student on how well his or her Abilities, Skills or Knowledge (ASK) are developing, it could be described as a formative assessment.

When the thesis write-up task is assigned to the student at the end of the research process and is used to measure the extent of learning, it could be described as a "summative" or "terminal" assessment. Apart from "formative" or "summative" assessment, the task may include observation and measurement of the application of the students' ASK (Abilities, Skills and Knowledge) of becoming a "career manager" which would constitute a "diagnostic assessment". This assessment task for certifying "career managers" is performed by the internal and external examiners of the subject program.

"Diagnostic" assessment is a valuable component of the measurement of the students' ASK (Abilities, Skills and Knowledge). Unfortunately, this function is often lost in the administrative machinery of the "assessment system" (Brown, et al., 1997). Simply speaking, the composites or aggregates of marks in the "formative cum summative" assessment may hinder the identification of a student's particular strengths and weaknesses (Brown, et al., 1997). Thus, student learning outcomes could not be thoroughly assessed.

With the feature of objective-referenced assessment in the assessment practices of the subject program, Tylerian Approach is of reference to this research. It is because that Tylerian Approach does not rely heavily on the test scores of the candidates as the Testing Approach which has the characteristics of criteria-referenced assessment (Sou, 2008). In short, Tylerian Approach covers a much wider range of outcome data in the light of the learning objectives.

3.7 Theoretical Proposition

In the process of program evaluation, evaluation produces information of a particular type depending on the evaluation approach or model used. There are many approaches to quality assurance or models for program evaluation, each aiming at producing a certain kind of information (Augustine et al., 2001; Butcher & Hope, 2006; Chaffee & Sherr, 1992; Clark, 2003; Cornesky et al., 1991; Daniel, 2006; Forestier, 2004; Mills, 2006; Tsui, 2010). The Tylerian Models for Evaluation, with its orientation to objectives, testing and experimental design (Astin, 1991; Bell, 1980; Keogh & Naylor, 1980; Waluconis, 1993), is suitable for this doctoral research.

Tylerian Evaluation Rationale has been the mainstay of educational evaluation theory for more than 70 years whose influence on the profession of program evaluation is enormous (Astin, 1968, 1977, 1991, 1992 & 1993; Astin et al., 1987; Bradley et al., 1993; Bunda & Muffo, 1993; Cole et al., 1996; Ewell, 1987; Steele & Lutz, 1995; Tyler, 1942 & 1950). In the Tylerian Models for Evaluation, program evaluation is a process for determining the degree to which student learning outcomes match one or more defined instructional or behavioral objectives. More generically speaking, it is a process for determining the congruence, after instruction, between behaviors and objectives.

In the accountability system, program evaluation plays an essential role in improving programs and curriculum by exploring how far the learning activities go toward the desirable learning objectives and expected behaviors. The process for evaluation includes analyzing objectives to identify their behavioral content, identifying situations or circumstances in which a student could exhibit those behaviors, and selecting or developing assessment tasks or tools to test those behaviors (House, 1980, 1983a, 1983b & 2003). All these can be viewed as study on student learning outcomes.

However, Tylerian Models for Evaluation have some inherent disadvantages and/or limitations (Rowntree, 1987). For instance, they do not offer explicit judgments of worth or merit of an educational program. They do not provide way to evaluate the objectives of the educational program. They do not provide way to derive standards by which performance-objectives discrepancies could be judged. They do not provide way to assess pattern strengths and weaknesses. Further, they are convergent in effects: premature closure, stultified creativity, locking onto early objectives; and usually focusing attention on the pre-post design.

On the other hand, Tylerian Models for Evaluation also have advantages and/or potential that could outweigh the above disadvantages or limitations. For instance, Tylerian Models for Evaluation are systematic, elegant, and rational. They are simple, easy to understand and apply. They are built on prevailing scientific tradition while their guiding paradigm is that of the positivist scientific method. They differentiate the concepts of measurement and evaluation. Above all, Tylerian Models measure a program by means of criteria-referenced information instead of the norm-referenced data for program evaluation.

Until the advent of criteria-referenced assessment in mid-1960s, the tests used with Tylerian approach to evaluation were almost exclusively of the norm-referenced assessment. Subsequently, criteria-referenced assessment is more apt for installation into Tylerian Models for Evaluation. The reasons for this are: (a) criteria-referenced assessment requires a higher degree of specificity for objectives than do norm-referenced assessment, (b) criteria-referenced assessment is deliberately constructed to assess particular objectives, and (c) criteria-referenced assessment yield information directly interpretable in terms of the objectives (Kinchin, 2003; Michael et al., 1997; Sou & Zhou, 2005a-b; Stufflebeam, 2000; Zhou & Sou, 2005a).

Most assessment tasks require measurement instruments, such as questionnaires, surveys, interviews, observations, or a variety of tests (Brown et al., 1997). Standardized instrumental tests of ASK (Abilities, Skills, or Knowledge) are probably the most frequently used of the criteria-referenced assessment tools available (Brown & Knight, 1994). The increased use of criteria-referenced assessment is probably due to the fact that it is designed to yield data that references narrowly defined student learning outcomes. The high-focus capability of criteria-referenced assessment also makes it ideally suitable for program evaluation.

Though criteria-referenced assessment is a powerful tool that can be used in program evaluation, it cannot accommodate the needs for information about student learning outcomes as well as student learning experiences. Deciding which approaches or models do and which do not can be a tricky issue. I cannot simply conduct this research by using a haphazardly chosen approach or an inappropriately installed model and then expects that the research questions may be answered by the information produced.

Instead, Tylerian objective-based approach sustains the choice of the program evaluation model for quality assurance in this research. On top of that, Tylerian Models for Evaluation also meet meta-evaluation standards, that is, criteria of technical adequacy: reliability, validity, and objectivity (House, 1980, 1983a, 1983b & 2003).

Consequently, the notion of specifying objectives as the principal feature in program evaluation has been carried further in the late-1990s. Sou (2008) argues that behavioral objectives, regardless of clarity, are inadequate to define a domain unambiguously because they permit too much variation in their content, form and difficulty. Sou (2008) also assumes that it is necessary to focus on the diagnostic function of the performance standards in program evaluation. He posits that performance standards detailing stimulus

and response attributes for particular assessment tasks can constitute the most cost-effective way of improving the quality of public education (Sou, 2008; Sou & Zhou, 2007).

Diagnosing the student learning outcomes in an assessment task can facilitate objective-based evaluation and boost educational quality assurance. In other words, assessment tasks can be quality-driven tools in educational programs. Assessment practices comprising assessment tasks, when studied thoroughly and understood properly, are a "can-do" area to explore in program evaluation for quality assurance (Brown & Knight, 1994).

3.7.1 Summary of Literature Review

To this end, I have studied the theories related to the program evaluation and completed a historical and comprehensive review on program evaluation approaches (Section 3.2). I have made reference to the American practices in program evaluation and compared four conventional approaches of educational quality assessment (Section 3.3). I have also reviewed four foundation models of program evaluation (Section 3.4).

I have analyzed the application of various assessment methods for program evaluation (Section 3.5). Those assessment methods are mainly applied towards the assessment of the student learning outcomes. Being evaluative components of a program evaluation model (Figure 1), "Assessment of Student Learning Outcomes" is as important as "Assessment of Student Learning Experiences". "Assessment of Student Learning Experiences" has been studied with reference to the findings of previous studies in mainland China (Section 3.6).

These two evaluative components communally contribute to another evaluative component, namely "Assessment Practices" of a sound program evaluation model for educational quality assurance. Eventually, the theoretical proposition (section 3.7) on (a) Effective Assessment Practices, (b) Desirable Student Learning Outcomes, and (c) Favorable Student Learning Experiences will become the analytical framework in Chapter V, VI and VII.

IV. RESEARCH METHODOLOGY

4.1 Overview of Methodology

UTS Human Research Ethics Committee approved this research vide Ethics Approval # HREC 2005-062P. Full ethics clearance was given to the methods and procedures employed in this research therefore the ethical issues would not be elaborated in this chapter. Instead, this chapter focuses on the methodology embracing five research methods and the associated procedures. The first section of this chapter depicts the methodology and its justification while the second section introduces the supporting research methods. The third section details the research procedures and their paradigm while the last section sums up the contents of this chapter.

4.1.1 Use of Mixed Methods

In many theses or studies, several research methods may be used. Authors and researchers increasingly attempt to mix methods to some extent, because it provides greater perspectives on the phenomena being studied (Abrahamson, 1983; Ainscow et al., 1994; Brennan, 1997; Cohen et al., 2003; Creighton et al., 2003; Easterby-Smith, 1991; Jennings, 1993). On the other hand, Perry (1999) considers that there will usually be only one major method that suits the research problem and associated research gaps. The major method may be either "qualitative" or "quantitative".

Qualitative methods may be more compatible with interpretive study. However, qualitative data may be difficult to be expressible than quantitative data. Sometimes, it may be equivocal to describable the phenomena being studied. Therefore, using both "qualitative" and "quantitative" methods in this research would be more reliable and manageable. These two mixed methods can strike a balance between each other and

facilitate the research processes to a great extent:

> *By the mystery and expressiveness of numbers, we can express what is inexpressible, describe what is indescribable, predict what is reasonable to expect, or infer a logical conclusion to a series of events. Statistic is a language that can speak where other tongues are mute. Words cannot express the concepts that have been reserved for the eloquence and expressiveness of statistics alone. (Leedy, 1994, p 243)*

Dr. Laurence Chalip of Griffith University, Australia teaches that science does not advance through methodological proscription (Sou, 2004). It advances as a consequence of meaningful questions, empirical work to answer those questions, and a thoughtful analysis of the assumptive underpinnings of our work (Sou 2008). In other words, doing robust social science requires that a disputatious community of social scientists attack its questions analytically, applying any and all methods that may be useful to find answers (Sou, 2004 & 2008).

Therefore, this research utilizes a mix of "qualitative" and "quantitative" investigations. "Interpretive Study" is the major qualitative methodology for this research. Interpretive study can be viewed as the systematic processes of observing, detailing, describing, documenting, and analyzing the patterns of a culture or subculture (Jennings, 1993). Interpretive study can also be viewed as the science of "cultural description" or "naturalistic description" or "analytic description" (Leedy, 1994; Vaus, 1996). Some researchers just address it as a constructivist approach or simply fieldwork while some quote participant observation or case study in defining interpretive study (Sou, 2007; Strauss, 1987; Strauss & Corbin, 1990).

4.1.2 Sources of Data

In this research, I used interviews, observation, and documentary sources in examining ordinary, routine, mundane aspects of a group or organization's life (Jennings, 1993). The materials for interpretive study were myths, symbols, rituals and expressive behaviors, which reflected how people interacted and communicated with each other (Sou & Chow, 2007). Through interpretive study, themes were derived from data, systematically gathered, categorized, and analyzed (Glaser, 1992; Glaser & Strauss, 1967; Strauss, 1987; Strauss & Corbin, 1990 & 1998). There were two major data sources for interpretive studies:

1. Archival study of 331 master thesis assessment forms completed by the internal and external examiners of AIOU
2. Serial interviews with the informants (6 internal examiners, 9 external examiners, and 2 registrars of AIOU) on 23 occasions

Study of the assessment forms (specimen at Appendix A) examined how the MBA candidates applied the knowledge and skills gained in the subject program. In those forms, the examiners assessed how the candidates had completed their theses by applying the professional knowledge and management skills learnt. The examiners gave marks t five aspects of each thesis. Interviews of the examiners and registrars ("Outline of Questions" at Appendix B) informed me how the MBA candidates performed in the viva. Simultaneously, the candidates reported to the examiners how they completed their theses. These two different data sources generated rich qualitative data about Student Learning Experiences (SLE) and Student Learning Outcomes (SLO) for analysis.

This research did not only produce a "yes" or "no" answer to the three research questions framed in Chapter 1.2. It told "how" well the Student Learning Experiences of the MBA candidates matched the claims of the educational program provider, that is, AIOU. Furthermore, direct measurement of Student Learning Outcomes told "in what way" the sampled MBA candidates' performance matched the claim of AIOU in cultivating career managers through the subject program. As mentioned in Chapter 1.3, the readers could judge whether AIOU provided a quality MBA degree program with "fitness for purpose" to its MBA candidates. To recap, "fitness for purpose" (HEFCE, 1993) means the examination of the links between particular student learning experiences and specific objectives of the subject program.

As mentioned in the above paragraph, this research did not aim to answer leading research questions – a question simply yields a "yes" or "no" answer. Therefore, I did not begin this research with a preconceived theory in my mind because I was not going to elaborate and extend existing theories. Instead, I began with the exploration of three areas of study (Student Learning Experiences, Student Learning Outcomes and Assessment Practices) and allowed the themes to emerge from the data. Themes derived from data was more likely to resemble the "reality" than was theme derived by putting together a series of concepts based on theories or solely through speculation, that is, how one thought things ought to work.

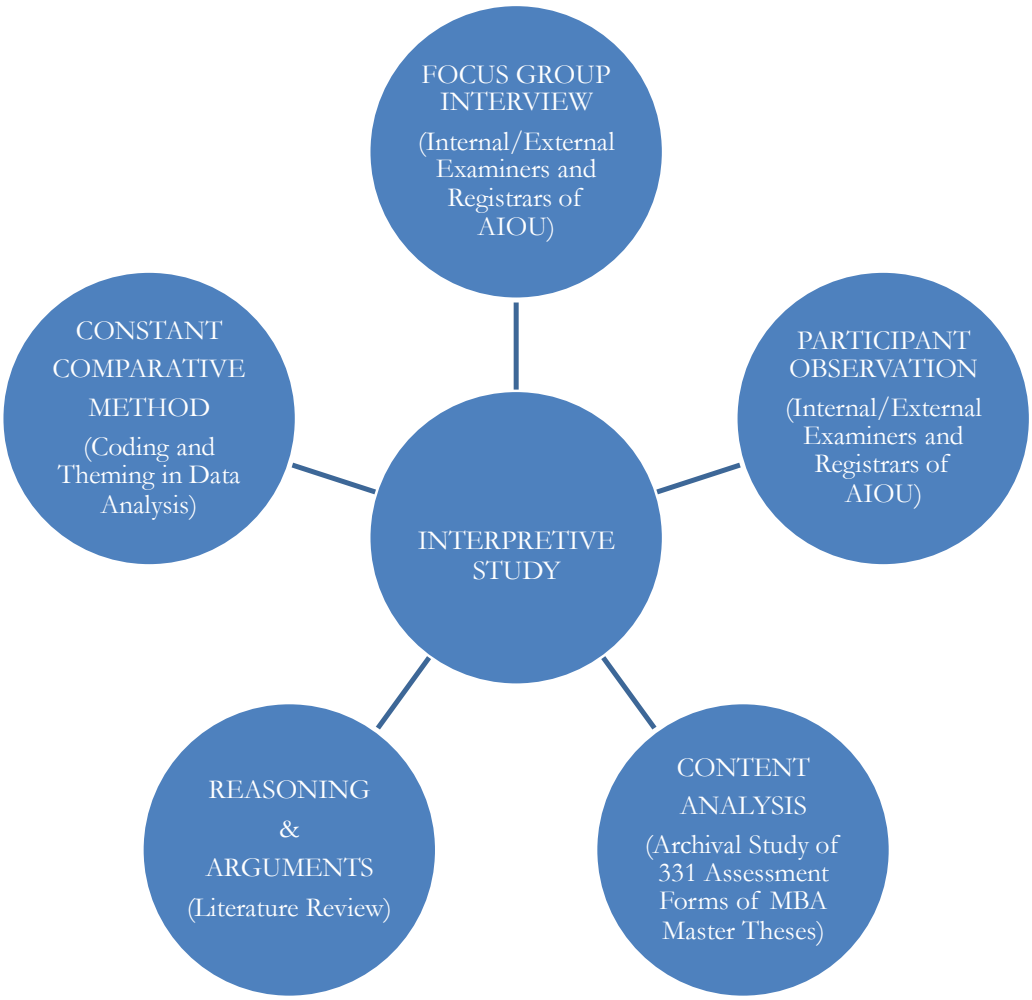
As for the focus group interviews, I began with unstructured and imprecise questions and expected imprecise answers which provided me with affluent data. Data were of crucial importance in any research and particularly in interpretive study. Working discreetly with the data could increase insight into the issue being researched. I believed that the data did not lie after careful analysis and would develop the theme. If the developed theme did not fit the data then there was something wrong with the theme. In such circumstances, I

needed to go back to the data to revise the theme (Sou, 2004 & 2008; Sou & Chow, 2007). During the data analysis processes, I therefore maintained an attitude of healthy skepticism: nothing was final until it was fully supported by data and the themes fully explained every aspect of that data.

4.2 Research Methods

In this section, the five research methods supporting interpretive study (illustration in Figure 8) will be briefed successively:

Figure 8: Methodology embracing Five Research Methods



1. "Focus Group Interviews" (Blumer, 1969; Bogdan & Biklen, 1992; Goetz & Lecompte, 1984) of 6 internal examiners, 9 external examiners, and 2 registrars of AIOU.
2. "Participant Observation" (Strauss & Corbin, 1990; Taylor & Bogdan, 1984) in a series of focus group interviews (Bodgan & Biklen, 1992; Glaser & Strauss, 1967, Strauss, 1987; Wax, 1971) with the internal or external examiners and the registrars of AIOU.
3. "Content Analysis" (Abrahamson, 1983; Holtsi, 1969; Vogt, 1993) of 331 assessment forms of MBA candidates completed by the internal and external examiners of AIOU.
4. "Reasoning and Arguments" (Nosich, 1982; Sou, 2001d) in literature review.
5. "Constant Comparative Method" (Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987) in data analysis.

4.2.1 Focus Group Interviews

Briefly, I employed interviews in conjunction with participant observation (Strauss, 1987; Strauss & Corbin, 1990) for data collection. Bogdan and Biklen (1992) state that an interview is a purposeful conversation used to produce rich, descriptive data about how participants interpret their world. Blumer (1969) add that interviewing is selected to reveal "the meanings that individuals have for things and people" (p. 2).

Focus groups comprised internal examiners, external examines and registrars of AIOU. Participation of the focus groups was voluntary. I interviewed the participants of the focus groups in accordance with the "Outline of Questions for Focus Group Interviews" (**Appendix B**). In Chapter V of this dissertation, 23 Focus Group Interviews participated by me were analyzed. Focus of study was the student learning experiences of

the subject program. In addition, findings of the focus group interviews became the foundation of the data analysis on the student learning outcomes.

"Focus Group Interviews" often brought up "cases" (Ouiment et al., 2001). Miles and Huberman (1994) define case as a social phenomenon of some sort occurring in a bounded context. The case may be a person, a group, an episode, a process, a community, a society, or any other unit of social life. It provides an opportunity for the intensive analysis of many specific details that are often overlooked with other methods (Theodorson & Theodorson, 1979). In this research, an interpretive case study is an intensive, holistic description and analysis of a bounded phenomenon such as an institution, a program, a person, a process, and a social unit. (Merriam, 1988).

In the data analysis within the following chapters, "cases" would be studied in detail. The general objective is to develop as full an understanding of that case as possible. I aimed to understand the case in depth, and in its natural setting, recognizing its complexity and its context (Stake, 1995). Above all, I had to differentiate the "cases" by its context and nature. A negative case refuted a construct while a discrepant case refined a construct (Goetz & LeCompte, 1984). During the interviews, the line between negative cases and discrepant cases often blurred. I did not mind what to name the "differing instances" of cases (Goetz & LeCompte, 1984). I just used both negative and discrepant cases to shape and refine the findings. I went through an active process of confirmation and readily changed my mind about findings when the data so dictated.

Glaser (1967) disagreed with recording or even taking notes during interviews in data collection session. I did agree that tape-recording would hamper the interview atmosphere and the informants might be conscious about the interviews and would not let out in-depth information. However, there might be a down side of not recording and

transcribing interviews as it opened the interviews to even more interpretation. Therefore, I took keyword field-notes during interviews and converted them to analytic memos immediately after the interviews. Eventually, practical certainty was achieved with a series of comparison and categorization.

4.2.2 Participant Observation

"Participant Observation" is the conventional name given to data collection that involves social interaction between a researcher and the informants in a naturalistic setting during which data are systematically collected (Strauss & Corbin, 1990; Taylor & Bogdan, 1984). Though meaning itself could be observed through the procedure of participant observation, it was still possible for me to observe social interactions that contributed to this research.

In this research, I struggled for a balanced and fair perspective. Sometimes, the participant observations confirmed as well as conflicted with my personal views, and/or with my reading and research experiences. I had to return to the field for more overt or covert observations. Such reflections through repetitive observations helped me to discover some possible factors that might be related to the discrepancies.

The interplay between interviews and participant observation was reciprocal. I, through interviews, provided clues as to what my focus should be in future occasions of participant observation. Furthermore, because it was impossible for me to observe all the social interactions that influenced this research; it was necessary for me to inquire through further interviews (See Bogdan & Biklen, 1992; Glaser & Strauss, 1967; Strauss, 1987; Wax, 1971).

During this research, I on and off returned to field when the foci of study shifted. My broad and general plans laid at the very beginning to establish credibility were redefined and augmented when circumstances required. Sometimes, I returned to the field in

respect of an anomaly in the data. Sometimes, I searched for a bit more information. To obtain more information and data for analysis, it was necessary for me to employ some supporting research methods such as Content Analysis, Reasoning and Arguments, and Constant Comparative Method.

4.2.3 Content Analysis

"Content Analysis" is any technique for making inferences by systematic and objective identifying special characteristics of messages (Holsti, 1969; Vogt, 1993). It was mainly used in the scrutiny of the archives and internal documents of AIOU. In reviewing the relevant literatures of parent discipline, it came up with theoretical samples. Theoretical samples provided concrete evidence of defendable propositions originated from the data analysis.

In analyzing data, I kept on examining artifacts of social communications between the internal or external examiners and the registrars of AIOU, for instance, field-notes of focus group interviews and analytic memos. Abrahamson (1983) suggests that "content analysis" can be fruitfully employed to examine virtually any type of communication. In other words, content analysis may focus on any materials or messages. However, it is true that only the manifest attributes of text may be coded. Thus, I did not solely rely on the inferences about latent meanings of materials or messages. Instead, I sought corroboration by independent evidence from other supporting research methods.

4.2.4 Reasoning and Arguments

"Reasoning and Arguments" (Nosich, 1982, Sou, 2000; Sou & Chow, 2007) is a research method used in a supporting role to formulate research problems, analyze qualitative data, and develop themes. In the literature review, "reasoning and arguments" was principally adopted for qualitative analysis. The central sense of reasoning and argument meant

understanding the issues, examining them carefully and thinking hard about them. During the research, reasoning skills and arguable thinking (Sou & Zhou, 2007):

- ✧ enabled me to construct arguments for beliefs I held;
- ✧ gave me chance to evaluate other people's arguments;
- ✧ helped me get an understanding of strategies in arguing and acknowledge of issues I was concerned; and
- ✧ helped me to question, examine and deal with my life and the things in it.

For the purpose of the interpretive study, I collected, coded, and analyzed the data obtained from the informants. Data "are inductively derived from the study of the phenomenon it represents and are verified through systematic data collection and analysis of the data" (Strauss & Corbin, 1990, p. 23). In this research, data collection and analysis stood in a reciprocal relationship with one another, that is, data collection leading to analysis, and analysis leading to further data collection.

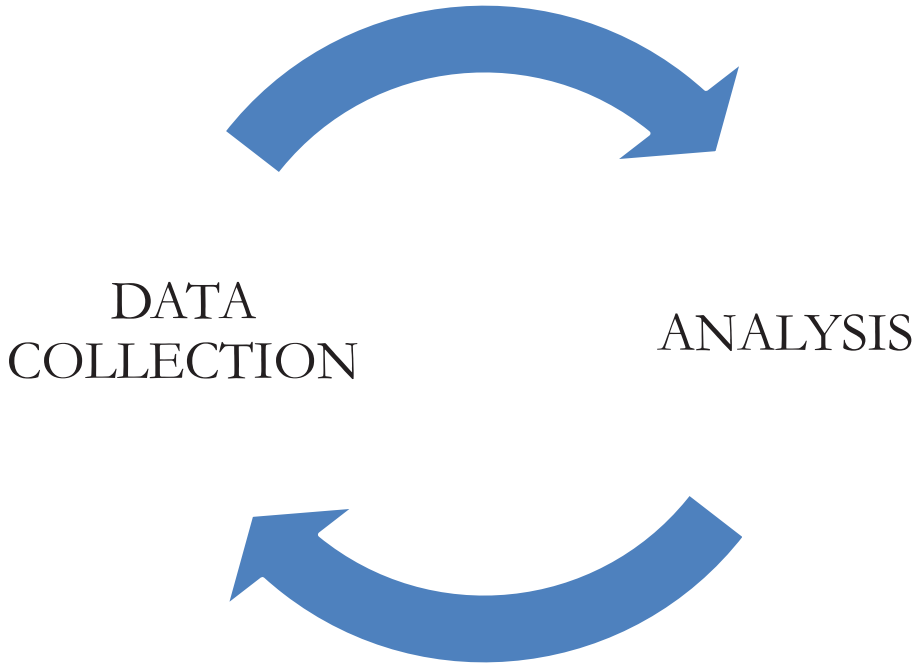
4.2.5 Constant Comparative Method

The reciprocal process of constantly switching back and forth between data collection and analysis (Figure 9) is referred to as the Constant Comparative Process (Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987). Constant Comparative Process is the essence of Constant Comparative Method. In describing Constant Comparative Process, Glaser and Strauss (1967) note that it has two functions: (1) as a procedure for analyzing data, and (2) as a means of developing theories. The Constant Comparative Method that produces "themes" instead of grounded theories is in fact an anomaly of grounded theory approach.

It was important for me to keep my mind and my eyes open in an interpretive study. There were lots of concepts from the two major data sources to be interpreted. Focus

group interviews were frequently the main source of the information from which I could develop the themes. Other than interviews, I could use participant observation, content analysis, reasoning and arguments, and constant comparative method to explore concepts, interpret the phenomenon, and develop themes that were implied by the qualitative and quantitative data.

Figure 9: Constant Comparative Process in Constant Comparative Method



(Source: Sou & Chow, 2007)

4.2.6 Interpretive Study

Interpretive study is a qualitative research methodology with "constant comparison" and "inductive" features which is different from qualitative quasi-deductive methods. The term "qualitative research" is confusing because it can mean different things to different people. Strauss and Corbin (1998) define qualitative research as "any type of research that produces findings not arrived at by statistical procedures or other means of quantification". It can refer to research about persons' lives, lived experiences, behaviors, emotions, and feelings as well as about organizational functioning, social movements, cultural phenomena, and interactions between nations.

Some researchers gather data by means of interviews and observations, techniques normally associated with qualitative methods. However, they code the data in a manner that allows them to be statistically analyzed. They are, in effect, quantifying qualitative data. Qualitative data may also include documents, films or videotapes, and even data that have been quantified for other purposes such as census data. In speaking about the quantitative investigation of this research, I was referring to the quantifying of qualitative data for the purpose of discovering concepts and relationships in raw data. Chapter V and VII of this dissertation were a typical examples of analysis of quantifying qualitative data.

The "data synergy" or "data triangulation" refers to works as "quantitative data can indicate directly observable relationships and corroborate the findings from qualitative data". Qualitative data can help understand the rationale of the theory and underlying relationships. Using multiple research methods combined with the judicious reference to the library materials, I could provide validity and reliability to the data and credibility to the themes which needed some theoretical concepts. (Sou, 2001d; Sou & Chow, 2007).

4.2.7 Application of Statistical Tools

To supplement qualitative data analysis, I used the following statistical tools to process some quantitative data from Chapter V to VII (for instance, variables in the markings of the thesis supervisor, internal and external examiners) with a view to learn their statistical significance and implications to this research:

- ✧ Pearson Analysis
- ✧ Analysis of Variance (ANOVA)
- ✧ Dunnett's tD Test
- ✧ Tukey HSD (Honest Significant Difference) Test
- ✧ S-N-K (Studenized Newman Keuls) Test

4.2.7.1 Pearson Analysis

Karl Pearson (1857–1936) asserts that there are several measures to describe the relationship between two variables, one of these descriptors being the covariance (Lohninger, 1999). Covariance is a measure of how much two variables change together whereas variance is a special case of the covariance when the two variables are identical. If two variables tend to vary together, then the covariance between the two variables will be positive. On the other hand, when one of them is above its expected value the other variable tends to be below its expected value, then the covariance between the two variables will be negative.

Assuming that the "Student Learning Experiences" was related to a circular process of the subject program to some extent, I could define a statistical measure for the relationship between them. In short, I tried to sum up the products of the "Student Learning Experiences" and various curricular processes for all data points. However inspecting this measure more closely, I discovered that this number was heavily dependent on the absolute

values of the coordinates. If I added a constant amount to all coordinates (which meant shifting the data in the x-y-plane without changing their mutual relationship), this measure would increase approximately by the square of the added value.

So what I needed was a value which was independent of any shifting along the axes (invariance of translation). In order to achieve this, I subtracted the mean of the x and the y data before calculating the average of the product terms. Eventually, the covariance was independent of any translations. It played an important role in multivariate statistical study of the correlation of "Student Learning Experiences" with various curricular processes.

4.2.7.2 Analysis of Variance (ANOVA)

ANOVA is a test of the statistical significance of the differences among the mean scores of three or more groups on one or more variables or factors (Phillips, 1992 & Vogt, 1993). In fact, it is an extension of the t-test, which is designed to handle two groups only. To be more specific, I used ANOVA to assess the statistical significance of the relationship between categorical independent variables and a continuous dependent variable. The procedure in ANOVA involves computing a ratio (F-Ratio) of the variance within the groups (error variance) to the variance between the groups (explained variance).

Virtually, the name 'Analysis of Variance' is somewhat misleading (Zhou & Sou, 2005b). It should be termed more appropriately as "Multi-Mean Comparisons" as ANOVA involve analyzing sums of squares, not variances. In the subject research, ANOVA was particularly compatible and advantageous since several factors emerged could be tested. For instance, there were several factors about the "Student Learning Experiences" on the teaching process[7], learning process[8], research process[9], and appraisal process[10] of the

[7] Perceived Degree of Importance/Satisfaction in Teaching Mode

subject program. Furthermore, those primary and secondary data for quantitative investigation could be ultimately verified against the data from the qualitative investigation. In Chapter V & VII of this dissertation, ANOVA was used for the analysis of quantitative data and led to verification of qualitative data throughout Chapter VI to VII.

4.2.7.3 Dunnett's tD Test

When I used the ANOVA and found a significant F, all that said was that the various means were not all equal. However, it did not say which means were different. I noted the problem with doing multiple comparisons was that the familywise error[11] of the experiment increased with each comparison I did. One way to control this was to try hard to limit the number of comparisons. Another way was to reduce the comparison level of alpha to compensate for the inflation caused by doing multiple tests. Dunnett's tD test allowed me to do post-hoc comparisons of the "Student Learning Outcomes" in a more precise manner. It was similar to the Tukey test but was used only if a set of comparisons were being made to one particular group (Phillips, 1992 & Vogt, 1993).

4.2.7.4 Tukey's Honest Significant Difference (HSD) Test

When I had several study groups that were compared to one control group, the Tukey's test served a much more general purpose. It calculated a new critical value that could be used to evaluate whether differences between any two pairs of means were significant. The critical value was a little different because it involved the mean difference that had to be exceeded to achieve significance. So I calculated one critical value and then the difference

---

[8] Perceived Degree of Importance/Satisfaction in Subjects and Perceived Degree of Importance/Satisfaction in Training Materials

[9] Perceived Degree of Importance/Satisfaction in Research Mode

[10] Perceived Degree of Importance/Satisfaction in Appraisal Mode

[11] Familywise Error (FWE) is also known as alpha inflation or cumulative Type I error. FWE represents the probability that any one of a set of comparisons or significance tests is a Type I error. As more tests are conducted, the likelihood that one or more are significant just due to chance (Type I error) increases. (Phillips, 1992 & Vogt, 1993)

between all possible pairs of means. Each difference was then compared to the Tukey critical value. If the difference was larger than the Tukey value, the comparison was significant.

When Tukey's Test was more conservative (family based), it controlled better than S-N-K Test for Type I error. It was generally more powerful than S-N-K if a family-based alpha was desired. This statistical tool was the most appropriate and powerful means to test all possible pairwise comparisons in studying the "assessment practices" of the subject program. The pairwise comparisons included the individual assessments made by the thesis supervisors, internal examiners, external examiners and the collective assessments made by the internal and external examiners.

4.2.7.5 Studenized Newman Keuls (S-N-K) Test

On the contrary, S-N-K Test was also powerful because it was contrast based. It was liberal in cases where I needed not worry about alpha. Therefore, it was a method of choice for accuracy and powerful for pair-wise comparisons. In analyzing the data about "assessment practices", it allowed me to start with largest difference and end with the statistical significance of the assessors (thesis supervisors, internal and external examiners) involved in the subject program. In Chapter VII of this dissertation, the foregoing statistical tools were used for the analysis of quantitative data.

4.3 Research Procedures

With a basic understanding of the research methods, the readers of this thesis can understand the choice of methodology with justified paradigm. In this section, the research procedures will be detailed. For comprehensibility, this section will be started with the selection of research samples and the introduction of terms regarding data

sources like Technical Literature[12] and Non-Technical Literature[13] (Berliner, 1996; Glaser & Strauss, 1967; Strauss and Corbin, 1990; Wittrock, 1986).

In AIOU, there were six internal examiners in the Faculty within its School of Graduate Studies whereas two registrars were engaged in the administration of its MBA degree program. On the other hand, there were nine external examiners selectively invited to join the internal examiners for the assessment of MBA theses and viva. When there were three to five viva voce in a month, the combinations of internal and external examiners depended on the availability of both the internal and external examiners thus varied occasionally.

The external examiners had their own commitments to their professions; therefore, some external examiners on list might attend one or two viva in a year. Eventually, I met six internal examiners and nine external examiners during the focus group interviews at the AIOU's campus in Macau Special Administrative Region. These internal or external Examiners were the available but most active ones in the pool of examiners. Moreover, the registrars of AIOU joined the focus group interviews whenever they were available. Given the small numbers of available samples, I selected all available examiners and registrars as part of the research samples.

In this research, I scrutinized 331 MBA thesis assessment forms provided by AIOU. The assessment forms were completed by the academic panels formed by the combination of the six internal and nine external examiners. The theses concerned were submitted by the

---

[12] "Technical Literature" was used: (a) to develop theoretical sensitivity to the area; (b) to develop secondary sources of data such as quotations from the focus groups under study; (c) to develop questions for use in finding additional data searches; (d) to develop a theoretical sample; and (e) as a means to partially validate theory as it is developed. (Strauss and Corbin, 1990)

[13] "Non-Technical Literature" included "letters, reports, newspapers and other materials that are considered primary data to be analyzed for theory development (Glaser et al., 1967, pp. 55-60)". Furthermore, it was possible to make use of library materials in triangulating field data with observations from other studies (Glaser & Strauss, 1967).

MBA candidates from 20 different higher institutions in mainland China which ran MBA degree programs with AIOU. The 331 MBA candidates were the most recent graduates of the MBA degree program from mainland China. They were selected from the aggregation of about 2,000 MBA graduates in a calendar year.

Consequently, focus group interviews (all internal and external examiners available) and archival study (scrutiny of assessment forms) had been conducted. Selected assessment forms were archived electronically for subsequent data analysis. Field-notes of the interviews and archival study were written during the research processes. Analysis of primary qualitative data and secondary quantitative or qualitative data led to holistic findings of this qualitative research.

Any research has three major processes: collecting data, interpreting and analyzing the data, and producing research reports. With a basic understanding of the research samples and data sources, readers would know how the data are collected. Readers will then be led to have a look at the data analysis processes. In line with the conceptual framework of literature review (Chapter III), I began data analysis with a systematic reorganization of the foregoing materials that included technical and non-technical literature.

4.3.1 Field-Notes and Analytic Memos

For the analytic framework, I continually kept some systematic substantive notes to record the qualitative data collected during the focus group interviews. Predominantly, his field-notes in chronological order and loose minute form were descriptive and aimed to provide a detailed portrait of the various focus group interviews in which I was involved. The field-notes included physical descriptions of the situations and the informants, details of conversations, and account of events (Bogdan & Biklen, 1992; Taylor & Bogdan, 1984). Field-notes led to preliminary analysis in the light of observation comments.

When I was developing a set of ideas from the data, I habitually made some observation comments on the qualitative data. Then, I wrote analytic memos for logical thinking (Geer, 1864, Sou, 2000), namely, (1) testing crude yes-or-no propositions, (2) working with negatively expressed hypotheses, and (3) working eventually with rough themes (concepts to be confirmed). Analytic memos enabled me to think logically and became the core of data analysis. Sometimes, they reflected the development of some concepts to be confirmed consecutively. Analytic memos also contained the actual product of the three types of coding: Open, Axial and Selective (Strauss & Corbin, 1998, p. 217) in Chapter V.

4.3.2 Open, Axial and Selective Coding

Open coding was the analytic process through which concepts were identified and their properties and dimensions were discovered in data (Strauss & Corbin, 1998, p. 101). Axial coding was the process of relating categories to their sub-categories, termed "Axial" because coding occurred around the axis of a category, linking categories at the level of properties and dimensions (Strauss & Corbin, 1998, p. 123) and Selective coding was the process of integration and refining the theory (Strauss & Corbin, 1998, p. 143).

The active process of coding and confirming the concepts could be regarded as case study that has been defined in the preceding section. Whenever a number of cases allowed, I just picked up the examples randomly from amongst the relevant categories. Then, I documented every assertion or narration made in the thematic analysis with no fewer than three examples. Whenever inconsistencies in themes emerged, they would be crossed checked by means of more observation or interviews as far as possible. Findings of case study were usually reported as narratives when the writing of a narrative was the telling of a tale. In the course of data analysis, I kept on his thematic writing. My writings were virtually telling tales.

4.3.3 Realist, Confessional and Impressionist Tales

When I was engaged in the thematic writing, I discovered that there were "Realist Tales, Confessional Tales and Impressionist Tales" (Margolis et al., 2003). Realist tales provided a rather direct, matter-of-fact portrait of a studied culture, unclouded by much concern for how I produced such a portrait. Confessional tales focused far more on me than on the cultural studies. Impressionist tales were personalized accounts of fleeting moments of fieldwork cast in dramatic form.

I attempted to vanish behind a steady descriptive narrative whilst telling "Realist Tales". In focus group interviews, I adopted the device of allowing various informants (internal examiners, external examiners and registrars) to speak themselves. However, I noted that there were confessional elements woven into my memo-writing of "Realist Tales". This was attributed to my reflections in telling tales. Thus, I eventually tended to tell "Impressionist Tales". It was because "Impressionist Tales" carried elements of both realist and confessional writing.

I elected to tell "Impressionist Tales" rather than merely "Realist or Confessional Tales" as I intended to highlight the dramatic characteristics of qualitative investigation. In other words, I was aware that I could not be independent of the data and so forth the thematic analysis and establishment. I would bring to the readers my reasoning as well as arguments and my reservations about the findings. Airing these was extremely valuable to understanding resultant themes as well as to fashioning future research projects (Ely, 1991).

As mentioned in the preceding paragraphs, the analytic framework was stemmed from an analytic induction, appropriate coding and followed by empirical categorization of data (Ely, 1991). Whilst engaged in pre-categorization, I was able to generalize from a small number of cases. When being faithfully to the qualitative data, analytic induction in memo form

was used to counterbalance the problem of causal interference. The main steps in preparing analytic memos were:

- ✧ defining the phenomenon to be analyzed.
- ✧ formulating a hypothetical explanation of the phenomenon.
- ✧ studying one case to see whether the hypothesis related to the particular case.
- ✧ reformulating the hypothesis if it did not fit the case or redefining the phenomenon in order to exclude the particular negative case.
- ✧ indexing the analytic memos for data categorization.

The redefinition of phenomenon and reformulation of hypothesis was continued until a universal relationship had been established in terms of the analytical framework. Case study related to the analytic framework was also under way in the course of data categorization. Gradually, I sorted out the data by three core categories, namely (1) student learning experiences, (2) student learning outcomes, and (3) assessment practices. These three categories of data were respectively analyzed in Chapter V, VII and VIII.

Afterwards, I read through all my field-notes and analytic memos. I discovered that there were regularities in the informants' conversations with me that provided certain themes. Sometimes, it seemed vital to check an observation with the informants whom had been observed. In such circumstances, I further interviewed the informants for cross-checks. Focus group interviews cum participant observations with a nagging thought ensured the fairness in interpreting the data by category.

Having everything together in loose minutes, I began by rereading the entire log, commenting on both my comments and the observations therein. I attempted to construct a general overview of the entire picture. My field-notes, observation comments and analytic memos were empirically sorted into categories. Constant Comparative

Process was a helpful tool for the data analysis in this regard. In practice, the line by line "coding" might be exactly the same as a general inductive approach to data analysis. Under Constant Comparative Method, I searched line by line through my written up field-notes for units of data, called "sub-category" that served as the basis for defining "categories" (Lincoln & Guba, 1985).

4.3.4 Data Categorization

Through the constant comparative process, I built relationships between the category and sub-categories to form a core category (Glaser, 1978; Glaser & Strauss, 1967; Lincoln & Guba, 1985; Strauss, 1987). A core category might evolve out of the list of existing categories. A core category was the central category around which all categories and sub-categories were integrated. A core category had analytic power. What gave it that power was its ability to pull the other categories together to form an explanatory whole. On the other hand, a core category accounted for considerable variation within categories.

Eventually, I identified 27 sub-categories in my field-notes over 23 serial interviews. By looking at each sub-category's similarities and differences, I grouped the 27 sub-categories into category sets that contained six categories and three core categories. As a result, I regarded "Student Learning Experiences", "Student Learning Outcomes" and "Assessment Practices" as three core categories of this research whereas "Program Evaluation" as the common thread connecting them (Figure 13 and Figure 14 in Chapter V).

In empirical categorization, I followed the five guidelines by which to judge the efficacy of categories in qualitative data analysis (Holsti, 1969; Sou & Chow, 2007):

1. The categories should reflect the purpose of the research and were congruent with research problems.
2. The categories should be exhaustive, that is, "all relevant items in the sample of documents under study must be capable of being placed into a category".
3. The categories should be mutually exclusive – no single unit of material should be placed in more than one category.
4. The categories should be independent in that "assignments of any datum into a category not affect the classification of other data".
5. All categories should derive from a single classification principle.

It was understandable that category-making process depended solely upon the knowledge, insight and imagination of the person who collected the data. I all the way endeavored to keep the analytic framework on track with the conceptual and analytical framework of this research. Besides, I always referred to the six criteria (Glaser et al, 1967; Strauss, 1998) that could be applied to a category to determine whether it qualified:

1. It must be central, that is, all other major categories could be related to it.
2. It must appear frequently in the data. This meant that within all or almost all cases, there were indicators pointing to that concept.
3. The explanation that evolved by relating the categories was logical and consistent. There was no forcing of data.
4. The name or phrase used to describe the core category should be sufficient abstract that it could be used to do research in other substantive areas, leading to the development of a more general theory.
5. As the concept was refined analytically through integration with other concepts, the theory grew in depth and explanatory power.

6. The concept was able to explain variation as well as the main point made by the idea, that is, when conditions varied, the explanation still held, although the way in which a phenomenon was expressed might look somewhat different. One also should be able to explain contradictory or alternative cases in terms of that central idea.

I satisfied that the above criteria were met in data analysis. There were several techniques that I used to facilitate identification of the core category and the integration of concepts. Among these were writing the storyline, using diagrams and matrixes, or reviewing and sorting memos by computer program for content analysis. With the use of these techniques, I could trim off excessive and filled in poorly-developed categories. Finally, the themes emerged were validated by comparing them to raw data or by presenting them to the informants for their reactions. Themes should be recognizable to the informants; and although it might not fit every aspect of their cases, the larger concepts should have applied (Strauss & Corbin, 1998, p. 161).

4.3.5 Theming

A theme could be a statement of meaning that ran through all or most of the pertinent data or phrases or single words in the minority that carried heavy emotional or factual impact (Ely, 1991; Sou & Zhou, 2007). Themes could be developed on the inferred statements, phrases or words that highlight explicit or implied attitudes toward life, behavior, or understandings of the target population or their culture. In developing the themes, I usually carried out the following tasks (Margot, 1991; Margolis et al., 2003; Sou, 2000):

1. studying and re-studying the raw data to develop detailed, intimate knowledge;
2. noting preliminary impressions;
3. listing tentative categories
4. refining categories by collating the tentative categories with the reviewed literature and returning to the entire raw database;
5. grouping data under the still-tentative categories and revising categories when circumstances warranted;
6. selecting the verbatim narration to link the raw data to the categories;
7. studying results and revising the themes if desirable;
8. writing phrases or words and then theme statements for the target population to speak from their points of view by linking data in and across categories;
9. integrating findings about the target population; and
10. comparing findings for commonalities or patterns, differences and unique happenings.

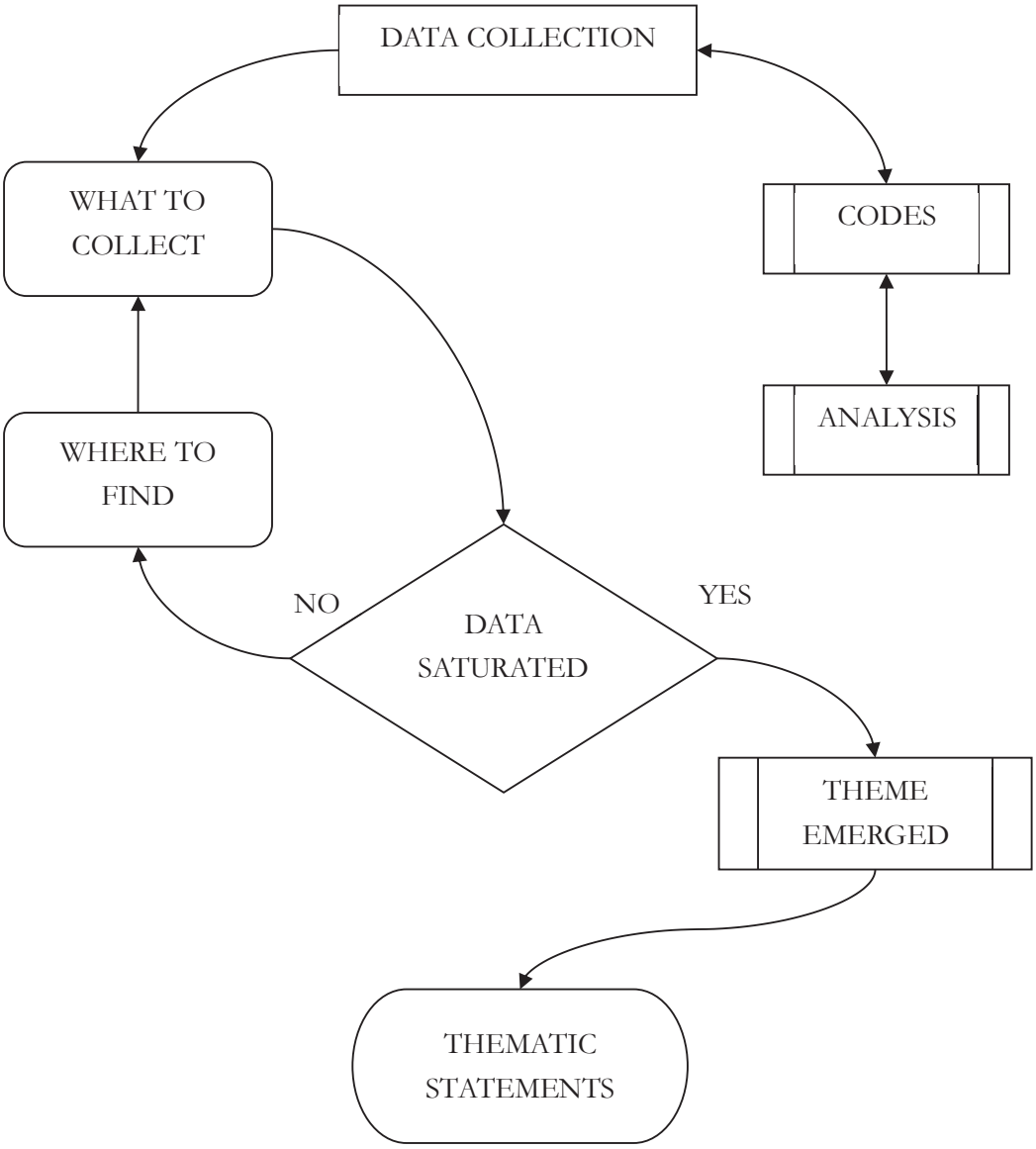
I established some themes with first building a set of categories. Whilst I progressed the data from categories to themes, the resultant themes were more grounded in the data. I was continually concerned about the trustworthiness of those analytic categories and themes. I attempted to make this research credible, produce results that could be trusted and establish findings to which they were worth paying attention. Hence, revision and refinement of still-tentative categories took place frequently. Moreover, the presentation of each of the themes should be bolstered by the data from participant observations as well as information from the Technical and Non-Technical Literature.

During data analysis (Figure 10), I jointly collected codes and analyzed data and decided what data to collect next and where to find them, in order to develop a theme as it emerged (Glaser & Strauss, 1967). When data collection was saturated, no more information could be added to the categories; it was deemed the time to sort and write the themes. In short, themes were generated through constant comparison and were grounded in data. With the aid of constant comparative method, "codes" and "categories" were weaved together conceptually to tell a story about human behavior, rather remaining suspended as a description of an event or feeling.

In the subject research, a theme was established on conditions that (1) it had appeared many times and/or for the majority of subjects under study or (2) it had appeared once or a very few times but carried important analytical impact. Themes arose when I found that there were direct and corroborative evidences collected in focus group interviews as well as participant observation. Amongst them, there were field-notes which were informed by theoretical schemes and analytic memos which developed conceptual scheme on the basis of the data.

Furthermore, there were substantive theoretical accounts such as technical literature that were concerned with substantive theory as well as data where I had generated the thematic statements. These themes were social phenomena associated with qualitative data and findings. The seven themes developed inductively from a corpus of qualitative and quantitative data were reported in Chapter VIII.

Figure 10: Data Collection Process for developing Thematic Statements



(Source: Sou & Chow, 2007)

4.4 Summary

This chapter describes the use of interpretive study in a qualitative educational research. Interpretive study is underpinned by symbolic interactionism (Blumer, 1969; Sou & Chow, 2007), a sociological perspective and jargon. Readers will see the academic rigor, tedious processes and multiple, but flexible methods in an interpretive study. Interviews, participant observation and Constant Comparative Process are time-consuming research activities in a self-disciplined research journey.

Regardless of the difficulties in an interpretive study, data analysis had been conducted step by step in this research. Firstly, data were organized chronologically or categorically and presented in a narrative that was largely descriptive, for instance, field-notes. Secondly, I moved from concrete description of observable data to a somewhat more abstract level. Concepts were used to describe phenomena in analytic memos.

Thirdly and continually, data were coded and classified into categories. The categories described the data and simultaneously interpreted the data to a certain extent. Then, I could make inferences, develop and establish some themes. Ultimately, I was moving up "from the empirical trenches to a more conceptual overview of the landscape" (Sou & Zhou, 2007). I was no longer dealing with observable but also with unobservable, and was connecting the two with successive layers of inferential glue (Holsti, 1969; Sou & Chow, 2007). I thoroughly analyzed the data thus making inferences and developing themes.

Theme-writing was the last step about data analysis – a step towards developing thematic statements that explained some aspect of the "assessment practices" and allowed the readers of this dissertation to generalize future activities. This interpretive study had been undertaken from 2004 to 2009. At the beginning of 2009, I had experienced a weary sense of relief, knowing that my data were sufficient because they were repeating themselves. In short, there was not one new category; not one new theme; but a few more examples to theorize the themes. Thus, I believed that the "mixed methods" had been applied successfully in this research.

At that stage, I satisfied that "theme-writing" was the cognitive process of discovering or manipulating abstract categories and the relationships among those categories. A Theme emerged when there was a perfect fit between it and the categorized data. Then, thematic statements were generated as findings of data analysis. In the latter chapters of this dissertation, I would bring public spotlight on my interpretive processes in establishing findings. Readers would have the information with which to judge for themselves whether the findings are reasonable. Furthermore, readers would discover that the findings in the qualitative and quantitative investigation should support each other to a certain extent.

As a prelude of the following chapters, data collected were used for analysis in Chapter V to VII. To write up the findings of this research, I recounted integrative field-notes and reviewed the memos until I was clear about the main analytical story. I began with reliance on the cache of field-notes and analytic memos, which provided the basis for writing. The writing required (1) a clear analytic story, (2) a sense of what parts of the story I wished to convey, (3) a detailed outline, and (4) a stack of pertinent memos to fill in the details of the outline. In Chapter V – VII, readers would be guided to read a full story.

Then, I constructed Chapter VIII and IX on two procedures (Strauss & Corbin, 1998, p. 251), "developing a clear analytic story by sorting through the diagrams and memos; and then working out a main outline that will fully incorporate all important components of that story". Glaser and Strauss (1967, p. 113) state "when I am convinced that my analytic framework form a systematic substantive theory, that is it is reasonably accurate statement of matters studied, and that it is couched in a form that others going into the same field would use … that I can publish my results with confidence".

V. DATA ANALYSIS – STUDENT LEARNING EXPERIENCES

5.1 Synopsis

This chapter reports on the processes and findings of data analysis regarding Student Learning Experiences (SLE). Data analysis and findings pertaining to Student Learning Outcomes (SLO) and Assessment Practices will be reported in Chapter VI and Chapter VII sequentially. Then data analysis in these three aspects and their findings will be consolidated in Chapter VIII. In the following paragraphs of this synopsis, there will be a general picture of the processes for the data analysis.

5.1.1 Conduct of Focus Group Interviews

From 30 January 2007 to 24 April 2008 (sampling period), 23 focus group interviews (Figure 11) were held in Macau Special Administrative Region. Voluntary participation of the focus group interviews eventually exhausted all the available six internal examiners, nine external examiners and two registrars of AIOU. One of the registrars was a doctoral candidate while the other participants of the focus group interviews had either Doctorates of Philosophy (PhD) or professional doctorates[14]. The external examiners included academics from the other universities or management practitioners in the commerce, industry or public sectors.

The registrars organized the MBA candidates to appear before the academic panels for their viva voce. The internal and external examiners formed various academic panels which assessed all the MBA candidates of AIOU in the sampling period. Assessment forms (Appendix A) of 331 MBA candidates from 20 higher institutions of mainland China were selected for further study. The higher education institutions involved were

[14] Doctor of Education, Doctor of Letters, Doctor of Literature, Doctor of Engineering, Doctor of Management or Doctor of Business Administration.

traditional universities, private colleges and national training establishments for the comrades (mid-career civil servants who were elected members of the Chinese Communist Party) in mainland China.

Figure 11: 331 MBA Candidates assessed and 23 Focus Group Interviews from 30.01.07 to 24.04.08

S/N	Dates	MBA Candidates assessed	Higher Education Institutions[15] of PRC	Participants of Focus Group Interviews		
				Internal Examiners	External Examiners	Registrars
1	30.01.07	13	1	2	5	2
2	31.01.07	6	1	2	5	2
3	14.03.07	16	2	2	5	2
4	18.04.07	8 + 11	1 + 3	2	4	2
5	19.04.07	7 + 9	4 + 5	5	5	2
6	22.05.07	8 + 5	6 + 7	3	3	2
7	21.06.07	15	8	3	5	1
8	22.06.07	7 + 8	5 + 9	3	5	1
9	18.07.07	17	10	4	6	2
10	19.07.07	16	11	4	6	2
11	02.08.07	7 + 5	12 + 7	3	5	2
12	30.10.07	13 + 2	13 + 14	2	4	2
13	31.10.07	8 + 10	7 + 15	2	3	1
14	20.11.07	8	16	2	5	2
15	05.12.07	14	10	4	3	2
16	11.12.07	15 + 1	17 + 18	3	4	1
17	13.12.07	8 + 7	19 + 20	4	4	2
18	15.01.08	15	6	2	4	2
19	16.01.08	14 + 1	17 + 13	3	4	2
20	09.04.08	15	8	1	4	2
21	10.04.08	8 + 6	13 + 8	3	4	2
22	23.04.08	14	6	2	3	2
23	24.04.08	8 + 6	5 + 15	3	2	1
	23 days	**331 MBA candidates**	**20 higher institutions**	**64 times**	**98 times**	**41 times**

(Source: Assessment forms and field-notes of the subject research)

[15] 1. Ex-Tianjin Zhong Xin International Institute of Further Studies; 2. Jiang Han University; 3. Changsha Han Shuo Academy of Management; 4. Guangdong Academy of Technology for Comrades; 5. Shenzhen Hua Lian College of Commerce and Industry; 6. Guangxi Academy of Economics and Management for Comrades; 7. Guangdong Academy of Economics and Management for Comrades; 8. Liaoling Academy of Economics and Management for Comrades; 9. Guangdong Academy of Technology for Comrades; 10. Northwest University; 11. Beijing Academy of Finance and Management for Comrades; 12. Tianjin Academy of Finance and Management for Comrades; 13. Chengzhou Center of Adult Education; 14. Shandong Tai Shan College of Management; 15. China Advanced Studies and Research University; 16. National Finance Commission Training Center for Comrades; 17. Tianjin Faculty of Business; 18. Anhui University; 19. National Economics and Trade Commission Shandong Office of Occupational Education; 20. Tianjin Academy of Finance and Trade for Comrades.

5.1.2 Study of Primary and Secondary Data

For the interpretive study, the data from Technical Literature and Non-Technical Literature (Berliner, 1996; Glaser & Strauss, 1967; Strauss and Corbin, 1990; Wittrock, 1986) went through open, axial and selective coding, and Constant Comparative Processes (CCP). In fact, literature review, question or hypothesis generation, data collection and analysis occurred simultaneously. In short, the "Technical and Non-Technical Literature" involved in the data analysis are tabulated as per Figure 12:

Figure 12: "Technical Literature" and "Non-Technical Literature" used in Data Analysis

Technical Literature (Secondary Data)	✧ Archives of AIOU ✧ Reports on previous research in China ✧ Publications of referential value to the subject research
Non-Technical Literature (Primary Data)	✧ 331 Assessment Forms of MBA Candidates ✧ 331 MBA Theses of the MBA Candidates from 20 Higher Education Institutions. ✧ Field-Notes of Focus Group Interviews ✧ Analytic Memos with Observation Comments

(Source: The subject research)

5.1.3 Categorization of Primary and Secondary Data

Eventually, there were three core-categories, six categories and 27 sub-categories generated in the course of data analysis (Figure 13):

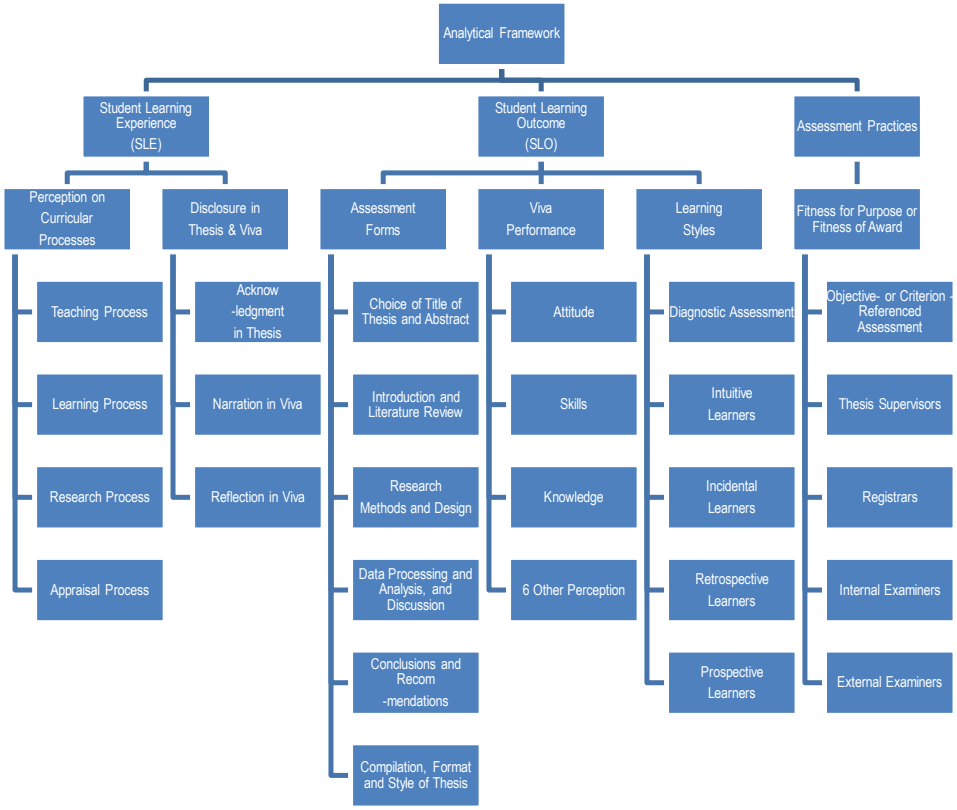
Figure 13: Core Category, Category and Sub-Category developed by Coding

Core Category	**Student Learning Experiences**		**Student Learning Outcomes**			**Assessment Practices**	**Total**
Category	2		3			1	**6**
Sub-Category	4	3	6	4	5	5	**27**

(Source: Analytic memos of the subject research)

Figure 14 is a diagram illustrating the analytical framework of this research. In the latter sections of this chapter and the succeeding chapters, the above core-categories, categories and sub-categories will be analyzed independently or communally. Readers will be guided through the journey of data analysis, data saturation and thematic emergence.

Figure 14: Diagram illustrating the Analytical Framework



5.2 Student Learning Experiences

This section reports the findings of the study of "Technical Literature" about student learning experiences in the Mainland. The findings can facilitate the further study of the "Non-Technical Literature" about the student learning experiences, student learning outcomes and the associated assessment practices. First, there will be a review of the AIOU's archive about the student learning experiences.

5.2.1 Quality Management System of AIOU

In September 2000, the administrators of AIOU sought ISO[16] certification for quality assurance. After a year's concerted efforts of the administrators, faculty and support staff, AIOU was awarded the ISO 9001:1994 certificate in December 2001. In January 2003, its ISO 9001 Certification was upgraded to the 2000 Edition (AIOU, 2006a-b). Emphasizing a "customer focused" and result oriented "process approach" management model and a "continual improvement" philosophy, the 2000 edition of this quality management tool realizes the compatibility of its quality management process model with different industries.

In 2000, this quality management tool viz. ISO 9001:2000 also extended its coverage to different industries. In the following years, there were quality assurance guidelines introduced to the education industry. This provides evidence that quality requirements (assessments) subsist in every organization including educational institutions. (Sou and Zhou, 2007; Zhou & Sou, 2005a-b)

The ISO 9001:2000 certification was awarded to AIOU for its "Provision of Curriculum Design and Tertiary Education". It evidenced that the administrators, faculty, and support staff of AIOU endeavored to meet the stringent requirements of certification under ISO within this scope of certified activities. However, the Certification had exclusion under clause 7.6 of ISO 9001:2000 QMS (Quality Management System) requirements (British Standards Institute, 2000, p.11):

[16] ISO stands for the International Organization of Standardization that endeavors to design a generic management system model applicable to any organization regardless of size, nature of business and industry sector. ISO 9000 Quality Management System (QMS) Standard sets out to establish, document and maintain an effective and efficient management system and ISO Certification demonstrates an organization's commitment to quality and its ability to satisfy customers' requirements. Through time, the ISO 9000 Standard has been revised to reflect the best practices and changing market expectation of progressive organizations and businesses.

7.6 Control of monitoring and measuring devices

The organization shall determine the monitoring and measurement to be undertaken and the monitoring and measuring devices needed to provide evidence of conformity of product to determined requirements (see 7.2.1). …

Unlike the other service organizations in modern business, AIOU is offering educational services or products as well as serving "customers" (MBA students) in the management education market. Such deviation from the ISO's origin for manufacturing industry or general service industry leads to the exclusion of clause 7.6. This reflects that monitoring and measuring educational services or product might not be viable to a certain extent as the student learning experiences and student learning outcomes are hardly quantifiable.

Despite the exclusion of this particular clause, clause 8 is still applicable to the educational institutions seeking ISO 9001:2000 certification. Clause 8 stipulates that the organization should plan and implement the monitoring, measurement, analysis and improvement processes needed (BSI, 2000, p. 11):

- ✧ to demonstrate conformity of the product;
- ✧ to ensure conformity of the Quality Management System; and
- ✧ to continually improve the effectiveness of the Quality Management System.

This clause suggests that the organization should determine applicable methods, including statistical techniques, and the extent of their use in monitoring and measuring Quality Management System processes and its performance. Then the organization could monitor information relating to customer perception as to whether the organization has met customer requirements. In the absence of applicable monitoring and measuring devices, is it possible to measure the performance of the Quality Management System in an educational institution?

Using ISO 9001:2000 quality management principles of "system approach" and "process approach", it would be pragmatic to study the "curricular processes" whilst systematically assuring educational quality. Since 2002, AIOU has adopted a "process approach" in the monitoring and measurement of its performance in the educational quality management system. Its MBA degree program is characterized by the "curricular processes" (Sou & Zhou, 2005a; Zhou & Sou 2005a), namely (1) Administrative Process, (2) Teaching Process, (3) Learning Process, (4) Research Process, and (5) Appraisal Process.

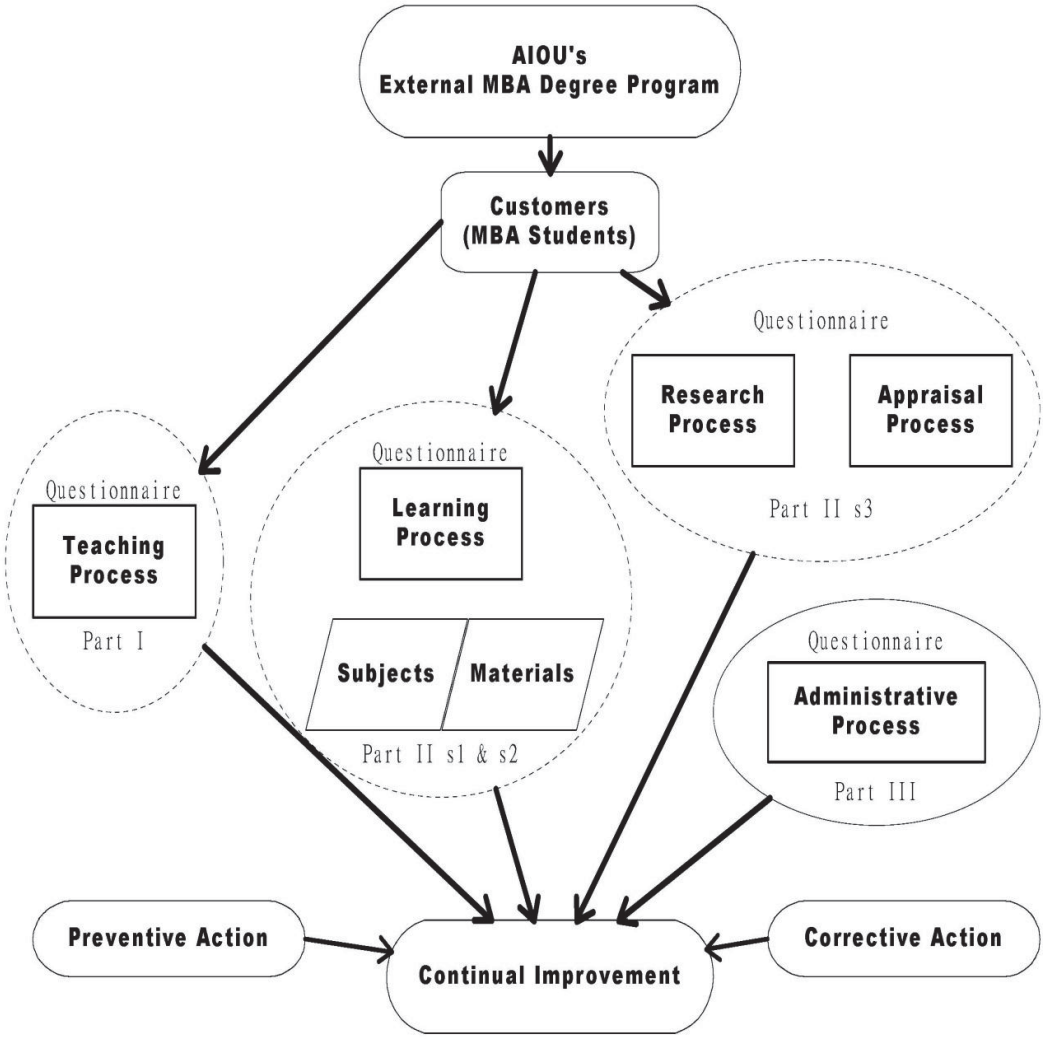
5.2.2 Program Evaluation Model for Management Education

In 2004, I embarked on a series of quantitative cum qualitative research studies (Appendix C) on a feasible prototype of program evaluation model for management education in the Mainland. Questionnaires for the "Conduct of Student Survey" were used to measure student learning experiences in the 2004-05 Graduate Survey on Perceived Importance and Satisfaction of AIOU's MBA Degree Cohort Program in the Mainland. Specimens of the tailor-made instruments (Questionnaire Part I – III used for the "Conduct of Student Surveys") in the previous research (Sou & Zhou, 2005b) are available for public reference in the "Sciencepaper Online[17]".

The 2004-05 Graduate Survey generated a return of 924 questionnaires for the study of "Teaching Process" and 1,074 questionnaires for the study of "Learning, Research and Appraisal Processes". It also generated a prototype of program evaluation model (Figure 15) for monitoring and measuring the various "curricular processes" in the Quality Management System of the AIOU's MBA degree program (Sou, 2006a-b; Sou & Zhou, 2007).

[17] A website administered by the Center for Science and Technology Development, Ministry of Education, People's Republic of China.

Figure 15: A Prototype of Program Evaluation Model for AIOU's MBA Degree Program



(Source: Sou & Zhou, 2005a)

5.2.3 Findings of 2004-05 Graduate Survey

The qualitative study of the 2004-05 Graduate Survey involved a literature review and focus group interviews of educational administrators, teaching staff or MBA candidates while the quantitative survey embraced questionnaire survey through convenience sampling. Combining qualitative and quantitative analysis, the "Perceived Degree of Importance/Satisfaction" of the MBA candidates of the subject program was worked out as per five matrixes:

- ✧ Teaching Process

    1. Perceived Degree of Importance/Satisfaction in Teaching Mode (Figure 16)

- ✧ Learning Process

    2. Perceived Degree of Importance/Satisfaction in Subjects (Figure 17)
    3. Perceived Degree of Importance/Satisfaction in Training Materials (Figure 18)

- ✧ Research Process

    4. Perceived Degree of Importance/Satisfaction in Research Mode (Figure 19)

- ✧ Appraisal Process

    5. Perceived Degree of Importance/Satisfaction in Appraisal Mode (Figure 20)

Figure 16: Perceived Degree of Importance/Satisfaction in Teaching Mode

Relative Important Teaching Mode	**Relative Unimportant Teaching Mode**
✧ Enthusiasm in Class ✧ Sound Preparation ✧ Logical Presentation ✧ Concise Presentation ✧ Stimulated Thinking	✧ Provision of References facilitating Effective Learning ✧ Provision of Advice on Extracurricular Activities ✧ Stimulation of Thinking by means of Assignment or Examination ✧ Application of Modern Pedagogy
Relative Satisfactory Teaching Mode	**Relative Unsatisfactory Teaching Mode**
✧ Enthusiasm in Class ✧ Sound Preparation ✧ Logical Presentation ✧ Effective Transfer of Knowledge	✧ Application of Modern Pedagogy ✧ Encouragement of Class Participation ✧ Provision of References facilitating Effective Learning ✧ Cultivation of Innovative and Research Capabilities ✧ Provision of Advice on Extracurricular Activities

Note: *The candidates reflected that they were not concerned much about the pedagogy. They were apt to the chalk and board methods in the teaching process. They liked to sit in class and listened. Reading the expatriate reference books did not necessarily help them solve the management problems in mainland China before of its unique business environment. Traditional research skills were adequate for China's business environment. The candidates did not think that extracurricular activities would add value to the subject program.*

(Source: Sou, 2006a & 2006b; Sou & Zhou, 2007)

Figure 17: Perceived Degree of Importance/Satisfaction in Subjects

Relative Important Subjects

Core Subject (2), (3), (9), (10) and (16):

- ✧ Organization and Management for Business
- ✧ Financial Management
- ✧ Human Resources Management
- ✧ Marketing Management
- ✧ Business Strategic Management

Relative Unimportant Subjects

Core Subject (4), (5) and (12):

- ✧ Quantitative Methods Analysis
- ✧ Production Management
- ✧ Small Business Management

Elective (2), (4) and (5):

- ✧ Auditing
- ✧ Comparative Management
- ✧ Total Quality Management

Relative Satisfactory Subjects

Core Subject (7), (8), (9) and (16):

- ✧ Economics for Business
- ✧ Organization Behavior
- ✧ Human Resources Management
- ✧ Business Strategic Management

Relative Unsatisfactory Subjects

Core Subject (4), (5) and (12):

- ✧ Quantitative Methods Analysis
- ✧ Production Management
- ✧ Small Business Management

Elective (2), (4) and (5):

- ✧ Auditing
- ✧ Comparative Management
- ✧ Total Quality Management

Note: *The candidates reflected that the Qualitative Methods Analysis was not applicable to their job settings while Production Management was pretty technical for them. Small Business Management was also inapplicable when the candidates mainly worked in sizeable companies or national enterprises. Auditing for the China-based corporations and national enterprises was unique and quite different from the auditing practices in the western countries. Again, the management styles in the East and West were not similar. Finally, Total Quality Management was a rather new subject in mainland China.*

(Source: Sou, 2006a & 2006b; Sou & Zhou, 2007)

Figure 18: Perceived Degree of Importance/Satisfaction in Training Materials

Relative Important Training Materials	**Relative Unimportant Training Materials**
Core Subject (3), (8), (9), (10) and (16): ✧ Financial Management ✧ Organization Behavior ✧ Human Resources Management ✧ Marketing Management ✧ Business Strategic Management	**Core Subject (4), (11), (12) and (15):** ✧ Quantitative Methods Analysis ✧ Cost Accounting ✧ Small Business Management ✧ Business Law
Relative Satisfactory Training Materials	**Relative Unsatisfactory Training Materials**
Core Subject (8), (9), (10) and (16): ✧ Organization Behavior ✧ Human Resources Management ✧ Marketing Management ✧ Business Strategic Management	**Core Subject (1), (4), (5), (6), (11) and (12):** ✧ Business Accounting for Executives ✧ Quantitative Methods Analysis ✧ Production Management ✧ Information Systems ✧ Cost Accounting ✧ Small Business Management

Note: *The candidates reflected that business accounting was a profession of a financial controller or chief finance officer of a sizeable company. Similarly, Cost Accounting was also their charter. Quantitative Methods Analysis was rather difficult because statistical tools were profound. Production Management was a bit technical for the managers while Information System would be shouldered by an engineer or chief technology officer of a sizeable company. Small Business Management was a subject for the owner of small business concern. Therefore, the candidates appeared to be unsatisfactory with the training materials in these subjects.*

(Source: Sou, 2006a & 2006b; Sou & Zhou, 2007)

Figure 19: Perceived Degree of Importance/Satisfaction in Research Mode

Relative Important Research Mode	**Relative Unimportant Research Mode**
✧ Guidance on the Choice of Research Methodology ✧ Guidance on the Application of Research Methods ✧ Close Supervision of Research Activities	✧ Guidance on the Choice of Research Tools ✧ Guidance on the Compilation of Thesis ✧ Guidance on Literature Review
Relative Satisfactory Research Mode	**Relative Unsatisfactory Research Mode**
✧ Guidance on the Choice of Research Topic ✧ Guidance on the Application of Research Methods ✧ Close Supervision of Research Activities	✧ Guidance on the Compilation of Thesis ✧ Guidance on the Choice of Research Tools ✧ Guidance on Literature Review

Note: *The candidates reflected that they were disappointed when their supervisors' guidance on the compilation of theses did not lead to high marks given by the examiners. Furthermore, the supervisor's guidance on the choice of research tools and literature review sometime attracted the internal and external examiners' criticism.*

(Source: Sou, 2006a & 2006b; Sou & Zhou, 2007)

Figure 20: Perceived Degree of Importance/Satisfaction in Appraisal Mode

Relative Important Appraisal Mode	**Relative Unimportant Appraisal Mode**
✧ Earnestness of Internal Examiner ✧ Earnestness of External Examiner ✧ Professional Knowledge of Internal Examiner ✧ Proliferation of Personal Knowledge	✧ Appropriateness of Viva Voce ✧ Arrangement of Viva Voce ✧ Length of Viva Voce ✧ Environment of Viva Voce
Relative Satisfactory Appraisal Mode	**Relative Unsatisfactory Appraisal Mode**
✧ Earnestness of Internal Examiner ✧ Earnestness of External Examiner ✧ Professional Knowledge of Internal Examiner ✧ Proliferation of Personal Knowledge	✧ Appropriateness of Appraisal Mode ✧ Appropriateness of Viva Voce ✧ Arrangement of Viva Voce ✧ Environment of Viva Voce

Note: *The candidates appreciated that both the internal and external examiners were conducting the viva voce in earnest. They found this aspect of appraisal process relative important and satisfactory. The candidates reflected that they were sometimes upset by the low marks given by the internal and external examiners. They perceived that the viva voce was an appropriate platform for them to earn high marks because a face-to-face discussion enabled them to present their theses to the internal and external examiners. The arrangement of the viva voce would be favorable when the candidates were allowed to have more interaction with the examiners. The candidates appeared to be more satisfactory if the environment of the viva voce would facilitate two-way communication with the examiners.*

(Source: Sou, 2006a & 2006b; Sou & Zhou, 2007)

All the returned questionnaires were processed by means of *t*-test, Analysis of Variance (ANOVA), Linear Multiple Regression Analysis, Correlation and Factor Analysis (Lohninger, 1999; Phillips, 1992; Vogt, 1993) through Statistical Product and Service Solutions (SPSS) version 15.0 for Windows. Quantitative analysis in the previous research was conducted on the conjecture that the educational quality of the AIOU's MBA degree program was positively correlated to its curriculum, training materials, teaching staff, research and appraisal processes.

Simply speaking, the sampled student regarded the program was a quality one when they perceived a relative high degree of satisfaction in the foregoing aspects of the subject program. In brief, the most significant findings of the quantitative analysis (Sou, 2006a-c; Sou & Zhou, 2005a-e & 2007; Zhou & Sou, 2005a-b) are as follows:

- ✧ Teaching Process (n = 924)
  - ⌑ Pearson analysis ($C_V = 0.874$; $\alpha < 0.05$) reveals that the "Perceived Degree of Importance in Teaching Mode" has a positive statistical correlation with the "Perceived Degree of Satisfaction in Teaching Mode"
  - ⌑ ANOVA ($R^2 = 0.72$) reveals that the "Overall Satisfaction of Teaching Process" = $0.03 + 0.23^{18} + 0.12^{19} + 0.13^{20} + 0.08^{21} + 0.09^{22} + 0.07^{23} + 0.09^{24} + 0.08^{25} + 0.08^{26}$

---

18 Variable 19 = Teaching process enhances the knowledge of student.

19 Variable 5 = Teaching process reflects the new development of research.

20 Variable 2 = Teaching process reflects the seriousness of the teacher.

21 Variable 7 = Teaching process reflects the focus of the subject.

22 Variable 10 = Teaching process satisfies the learners' needs.

23 Variable 13 = Teaching process cultivates the learners' abilities in renovation and researches.

24 Variable 6 = Teaching process reflects a clear learning objectives.

25 Variable 4 = Teaching process embraces appropriate course content.

✧ Learning Process (n = 1,074)

- Pearson analysis ($C_V = 0.905$; $\alpha < 0.05$) reveals that the "Perceived Degree of Importance in Subjects" has a positive statistical correlation with the "Perceived Degree of Satisfaction in Subjects"
- ANOVA ($R^2 = 0.72$) reveals that "Overall Satisfaction of Subjects" = $0.16 + 0.26^{27} + 0.12^{28} + 0.07^{29} + 0.15^{30} + 0.08^{31} + 0.10^{32} + 0.08^{33} + 0.04^{34}$
- Pearson analysis ($C_V = 0.42$; $\alpha = 0.86 > 0.05$) reveals that the "Perceived Degree of Importance in Training Materials" does not have a distinctive statistical correlation with the "Perceived Degree of Satisfaction in Training Materials"
- ANOVA ($R^2 = 0.65$) reveals that "Overall Satisfaction of Training Materials" = $0.14 + 0.15^{35} + 0.22^{36} + 0.12^{37} + 0.14^{38} + 0.09^{39} + 0.11^{40} + 0.08^{41}$

---

26 Variable 1 = Teaching process reflects the sound preparation of the teacher.

27 Variable 24 = Thesis Compilation and Research Methodology.

28 Variable 11 = Cost Accounting.

29 Variable 13 = Strategic Management of Change

30 Variable 10 = Marketing Management.

31 Variable 15 = Business Law.

32 Variable 2 = Organization and Management for Business.

33 Variable 17 = Corporate Recovery.

34 Variable 1 = Business Accounting for Executives.

35 Variable 7 = Economics for Business

36 Variable 20 = Thesis Compilation and Research Methodology

37 Variable 5 = Production Management

38 Variable 9 = Human Resources Management

39 Variable 12 = Small Business Management

40 Variable 1 = Business Accounting for Executives

41 Variable 13 = Strategic Management of Change

- ✧ Research and Appraisal Processes (n = 1,074)
  - ⌑ Pearson analysis (C_V = 0.627; $\alpha < 0.05$) reveals that the "Perceived Degree of Importance in Research and Appraisal Modes" has a positive statistical correlation with the "Perceived Degree of Satisfaction in Research and Appraisal Modes"
  - ⌑ ANOVA (R^2 = 0.708) reveals that the "Overall Satisfaction of Research and Appraisal Processes" = 0.15 + 0.37[42] + 0.14[43] + 0.17[44] + 0.12[45] + 0.12[46] + 0.03[47] + 0.07[48] + 0.08[49]

In parallel with qualitative study, the Graduate Survey 2004-05 led to the following findings:

1. Curriculum with its associated curricular processes was the premise of the AIOU's MBA degree program.

   The graduates perceived that the core subjects, electives and the various curricular processes (Teaching, Learning, Research and Appraisal) of the AIOU's MBA degree program were essential elements of a quality MBA degree program.

2. Training materials was closely linked with the educational quality of the AIOU's MBA degree program.

   The graduates perceived that a quality MBA degree program should have appropriate training materials.

---

42 Variable 22 = Research and appraisal processes enhance the students' overall knowledge level in business administration.

43 Variable 13 = Appraisal process in the form of 30-minute viva is appropriate.

44 Variable 20 = Appraisal results is appropriate.

45 Variable 1 = Research process reflects the proper mentoring in research methodology.

46 Variable 16 = Appraisal process reflects the professionalism and expertise of the internal examiner.

47 Variable 14 = Appraisal process reflects the sound preparation of the internal examiner.

48 Variable 3 = Research process reflects the proper guidance of the thesis supervisor in the conduct of the student's master research.

49 Variable 9 = Appraisal process reflects the appropriateness of viva voce.

3. Good teaching staff ensured the quality of the AIOU's MBA degree program.

   The graduates perceived that a quality MBA degree program should be delivered by good teaching staff.

4. Research and appraisal processes reflected the quality of the AIOU's MBA degree program.

   The graduates perceived that a quality MBA degree program should comprise an appropriate action research project (student learning experiences) and an effective assessment practices (student learning outcomes).

5.2.4 Coding of Secondary Data

In this research, Open Coding[50] pointed at the prototype of program evaluation model for the AIOU's MBA degree program (Figure 15) that was established on the principles and elements of ISO 9001:2000. The prototype illustrated the various processes of evaluating the quality of the subject program. Specifically, Axial Coding[51] led to the five curricular processes (Administrative, Teaching, Learning, Research and Appraisal) in the said model. Amongst these five processes, some of them could throw light to the student learning experiences, student learning outcomes, and assessment practices. Hence, "Selective Coding" was continually applied in the analytical framework of this research (Figure 14).

Selective Coding[52] narrowed down to the "Research Process" and "Appraisal Process" of the subject program. Why did "Selective Coding" lead to these two processes? It was because the findings of the Graduate Survey 2004-05 informed that these processes were respectively related to student learning experiences and student learning outcomes. They

---

50 "Open coding" was the analytic process through which concepts are identified and their properties and dimensions were discovered in data (Strauss & Corbin, 1998, pp. 101 & 217).

51 "Axial coding" was the process of relating categories to their sub-categories which occurred around the axis of a category, linking categories at the level of properties and dimensions (Strauss & Corbin, 1998, pp. 123 & 217).

52 "Selective Coding" was the process of integration and refining the theory (Strauss & Corbin, 1998, pp. 143-217).

were the foci of this research. I therefore screened the field-notes and analytic memos and categorized the concepts for analysis.

Simultaneously, review of those 331 MBA theses and the field-notes from the focus group interviews directed the focus of the study to the "Disclosure in Theses and Viva". Therefore, the first core-category to be "coded" is "Student Learning Experiences" incorporating two categories viz. "Perception on Curricular Processes" and "Disclosure in Theses and Viva" (Figure 14). In the following paragraphs, there will be elaborations of the sub-categories under these two categories (Figure 14).

5.2.4.1 Teaching Process

Firstly in the "Teaching Process" (Figure 16), the MBA candidates perceived that "Stimulated Thinking" was relative important but "Stimulation of Thinking by means of Assignment or Examination" was relative unimportant. They perceived that "Effective Transfer of Knowledge" was relative satisfactory while "Cultivation of Innovative and Research Capabilities" was relative unsatisfactory. This reflected that the MBA candidates were not keen on examination to test their thinking. However, the program already offered them the teaching and learning processes to stimulate their thinking. Furthermore, they were satisfied that the processes could effectively transfer knowledge to them. However, the processes did not necessarily cultivate their innovative and research capabilities.

5.2.4.2 Learning Process

Secondly in "Learning Processes" (Figure 17 and Figure 18), the MBA candidates perceived that some subjects and/or training materials were relative important and satisfactory. The relative important subjects or training materials included "Organization and Management for Business, Financial Management, Organization Behavior, Human Resources

Management, Marketing Management, and Business Strategic Management". The relative satisfactory training materials included "Economics for Business, Organization Behavior, Human Resources Management, Marketing Management, and Business Strategic Management". Consistently, the relative unimportant and unsatisfactory subject and training materials included "Quantitative Methods Analysis". It reflected that the MBA candidates did not perceive the importance and satisfaction of this subject and its training materials.

5.2.4.3 Research Process

Thirdly in the "Research Process" (Figure 19), the MBA candidates perceived that the "Guidance on the (a) Choice of Research Methodology, (b) Application of Research Methods, and (c) Close Supervision of Research Activities" was relative important and satisfactory. On the other hand, they perceived that the "Guidance on the (a) Choice of Research Tools, (b) Compilation of Thesis, and (c) Literature Review" was relative unimportant or unsatisfactory.

5.2.4.4 Appraisal Process

Fourthly in the "Appraisal Process" (Figure 20), the MBA candidates perceived that "Proliferation of Knowledge" was relative important and satisfactory. However, they subjectively perceived that the "Appropriateness of Appraisal Mode and Viva Voce" and the "Arrangement and Environment of Viva Voce" were relative unimportant and unsatisfactory. The relative unsatisfactory student learning experiences in the "Research and Appraisal Processes" was worth further exploration more objectively.

5.2.5 Perception of the Candidates

In the focus group interviews, the registrars commented that the findings of the 2004-05 Graduate Survey were still prevailing after 2005. In the 331 MBA theses reviewed from 2007 to 2008, it was observed that the sampled MBA candidates customarily acknowledged the teachers' edification in the teaching activities, the enlightenment from the subjects and training materials in the learning activities, and the supervisors' guidance in the research projects.

In other words, the sampled candidates appeared to enjoy the teaching, learning and research processes. When the 331 candidates appeared before the academic panels, they usually reiterated the acknowledgements during their "10-minute narration in the viva". When they were questioned about their experiences in the "Appraisal Process", they reflected that the viva enabled "proliferation of their personal knowledge".

Despite the above positive cases, there were some negative cases pertaining to the student learning experiences in the viva. One of the registrars, who had close liaison with the educators and students of the Mainland, quoted some unfavorable student learning experiences according to the MBA candidates and their supervisors. They commented that there were different assessment criteria between the thesis supervisors and the internal or external examiners. The registrar added that this phenomenon had been a long-standing problem since 2003.

In 2003, there was once a seminar for academic panels hosted by the AIOU's rector with attendance of five internal examiners, seven external examiners and two registrars. In the seminar, the rector debriefed the members of academic panels on the opinions collected from Annual Educators' Conference 2003 that was held in the Mainland. The opinions concerned the master research, thesis and viva are translated as follows (Source: Excerpt of

the Chinese Minutes of a Seminar for Academic Panels dated 05.08.03):

1. There was underrating in the viva. Generally speaking, the academic panels in Macau under-marked the theses at about 10 scores in contrast with the thesis supervisors in the Mainland. For instance, the supervisor awarded 80 marks to the candidate whereas the academic panel only awarded 70 marks to the same candidate.
2. There were variable assessment criteria between different academic panels. The marking difference was quite significant.
3. There were discrepancies between the scores and the comments in the assessment forms. Some theses with good comments earned low scores while some theses with rather bad comments earned high scores.
4. Some examiners did not study the theses well in advance and prepare for the viva thoroughly. Some even started reading the theses during the viva.
5. Some "negative" comments such as "impractical research" or "unrealistic solutions" from the academic panels hurt the feeling of the candidates who felt that their efforts were not appreciated.
6. Academic panels laid too much emphasis over the format of the theses and the utilization of formal research methods.

5.2.6 Interaction between the Examiners and Candidates

The foregoing opinions were enlightening in this research. In the focus group interviews, the registrars quoted that some candidates had derived a lot of pleasure from the viva though some did not. When the examiners were invited to exchange their experiences, four approaches of their interaction with the candidates (Figure 21) were generalized with the aid of field-notes and analytic memos.

Figure 21: Four Approach of Interaction between Examiners and Candidates in the Viva

Approach	Interaction
Tell and Sell	The examiner mainly intents to tell the candidate what is "right and wrong" with the action research and what changes are necessary. The examiner strives to persuade the candidate to accept the judgments by assuming the role of all-knowing judge and using various persuasive techniques to get the candidate's acceptance. The candidate sometimes becomes defensive because of the one-sided discussion. This approach rests on the assumption that the examiner's judgment is complete and accurate and that the candidate cannot make any significant contributions. Generally, it does not result in the candidate's assessment form for proposed improvement activities nor result in consensus or acceptance of the rating.
Tell and Listen	The examiner serves as a judge and tells the candidate what is right or wrong regarding the action research. It assumes that the candidate has all of the answers and merely needs to be communicated with. After telling, the examiner does not press the candidate to accept the judgments, but allows expression of feelings and opinions. Even if the candidate becomes defensive, the examiner listens to the complaints and remains understanding and sympathetic. This approach assumes allowing the candidate to express feelings, in the context of the examiner being empathetic, is likely to prompt acceptance of the judgments. There is more 2-way communication, but the candidate is not contributing through self-evaluation or reflecting on performance.
Problem-Solving	This approach stresses 2-way communications and shifts the examiner's role from judge to helper. The examiner's task is to prompt the candidate to do a self-evaluation and identify strengths and areas of the action research where problems may exist. It also emphasizes the candidate identifying goals for improvement. The viva begins with the candidate sharing a self-evaluation. The examiner assists by drawing out ideas, providing reflective analysis, and supplementing incomplete information. Next, the candidate suggests areas for improvement and a course of action for doing so. A limitation of this approach is that the candidate may not identify the problems which the examiner feels are to be addressed.

Approach	Interaction
Composite	This approach combines the foregoing ones, with a focus on collaboration in identifying representative student learning outcomes information and developing solutions to areas which need improvement. The viva begins by the candidate conducting self-evaluation. The examiner then supplements information that the candidate has not included. After the student learning outcomes are identified through the candidate's self-evaluation and the examiner's input, collaborative problem solving begins for those areas to be improved. It is also important for the examiner to highlight those areas in which student learning outcomes had met expectations and provide positive reinforcement. The candidate's developmental needs and wants, as well as career aspirations, are also discussed.

(Source: Field-notes of the subject research)

The appraisal process in the form of a viva can be a good opportunity for interaction and communication between an examiner and a candidate. An examiner can take the chance to formally inform the candidate how he or she has conducted the action research against the prescribed assessment criteria. On the other hand, a candidate can be guided to go through a self-evaluation. The approach of an examiner adopted towards a viva can affect its effectiveness as well as the student learning experiences.

All the examiners welcomed the Composite Approach when they were invited to choose amongst those four approaches detailed in Figure 21. During the viva, the examiners could collaborate with the candidate in identifying student learning outcomes and areas which needed improvement. The candidate began the viva by conducting self-evaluation. The examiners then supplemented information that the candidate had not included.

After the student learning outcomes were identified through the candidate's self-evaluation and the examiners' input, collaborative problem solving addressed those areas to be improved. It was most favorable for the examiners to highlight those areas in which student learning outcomes had met expectations and provide positive reinforcement. Moreover, the candidate's developmental needs and wants, as well as career aspirations, were also discussed.

The Composite Approach would appear to be the most favorable one but one can predict that a viva under such an approach may last for an hour. Due to time limitation, some examiners used the 'Tell and Sell' approach rather than the others. The examiners said that the candidates occasionally asked for more time to explore their student learning outcomes in the viva. This phenomenon was partially attributed to the design (Dimmock, 2000; Peterson et al., 1999a-c; Sou, 2005c-d) of the appraisal process as well as the time constraint of the viva.

Quantitative and qualitative analyses in this chapter throw light to the favorable and less favorable student learning experiences in various curricular processes of the subject program. Amongst the processes, some candidates showed their dislikes to the appraisal process for the assessment of student learning outcomes. Therefore in the following chapter, the focus of study will be shifted to the student learning outcomes. Through the study of the assessment forms for the diagnosis of student learning outcomes, there can be a greater understanding about the assessments made by the thesis supervisors, internal and external examiners and the actual learning outcomes of the candidates.

VI. DATA ANALYSIS – STUDENT LEARNING OUTCOMES

6.1 Synopsis

The previous chapter analyzes the data about "Student Learning Experiences". This chapter reports the processes and findings of data analysis regarding Student Learning Outcomes (SLO). To recap, this core category (Figure 13 and Figure 14) has 3 categories called "Assessment Forms", "Viva Performance", and "Learning Styles". These three categories respectively have six, four and five sub-categories:

- ✧ Assessment Forms – (1) Choice of Title of Thesis and Abstract, (2) Introduction and Literature Review, (3) Research Methods and Design, (4) Data Processing and Analysis, and Discussion (5) Conclusion and Recommendations, and (6) Compilation, Format and Style of Thesis.
- ✧ Viva Performance – (1) Attitude, (2) Skills, (3) Knowledge, and (4) Six Other Perception viz. Achievements, Potential, Intelligence, Aptitude, Motivations and Personality.
- ✧ Learning Styles – (1) Diagnostic Assessment, (2) Intuitive Learners, (3) Incidental Learners, (4) Retrospective Learners, and (5) Prospective Learners.

6.2 Student Learning Outcomes

An assessment form is the instrument for the assessment task in the subject program. It consists of six aspects to be assessed by the examiners. Each aspect carries 10 to 30 marks contributing to a full mark of 100 (Figure 22). A candidate earned a total score of 80 or above will be graded as "outstanding" while a total score of 70 to 79 will be graded as "credit". The pass mark is 60 to 69 while a candidate will be failed with a total score of 59 or below. In this section, those six aspects contributing to the total score will be

analyzed communally.

Figure 22: Six Aspects of a Thesis to be assessed by Examiners

Item	Aspect to be Assessed by Examiners	Maximum Contributory Score
1	Choice of Title of Thesis and Abstract	10 of 100
2	Introduction and Literature Review	20 of 100
3	Research Methods and Design	10 of 100
4	Data Processing and Analysis, and Discussion	30 of 100
5	Conclusion and Recommendations	20 of 100
6	Compilation, Format and Style of Thesis	10 of 100

6.2.1 Study of the Assessment Forms

Scrutiny of 331 assessment forms revealed that about one-fifth of the candidates tended to work on a rather big topic. For instance, a big topic might be a study on the whole industry involving state-owned enterprises, small and medium enterprises. To name a few but not all, titles of those big topic theses were (Figure 23):

Figure 23: Some Big Topics of MBA Research Projects

"Marketing Strategy for Digital Television in China"
"Corporate Culture fits for International Competition"
"Analysis of Cost Management in Modern Enterprises"
"How to attract Talents to join Newly-Founded Small Enterprises"
"Study on e-Business and Profit Model of Small-and-Medium Enterprises in China"
(Source: Assessment forms of the sampled MBA candidates)

With reference to the field-notes, the examiners concurred that an MBA research project should be confined to a well-defined and specific topic. Too large a project became unmanageable. Examiners opined that the supervisors should control the scope of study on the candidates' research project. The registrars shared the same view. However, the registrars said that it was sometimes difficult to change the mindsets of the supervisors. Therefore, some candidates were guided to work on large topics.

AIOU stipulated that each thesis should have an abstract in Chinese and English. Examiners found that the great majority of English abstracts were questionable. Grammatical errors and structural problems were common in the English abstract. This is the excerpt of an English abstract (Figure 24):

Figure 24: Excerpt of the English Abstract of an MBA Thesis

> *Flow out to go into in great quantities along with fast development and foreign capital bank of the our country share system commercial bank, bank the talented person's fluxion is increasingly multifarious, how draw on and keep a talented person, become each line governor to pay attention to of focus. The establishment rises I set since fair reasonable, and then transfer employee to work well the actively sexual salary system is the key which resolves that problem. The salary system is the most basic incentive system in the …*
>
> *The thesis led to go into modern salary theories first and carried on elaborating to it, secondly to the JS bank branch office present condition in Henan carry on analysis, pass and same profession salary the contrast discovered shortage and margin, finding out the problem output reason, combining the oneself body actual circumstance, drawing lessons from a same profession experience of the forerunner, being considering well part of, individual of, at present benefits foundation up, fixed attention on bank and the whole benefits of the employee and farsighted benefits design a set of system project, it includes six modes, respectively is wages total amount management style, wages management and performance mode, wages structure adjust mode, post value the evaluation mode, post salary management style, post grade adjust mode; 6 constitute an organic whole, pass the organic movement of each statures mode, the wages of the formation JS branch office science of bank Henan assign mode of whole circulate mechanism. …*

(Source: Thesis of a sampled MBA candidate)

The above exhibit demonstrates the English proficiency of the candidate concerned. Sometimes, the candidates did not have the proficiency of writing an English abstract. They would use the on-line translation tool to transliterate their Chinese abstracts. Therefore, the English abstracts were not always comprehensible. Understandably, English is the second language of the candidates from the Mainland hence it is reasonable to assume that their English proficiency may not meet the generally acceptable academic

standard. Above all, the subject program did not prescribe that the candidates should attain proficiency in English before enrolment.

A newspaper with daily average distribution of 280,034 in Hong Kong (am730, 2007) quoted the findings (Figure 25) of a Sino-Australia research agency about the English proficiency of the Asian graduates. The agency through Australia China Alumni Association found that the English proficiency of over one third of Asian graduates in Australia were poor. Actually, they should not be admitted to the Australian universities. It also commented that over 40 per cent of the graduates from mainland China could not work in Australian with their English proficiency.

Figure 25: English Proficiency of Asian Students

Item	Country or City	Percentile of Asian Students without English Proficiency
1	South Korea	55.5
2	Thailand	50.9
3	Nepal	47.9
4	Taiwan Province of China	47.4
5	**Mainland China**	**43.2**
6	Hong Kong Special Administrative Region	42.9
7	Japan	36.8
8	Vietnam	32.9
9	Indonesia	32.0
10	Malaysia	23.5

(Source: am730, 2007)

The above survey shows that over 40 per cent of them are not proficient in English too. In Hong Kong Special Administrative Region, all providers of full-time degree programs except the Open University of Hong Kong make completion of a defined number of credits in English language course a requirement for graduation (Butcher and Hope, 2006). In contrast, AIOU at Hong Kong and Macau Special Administrative Regions reckons that

English language training is not a must for the subject program.

Administrators of AIOU considered that adult learners like the candidates of the subject program already in full-time employment could decide for themselves whether they wished or needed to take language-enhancement courses to attain their graduation goals. Nonetheless, the examiners believed that career managers in mainland China should have acceptable English proficiency in order to do global business. The possible way forward may be incorporation of a guidance session into the subject program. Through the guidance session, AIOU may introduce the importance of an acceptable English proficiency to the students and recommend to them the possible ways for their language enhancement.

As mentioned in the above paragraphs, the candidates sometimes introduced rather big research topics. Examiners and registrars reasonably believed that the candidates' macro mindsets were stemmed from the long-standing planned economy in mainland China (China Education Center, 2011; Gale Group, 1999). This macro-mindsets phenomenon was prevalent amongst those candidates working in the public sector or state-owned enterprises. On the other hand, the candidates working in the private sector or small-or-medium enterprises were more likely to choose specific topics for their research. The possible way forward may be a change of the candidates' mindsets. The thesis supervisors may encourage those candidates who are working in public sector or state-owned enterprises to be entrepreneurs.

Though the other research topics were comparatively specific, most candidates introduced a wide literature review to support their research projects. Examiners often found that the literature reviews were not entirely relevant to the research topics. When the examiners questioned the candidates about the relevance of their literature reviews to the

research topics, the candidates appeared to be a bit perplexed. They were most likely to associate their literature reviews to those subjects that they perceived to be relevant and important (Figure 17), such as Financial Management, Organization Behavior, Human Resources Management, Marketing Management or Business Strategic Management. However, it was always not the case. The examiners discovered that there was always one primary theoretical theme and one secondary theme in each thesis.

When the examiners further questioned the candidates about their reflections upon the "Research Methods and Design" and "Data Processing and Analysis, and Discussion", about half of the sampled candidates were not able to differentiate qualitative from quantitative research. They usually claimed that qualitative cum quantitative methods had been used in their research. They often misunderstood that quantifying qualitative data for analysis was quantitative research. Actually in their cases, there were no statistical procedures adopted in the data processing and analysis.

"Data Processing and Analysis, and Discussion" is a significant aspect to be assessed which carries a maximum of 30 scores contributing to the total mark. Amongst the 331 assessment forms, the modal score of this aspect was 18 to 21. When the candidates discussed the issues or problems to be tackled in their action research projects, the examiners perceived four kinds of reflection (Hatton & Smith, 1995) that were manifest in the theses about the candidates' learning in action (Figure 26).

Hatton and Smith (1995) define "reflection" as "deliberate thinking about action with a view to its improvement". They also propose that the candidates' discussion with a more experienced person can facilitate dialogic and critical reflection. In the subject program, the thesis supervisor is the experienced person who can facilitate a candidate's dialogic and critical reflection. However, the thesis supervisor would also hinder the candidate's

dialogic and critical reflection if asking the candidate to reflect too early and too broadly.

Figure 26: Candidates' Reflection and Examiners' Perception

Item	Candidates' Reflection	Examiners' Perception
1	**Descriptive Writing**	The candidates did not have any reflection.
2	**Descriptive Reflection**	The candidates provided some reasons based on personal judgments.
3	**Dialogic Reflection**	The candidates explored possible reasons and approaches that might be rooted in their reading of the relevant literature.
4	**Critical Reflection**	The candidates explored reasons and approaches and the underlying assumptions and concepts. The exploration was based upon an evaluation of context which took account of personal, historical and social influences.

(Source: Hatton and Smith, 1995; Analytic memos of the subject research)

Hence, the thesis supervisor should emphasize the importance of providing the candidates with practice in the skills of reflection. Otherwise, the candidates tended to utilize descriptive writing or descriptive reflection in conducting their action research projects. Examiners opined that thesis supervisors played an important role in guiding the candidates in conducting professional research. Both the internal and external examiners sometimes remarked that they were failing the thesis supervisors but not the candidates.

6.2.2 Review of Candidates' Performance

In the focus group interviews, the examiners expressed that they were more than pleased to meet candidates being reflective practitioners (Schön, 1983, 1988 & 1991). Reflective practitioners were those candidates as prospective career managers who were able to learn how to frame and reframe the complex issues or ambiguous problems and then interpret and modify their practice as a result. The most satisfactory and promising candidates were prospective career managers with "reflection in action" as well as "reflection on

action" (Schön, 1983, 1988 & 1991). The former reflection facilitated immediate decision-taking while the latter reflection provided a longer and deeper view.

Reflective candidates (Schön, 1983, 1988 & 1991) with a longer and deeper view can be nurtured to be "understanding seekers" (Entwistle, 1987 & 1992). Examiners also expressed that they "loved" to meet candidates as "understanding seeker" than "knowledge seeker" (Entwistle, 1987 & 1992). Knowledge seeking candidates searched for facts and information. Their learning was mechanical and superficial. They were not interested in speculating, playing with theories or searching for deeper meanings. In comparison, understanding seeking candidates were less interested in facts and more interested in searching for personal meaning in their action research projects. They related what they had learnt from the subject program to their earlier experience and explored potential connections, linkages or discrepancies.

Knowledge seeking candidates appeared to be serialists who had built up their picture of a research topic gradually while understanding seeking candidates appeared to be holists who had grasped a global picture before exploring any detail (Entwistle, 1987 & 1992; Gibbs, 1990). Therefore, understanding seeking candidates tended to be intrinsically motivated in their action research projects. They were likely to be deep problem solvers and to be innovative and constructive in reaching sound conclusions and making appropriate recommendations (Entwistle, 1987 & 1992).

It appears that understanding seekers may have more desirable student learning outcomes than the knowledge seekers. It is noteworthy that these two learning orientations are not mutual exclusive. Though there were only a few candidates who were equally comfortable with both learning orientations, some candidates showed a willingness to switch to the style of learning demanded by the assessment task if the thesis supervisors offered proper

guidance. Both the internal and external examines had quite identical comments on this phenomenon.

The above phenomenon substantiated the notion (Entwistle, 1987) that most candidates with a predominant learning orientation might be influenced by the conditions of thesis supervision and assessment. Thus, it was rational for the examiners to criticize the thesis supervisors who failed to encourage deeper learning (Gibbs, 1990) during the research processes.

In the subject program, the research processes undertaken by the candidates lead to the "Conclusion and Recommendations" in a thesis. This aspect of student learning outcomes to be assessed carries a maximum mark of 20-out-of-100. This assessment task for the examiners to diagnose the candidates' learning outcomes promotes deeper learning. Unfortunately, examiners perceived that about one-third of sampled candidates had failed to demonstrate personal understanding. Sometimes, they appeared to be reproducing the thesis supervisors' understanding (Entwistle and Marton, 1994). Examiners found that some candidates used visualization and structuring that that understanding went beyond those analyzing processes to perceiving the whole of the domain of knowledge.

6.2.3 Observation of Candidates' Learning Outcome

Using Structure of the Observed Learning Outcome (Biggs, 1987, 1996 & 1999; Biggs & Collis, 1982), the examiners could only award the overall grading of "outstanding" to a very limited number of candidates. With reference to Bigg's schema (Figure 27), "Extended Abstract" student learning outcomes could be regarded as outstanding performance but it was scarce in the sampled candidates. "Pre-Structural" student learning outcomes could hardly be regarded as satisfactory. Examiners did not hesitate to fail those candidates with pre-structural student learning outcomes.

Figure 27: Five Levels of Bigg's Schema of Structure of the Observed Learning Outcome

Level	Bigg's Schema	Findings of the Subject Research
1	**Pre-Structural**	The candidates did not attack their action research appropriately. They did not understand the points or the relevant theories.
2	**Unstructural**	The candidates picked up or used one or a few aspects of the theories relevant to their action research. Their understanding was nominal.
3	**Multistructural**	The candidates learned several aspects of the theories relevant to their action research but they treated those aspects separately.
4	**Relational**	The candidates integrated the components of their action researches into a coherent whole with each part contributing to the "Conclusion and Recommendations".
5	**Extended Abstract**	The candidates re-conceptualized the integrated whole at the relational level to a higher level of abstraction. Their "Conclusion and Recommendations" enabled generalization to a new topic or area; or the generalization was turned reflexively on oneself.

(Source: Hatton and Smith, 1995; Analytic Memos of the subject research)

Are there any other aspects of student learning contributory to desirable student learning outcomes? Analytic memos provide some hints to this question. Learning orientation is more likely to be associated with a set of characteristics of a learning environment than with just one characteristic (Carpenter and Tait, 2001; Entwistle, 1987 & 1992; Ramsden, 1988 & 1992). By the way, it is necessary to consider the learning environment of the candidates and their learning orientation which may influence the nurture of "Understanding Seekers" as well as "Extended Abstract" (Carpenter and Tait, 2001; Entwistle, 1987 & 1992; Ramsden, 1988 & 1992) student learning outcomes?

Before studying other aspects of student learning and answering the above question which may influence student learning outcomes, I wished to first complete the analogy of the remaining aspect of the assessment form, namely "Compilation, Format and Style of Thesis". This aspect carries a maximum score of 10 but it always brings up common errors that could fail a candidate. Some examiners awarded six or lower marks to the candidates who did not present their theses in the correct format and style therefore the candidates could not get an overall pass mark. Obviously, the format and style should not be critical failure factors of a thesis. Though some examiners particularly the internal ones of academic background were quite conscious with the format and style, they would not simply fail those candidates who had not adopted a 5-Chapter format in writing their theses.

Examiners remembered that some candidates previously tended to use 3-chapter format in writing theses. They compiled their theses with a prolog followed by the main text and then ended the theses with an epilog. Both the prolog and epilog were rather short and hindered the readability of a traditional thesis in 5-chapter or 7-chapter format. Either internal or external examiners also found such writing style annoying. They opined that 5-chapter thesis was the most readable format.

Like the 3-chapter format, another writing style was once very prevalent and also annoyed the examiners. Some candidates working in public sector or state-owned enterprises were used to start their theses with the following preface (Figure 28):

Figure 28: Translation of the Preface of an MBA Thesis

This research is based on Marxism-Leninism, Mao Zedong Thought, Deng Xiaoping Theory, and Jiang Zemin's Theory of Three Represents …

(Source: Theses of sampled MBA candidates)

When the examiners questioned the candidates about the relevancy of Marxism-Leninism, Mao Zedong, Deng Xiaoping, and Zhiang Zemin to their research projects, the candidates answered that it was a practice for them to start an essay like this. Some even confessed that those philosophers or politicians definitely played no role in the theses concerned. The candidates added that quoting those philosophers or politicians in daily life was an ideology that had been grounded in their organizations and formal writings. Definitely, such writing style would not fail a candidate if the materials in the thesis demonstrated the desirable student learning outcomes.

On the contrary, the examiners would definitely fail the candidates whose materials in the theses were proved to be plagiarized. The subject research uncovered some cases of plagiarism in the theses. Examiners said that they could bear descriptive writing or descriptive reflection (Hatton & Smith, 1995) as per the 4 level of action learning outcomes. (Figure 26). They could also stand the candidates who copied a lot from the literature that reflected pre-structural or unstructural learning outcome as per Bigg's schema (Figure 27). However, the examiners did not accept plagiarism without proper referencing.

The examiners understood that the Chinese candidates might not familiarize themselves with the western referencing protocols. Candidates sometimes mistook that they only needed to reference actual quotations, but not ideas. The worst scenarios involved non-attributing of materials, or patching the ideas of others together into a thesis. Though the candidates had sometimes not intended to deliberately cheat, improper referencing did constitute poor academic writing and was a reflection of a lack of understanding of academic conventions.

The worst scenarios involved the candidates copying or buying papers to submit as their own works. Examiners could not tolerate such frauds but unfortunately, examiners detected such cases of plagiarism. For instance, the examiners once detected a candidate, a mid-career traffic policeman, who had committed plagiarism in the compilation of his thesis. The subject candidate copied about 50 per cent of the thesis of another candidate. The examiners commented that it was a shame for a law enforcement officer who upheld the justice in his career but stealing another candidate's intellectual property.

Examiners also regretted to find that some candidates had stolen or bought the other people's intellectual property. An internal examiner was frustrated with the candidates' plagiarism. He quoted the title of a movie, namely "Catch Me If You Can" whilst assessing the candidates. He endeavored to sort out those candidates who plagiarized in their theses and regarded their misdemeanors had challenged the assessment tasks of the examiners.

To uphold the assessment system, the internal examiner used some powerful online search engine to spot those candidates who bought theses through the internet. The particular examiner detected six cases of cheating or plagiarism amongst seven candidates whom he had examined in a morning. Did such misdemeanors reflect the personal qualities particularly the learning attitudes of the candidates?

The learning attitudes of those candidates who had committed plagiarism appeared to be questionable. The examiners feared that questionable learning attitudes would not lead to favorable learning outcomes. As mentioned in page 120, student learning outcomes could be fully assessed after the perusal of the assessment forms. At this stage, the focus of data analysis can be shifted to the personal qualities of the candidates reflected in their viva performance. After all, the candidates' performance in the viva voce reflected their learning outcomes through the subject program.

As planned, participants of the focus group interviews were encouraged to think about the student learning outcomes (Appendix B) of the candidates in nine aspects, namely **Attitudes**, **Skills**, **Knowledge**, **Achievements**, **Potential**, **Intelligence**, **Aptitudes**, **Motivations**, and **Personality** (Brown, Bull & Pendlebury, 1997). Amongst the above nine personal qualities, the first three with the acronym ASK (Attitude, Skills, Knowledge**)** appear to be comparatively significant than the others.

Content analysis of the field-notes reveals that the examiners frequently comment on the candidates' ASK during the focus group interviews. This phenomenon proposes that the examiners attach importance to ASK rather than the other personal qualities of the candidates. This proposition contrasts the original conjecture in literature review (Chapter III) with the same acronym in which A represents "Abilities". In this chapter, "Attitudes" is more appropriate than "Abilities" for the study of student learning outcomes. Coding through the field-notes and analytic memos informs that the examiners appreciate the learning attitudes of the candidates rather than their abilities.

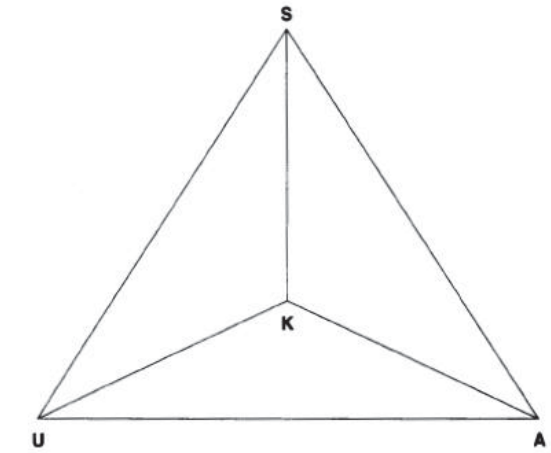
The examiners opined that the candidates should have the abilities of completing an action research on their own after completion of the subject program. However, their learning attitudes might have driven them to plagiarism even though they did possess the skills and the knowledge in doing a master research. Obviously, the examiners see the importance of candidates' ASK in higher education. Hence, it should be worthwhile to study ASK and their relationship with the candidates' learning outcomes.

First, **skills** can be defined as "an organized and coordinated pattern of mental and/or physical activity" (Haper & Kember, 1989). This definition makes no mention of "fitness for purpose" or "fitness of award". Brown et al., (1997) prefer to define skills as "constructs used to describe goal-directed sequences of action that may be learnt and

routinized". Once the candidates learned them from the subject program, they should be able to adjust their action to the tasks in hand and solve the problems in their action research projects. In other words, problem-solving by a routine method is a common example of "skills in action".

All skills involve cognitive, perceptual, and motor components in varying proportions. The cognitive component of a skill is directly proportional to the complexity of the skill. Cognitive and perceptual components are more important than motor component when the candidates are attempting to solve a problem in their action research projects. All the skills earned from the major and electives of the subject program are related to each other in a complex way. Brown et al. (1997) simplify their relationships by using a tetrahedron (Figure 29).

Figure 29: Tetrahedron showing Relationships of Attitudes, Skills, Knowledge and Understanding



(Source: Brown et al., 1997)

In higher education like the subject program, the scope of skills is enlarged to cognitive skills, such as reading, writing, problem-solving in the research projects. Social skills are also demonstrated in discussion or negotiation during the problem-solving processes. Thus, integral skills gained by the candidates become not just the basic but the centerpiece of much of the work in the subject program. Use of **Skills** (S) in different contexts is based upon the cornerstones of **Attitude** (A), **Knowledge** (K) and **Understanding** (U).

The essential cognitive skills upon SKU are creativity, critical thinking, evidence evaluation, information handling, problem-solving, and rational argument (Brown et al., 1997). The essential social skills upon SAU are working or communicating with the others in various roles (Brown et al., 1997). The essential cognative skills upon SAK are achieving, independence, initiative, risk-taking, and willingness to change (Brown et al., 1997).

Cognitive, social and cognative skills contribute to a candidate's capacity to self-evaluate and to clarify values. Finally, self-evaluation, value clarification and the essential **skills** provide the structure of KUA for the career choices and personal development of the candidates. The SKUA tetrahedron proposes that skills and their relationships may be seen from different perspectives. (Brown et al., 1997)

In the subject program, the skills of the candidates are taught through the medium of the core subjects and electives of an MBA degree program. Each subject has different scopes and interpretation of skills. Problem-solving in "Financial Management" is not identical with "Human Resources Management". Evidence evaluation in "Marketing Management" is not identical with "Business Strategic Management".

Nonetheless, there is a commonality for the high-level skills at the heart of higher education. At the heart of this research, the examiners and the registrars reflected a long-standing and traditional expectation from the educationists about the ASK (Attitudes,

Skills and Knowledge) on the candidates (Figure 30). This expectation appeared to be reasonable and predictable. All the examiners went through a long learning journey in their life; they in the focus group interviews unanimously agreed with such expectation on higher education.

Figure 30: Expectation about the ASK on the Candidates of Higher Education

> *… education … gives a man a clear conscious view of his own opinions and judgments, a truth in developing them, an eloquence in expressing them, and a force in urging them. It teaches him to see things as they are, to go right to the point, to disentangle a skein of thought, to detect what is sophistical, and to disregard what is irrelevant. It prepares him to fill any post with credit, and to master any subject with facility. It shows him how to accommodate himself to others, how to throw himself into their state of mind, how to bring before them his own, how to influence then, how to come to an understanding with them, and how to bear with them …*
>
> (Source: Newman, 1853)

The other six aspects of the analysis on the student learning outcomes are directed to the effects of **Motivations** and **Personality** on learning (Wittrock, 1986). Literature reveals that the way candidates perceive themselves and the way they account for their academic enhancement has a strong bearing on their motivations (Berliner, 1996; Wittrock, 1986). According to the supervisors, examiners and registrars of AIOU, the candidates expressed a genuine concern over the inadequate provision for tertiary education in mainland China. When the inadequacy was exacerbated, the candidates traditionally set a high value upon tertiary education.

In such circumstances, the candidates were naturally eager to enhance their career opportunities through investment in training and intellectual development (Butcher & Hope, 2006; Dhanarajan, 1998). Therefore in the subject program, **Motivations** of the candidates in seeking management education opportunities appeared to be high. In summary, the examiners and the registrars of AIOU believed the candidates were motivated to participate in the subject program.

Study of the relevant literature proposes that candidates are likely to initiate learning, sustain it, direct it and actively involve themselves in it when they believe that **Achievements** are caused by their own efforts. Moreover, praise, reward or other positive reinforcements are likely to enhance **Motivations** only if candidates perceive them to be related to factors over which they have control. Thus, building up candidates' sense of control over their own work, giving them opportunities for their own learning, and helping them to develop their **potentials** can all help to make them more successful and effective learners (Butcher & Hope, 2006; Dhanarajan, 1998).

Finally, **Intelligence** and **Aptitudes** are two aspects which influence the student learning outcomes in a subtle way. Intelligence is believed to be inborn to a certain degree (Butcher & Hope, 2006). Therefore, it is not a factor over which the candidates can have control. On the contrary, Aptitudes are something that can be changed. They influence the learning styles of the candidates which may or may not lead to the desirable student learning outcomes of the subject program. In the following section, there will be a diagnosis of the learning styles of the candidates.

In the focus group interviews, the internal and external examiners consented that the assessment task of the subject program included observation and measurement of the application of the candidates' ASK (Attitudes, Skills and Knowledge). They diagnosed whether the candidates were entitled to a MBA degree and hence the "career manager" certification. Therefore, the assessment task could be regarded as a "diagnostic assessment" as defined in Chapter III. The examiners were diagnosing the "fitness for purpose" as well as "the fitness of award" of the subject program and its candidates.

Brown et al. (1997) criticize that the function of measuring the candidates' ASK is often lost in the administrative machinery of the "assessment system". They believe that the

composites or aggregates of marks in a "formative cum summative" assessment may hinder the identification of a candidate's strengths and weaknesses. Thus, student learning outcomes could not be thoroughly assessed. In the focus group interviews, the examiners disproved this belief.

The examiners, particularly the external ones, believed that student learning outcomes could be diagnosed through the assessment of the candidates' research projects and their performance in the viva. Besides, the assessment task also threw light to the correlation of Student Learning Experiences and Student Learning Outcomes. In the viva, the examiners perceived that the candidates were "career managers learning from experience" (Morgan & O'neilly, 2002; Mumford, 1995; Sou, 2000). They also perceived four styles of the candidates' experiential learning, namely: intuitive, incidental, retrospective, and prospective (Mumford, 2008). These four learning styles were demonstrated by the candidates in their research projects and viva.

Intuitive learners learn from experience but not through a conscious process (Mumford, 2008). Candidates being intuitive learners reflected that learning was a consequence of having experiences. They could talk about a variety of different experiences in their action research projects. They were able to describe the happening and the outcome in detail. However, they rarely referred these things to the learning or developmental processes.

They saw managing and good business practices as synonymous with learning. Hence, they found it difficult to articulate how they learned to the Student Learning Experiences and what they earned to the Student Learning Outcomes. They were content that learning occurred as if through some natural process of osmosis. They trusted that learning was a natural and effortless process; therefore, they had a favorable student

learning experiences. They did not accept that it was advantageous to make the learning process more explicit, deliberate and conscious, even for themselves (Mumford, 2008).

Candidates being intuitive learners responded to the examiners during the viva in a unique manner. They typically used the following quotes (Figure 31):

Figure 31: Typical Quotes from the Candidates with Intuitive SLE

"I do that already without calling it a theory."
"I just complete an action research but I can't tell you how."
"I suspect that I am doing it all the time without realizing I'm doing so."
"I'm sure I'm learning all the time but I can't be more specific with the theories concerned."

(Source: Field-notes of the focus group interviews)

Incidental learners learn by chance from activities that jolt them into conducting a post-mortem (Mumford, 2008). When something extraordinary crops up or goes unplanned, they provoke incidental learners. Besides, mishaps and mistakes also provide incidental learners the jolts. Candidates being incidental learners tended to use the benefit of hindsight as a way of rationalizing or justifying what happened. During the viva, they found it easier to conduct the post-mortems by talking things over with the examiners. They typically used the following quotes (Figure 32):

Figure 32: Typical Quotes from the Candidates with Incidental SLE

"I only learn from my mistakes."
"If I know how to do something, I'm not going to learn from it."
"It's the originality of the experience that provokes more reflection."
"I learn from the unfamiliar parts of my job, not from the bits I am already familiar with and have already mastered."

(Source: Field-notes of focus group interviews)

Retrospective learners learn from experience by look back over what happed and reached conclusions about it (Mumford, 2008). Like incidental learners, the retrospective learners are provoked by mishaps and mistakes. Candidates being retrospective learners were inclined to draw lessons from routine events and successes. They learned from a diverse range of small and large, positive and negative, experiences.

Learning sequence of retrospective learners is: happenings – review – conclusions. By reviewing, the candidates acquired skills and knowledge or even changed their attitude. The examiners were happy to see that the candidates had their skills and knowledge confirmed and reinforced. They were happier to see the desirable student learning outcomes, that is, considered conclusions were knowingly reached.

Typical quotes from candidates with retrospective Student Learning Experiences were (Figure 33):

Figure 33: Typical Quotes from the Candidates with Retrospective SLE

"It helps to hold things up to the light."
"Reviewing is essential to put things into perspective."
"I concentrate on reviewing what happened after an experience."
"I never really understand something until I review it thoroughly."

(Source: Field-notes of focus group interviews)

Prospective learners demonstrate all the retrospective elements in their learning (Mumford, 2008). Additionally, they plan for the learning before an experience takes place. Prospective learners see future events as opportunities to learn. They tune their antennae into the possibility of learning from a whole variety of experiences (Mumford, 2008).

Learning sequence of prospective learners is: plan – learn – review – conclusions. Candidates being prospective learners expected to learn might not in fact materialize quite as planned. However, planning process enabled them to learn from the situation. Typical quotes from the candidates with Prospective Student Learning Experiences (Figure 34):

Figure 34: Typical Quotes from the Candidates with Prospective SLE

"I learn because I come here expecting to do so."
"Sorting out what I want to get in advance increases my chances of getting something worthwhile."
"There is no substitute for thorough planning, not only to complete the action research but also to learn from doing it."

(Source: Field-notes of focus group interviews)

The above description of the learners is drawn from the analysis of what the examiners said in the focus group interviews. The examiners, for example, did not say that a candidate was an intuitive learner or a prospective learner. The four titles are convenient but necessary shorthand developed and categorized for this research. Individual candidate might quote examples of their learning activities over the action research projects which could fit variously all four of the learning styles but might have a predominant learning style. Field-notes of the focus group interviews and analytic memos substantiated this conjecture.

Content analysis of the field-notes and analytic memos reports that the four learning styles have the following percentiles amongst the 331 candidates:

- ✧ 28% of intuitive learners
- ✧ 30% of incidental learners
- ✧ 29% of retrospective learners

✧ 13% of prospective learners

Apparently, some candidates were prone to demonstrate only one of the four styles. This research reveals that candidates (n=93) working in public sector prone to be intuitive learners; candidates (n=99) working in state-owned enterprises prone to be incidental learners; candidates (n=96) working in the private sector prone to be retrospective learners; and candidates (n=43) working in the transnational enterprises prone to be prospective learners. Field-notes of the focus group interviews and analytic memos also substantiated this conjecture. Content analysis of the sampled candidates' thesis informs the relationship of their background with the corresponding categorization of their learning styles.

Further study of these phenomena reveals that initial preference for one of the four styles needs not be a constraint for candidates' capacity to learn from experience in the future. Once the candidates have established the desirability of learning more effectively from work experiences, it is possible for them to extend their predominant learning style to other styles which they have not used before.

This proposition is established that favorable student learning experiences can be associated with the desirable student learning outcomes. If the candidates think harder about learning from the past and recent experiences and plan ahead for other learning experiences, they would not find the new learning style(s) unfamiliar that they should reject. Instead, the new learning style(s) may lead to more favorable student learning experiences which they have never been advised to use or helped to use before. Eventually, the new learning style(s) may help the candidates achieve the desirable student learning outcomes too.

With the guidance of the supervisors, the candidates in their real life spent time thinking about events that had occurred, why they had occurred and what they might do next time. However in the action research projects, the candidates might fail to show how to do this in relation to their learning. Similarly, the candidates plan ahead for their managerial work but they have rarely been guided on how to associate more effective learning with their managerial planning. Hence through coding processes as detailed in Chapter 4.3, some candidates tended to be intuitive (28%), incidental (30%) or retrospective learners (29%) rather than prospective learners (13%). Such categorization and conjecture was evidenced in my analytic memos.

Nonetheless, there were some positive cases discovered in this research. Some candidates had demonstrated favorable student learning experiences and desirable student learning outcomes. Further study of this phenomenon reveals that the candidates concerned are mainly prospective learners doing action research projects successfully. In the following paragraphs, the contributing factors to their learning success are explored.

Traditionally, learning success is measured by the quantity of programmed knowledge that a learner acquires (Koo, 1999a-b). If the candidates merely acquire knowledge, can they be described successful learners with desirable student learning outcomes? According to a primitive but popular learning equation (Beaty et al., 1997; Chan, 1994; Chan and Anderson, 1994; Gregory, 1994; Keys, 1994; Mumford, 1995; O'Neil, 1996), Learning (L) comprises two elements: Programmed Knowledge (P) and Questioning Insight (Q):

$$L = P + Q$$

To add value on the learning process and achieve desirable student learning outcomes, the element of Q should always be borne in mind (Sou, 2000). In other words, desirable student learning outcomes depends on an appropriate mix of the P and Q. In the context of career managers, Q is the need to resolve a managerial problem and so forth the identification of further management opportunity (Mumford, 1995). Action Learning is valuable with its emphasis and balance of both elements (Koo, 1999a-b; Mumford, 1995). Career managers would find action learning process so rewarding (Figure 35).

Figure 35: Ideal Quotes from the Candidates with Action Learning Experiences

I hear and I forget.
I see and I remember.
But I do and I understand.
What I have to learn to do, I learn by doing.

(Source: Field-notes of focus group interviews)

According to the examiners, the candidates during the viva often praised that their action research projects were rewarding learning experiences. The great majority of the candidates claimed that the action research projects help them solve some burning issues of their respective organizations. Some candidates said that the action research projects enabled them to explore new markets, introduce new products, expand the customer bases, etc. Some candidates even reported that their business turnovers were increased by millions of US dollars.

Career managers should experience that learning is not solely the acquisition of fresh programmed knowledge. Wisdom of the grand philosophers from the books may not necessarily be the solutions of the problems in the reality. Action learning involves testing out ideas and then modifying the respective ideas as a consequence (Bourner et al., 1996). A career manager should regard the subject program as a form of learning by doing that is, working on real problem focusing on learning and implementing solutions (Sandelands,

1998).

When career managers endeavor to make new ideas by placing themselves into natural experience, they seek to make meaning from their managerial experiences (Raelin, 1997). They learn more effectively in their own job settings whilst knowing best about work, at work and through work, within a structure that encourages learning (Peters, 1996). Action learning and action research provide a situation where prospective career managers at whatever level could become de facto 'in-house consultants' (Koo, 1999a-b) for their employers.

A prospective learner doing an action research project may be confident that the organization would benefit greatly from the findings of his or her research project that can stand the rigor of academic requirement as well as solve real problems therein. Furthermore, action learners in the supervisory or managerial capacity would definitely be benefited from their pragmatic job-related studies. They may lead to academic awards and consequently help them in future career development.

VII. DATA ANALYSIS – ASSESSMENT PRACTICES

7.1 Synopsis

The previous two chapters successively analyze the data about Student Learning Experiences and Student Learning Outcomes. On this analytical foundation, this chapter reports the processes and findings of data analysis regarding "Assessment Practices". To recap, "Assessment Practices" is a core category (Figure 13 and Figure 14) to be studied. Under which, there are one category called "Fitness for Purpose or Fitness of Award" and five sub-categories, namely (1) Objective/Criteria-Referenced Assessment, (2) Thesis Supervisors, (3) Registrars, (4) Internal Examiners, and (5) External Examiners. In the following sections, these five sub-categories will be analyzed one by one.

In implementing effective assessment practices, there appears to be a grave concern over the issues of fairness, reliability and validity (Brown et al., 1997). In the review of "Technical Literature", these issues can be construed as follows (Figure 36):

Figure 36: Three Issues in Assessment Practices

Fairness	Equality of opportunity and treatment.
Reliability	Consistency of approach.
Validity	Appropriateness of methods of truth-seeking.

(Source: Brown et al., 1997)

"Fairness", "Reliability" and "Validity" of assessment practices are deep problems. Atkins et al. (1993) criticize the procedures of many universities for student assessment. Barnett (1994) further criticizes the conflicts of purpose, variability across and within subject assessments, the sample of activities assessed, the methods of assessment and the grading system of worth in the English test system. With such criticism, is it possible to

implement assessment practices in an effective way? Furthermore, is it possible to seek the truth of an educational system as questioned by Rowntree (1987, p1)?

> *If we wish to discover the truth about an educational system, we must look into its assessment procedures. What student qualities and achievements are actively valued and rewarded by the system? How are its purposes and intentions realized?*

In this research, review of previous studies in Chapter V revealed that the candidates had relatively less favorable student learning experiences in the viva. In the focus group interviews, the registrars reported that the candidates had actually commented on the "fairness" and the "reliability" of the viva. Occasionally, some candidates came up to the registrar and queried about the unfavorable results in the viva or solicited her assistance in re-submission of the theses for grading. They perceived variability across the academic panels and the thesis supervisors, the methods of assessment and the grading system of worth. When the candidates from the same institution were assigned to different academic panels for the viva, they might get different assessments from the examiners. Obviously, this was a "reliability" issue. However, it might also be the cause of the "unfairness" feeling of the candidates.

This phenomenon was thoroughly investigated in this research. The investigation started with an exploration of relevant literature in terms of "fairness", "reliability" and "validity" of effective assessment practices. To this end, the nature of "reliability" and "validity" and their implications to assessment practices will be examined. Unlike Chapter V, statistical techniques will be used as they are based essentially on measures of agreements and differences and range in complexity from correlations and analyses of variance to factor analyses, multi-variate analysis and beyond (Ebel & Frisbie, 1986; Grounlund, 1988). Instead, the focus of study will be laid on the underlying concepts that are crucial to "fair" and effective assessment practices.

The standard approaches to reliability and validity are derived from psychometrics. They are based on the notion of an ideal which can be achieved if only one can reduce the errors. In management education, there is a range of values involved at the higher levels of abilities, skills and knowledge of which their integration could hardly assume that there is only one ideal. That is to say … management problems can be solved by the concerted efforts of a manager or his/her team(s) by collectively applying their abilities, skills and knowledge. Different manager and combination of team(s) may come up with different solutions. Amongst all those possible solutions, there may be more than one ideal solution that a manager can choose.

Hence, non-statistical approaches such as the use of judgments in the sampling of research project, definition of research problems, identification of the burning issues, application of research methods and proposition of remedial action are required in the assessment practices of management education. Nonetheless, non-statistical judgmental approaches of student assessment are also based on the underlying concepts of reliability and validity – notions of precision and accuracy. (Atkins et al., 1993; Barnett, 1994).

Brown et al. (1997, p. 234) tell an analogical story about "watch and time". The mechanism of a watch may be precise (reliable), it may measure the minutes and hours consistently, but the time show may be wrong. The time shown on a particular occasion may be correct (valid) but the watch may have stopped or its variable rate of loss or gain rather than its consistency may have provided the result. Furthermore, the watch-keeper may also be a variable when he or she fails to read and interpret the watch dial correctly.

The above analogy can be applied to interpreting the results of the assessment practices in the subject program. Even if guidelines and assessment forms are provided, ultimately the assessment instrument is the examiner in conjunction with the particular guidelines,

assessment forms and procedures. In analyzing data about the assessment practices, the focus can be placed upon the fairness, reliability and validity of the assessment task and its actual nature. Considering "fitness for purpose" and "fitness of award", the thesis supervisors, the registrars, the internal and external examiners should be able to tell whether the assessment practices of the subject program are effective or not.

7.2 Assessment Practices

With reference to the relevant literature, taking a sample of what the candidates do, making inferences and estimating the worth of their actions (Brown et al., 1997) may be regarded as effective assessment practices. Obviously, assessment practices of the subject program equates to:

Sampling of Candidates + Making Inferences + Estimating Worth of Actions

First, sampling (choice of a research topic) is undertaken by the candidates themselves and/or their thesis supervisors. Sometimes, the candidates may choose a research topic that is incidentally assigned by their employers to be a company project. Anyway, this is a sample of what the candidate do. It involves the learning tasks: selecting a research topic, identifying the burning issues, conducting an action research, reporting the findings, proposing the solutions, and writing a thesis.

After taking a sample of research project, the examiners could make inferences from the research thesis and through the viva about the student learning outcomes such as attitudes, skills. knowledge, achievements, potential, intelligence, aptitudes, motivations, and personality (Brown, et al., 1997). With the inferences, the examiners estimate the worth of the candidates' actions. The estimation is in the form of grades, marks, recommendations.

These three aspects of assessment task as learning task have respective shortcomings which lead to problematic assessment practices. **Sampling** (choice of a research project) may not be representative of the candidate's capabilities or may not match the learning objectives of the subject program. It may be drawn on too narrow a domain, such as one of the core subjects. Besides, it may be over-weighted towards particular skills or methods instead of applying integrated skills and knowledge gained from the subject program. (Brown, et al., 1997)

The **Inferences** drawn about the candidate's research may vary widely from examiner to examiner. The variations may be more significant when explicit criteria or marking schemes are not used. The **Estimation of Worth** in terms of the marks or grades may also vary. In the subject program, the variation in grading and marking would even lead to unjustifiable decisions by the academic panels.

Student assessment can be based upon the procedure of "Sampling", "Making References" and "Estimating Worth". On this basis, assessment practices would carry some common weaknesses (Brown, et al., 1997, pp. 251-52):

- ✧ The sample does not match the stated outcomes.
- ✧ The sample is drawn from too narrow a domain.
- ✧ The sample is too large or too small.
- ✧ Absence of well-defined criteria.
- ✧ Unduly specific criteria.
- ✧ Variations in the inferences drawn by different assessors of the sample.
- ✧ Variations in estimates of worth.

Analysis of the sampled candidates' theses and assessment forms disproves some of the above weaknesses. During the focus group interviews, the registrars, internal and external examiners concurred that an action research project was an appropriate sample with well-defined and specific criteria. The candidates were required to draw a live problem from the domain of business administration and applied their knowledge and skills gained from the subject program to solve the problem. The candidates proposed solutions in their theses from which the examiners drew inferences how the candidates applied their skills and knowledge gained from the subject program.

In estimating the worth of the proposed solutions, internal and external examiners were supposed to adopt the well-defined and specific criteria that were prescribed by AIOU. The internal and external examiners reflected that they were committed to a fair, reliable and valid assessment. However, some of them noticed that there were variations in the inferences drawn by different examiners as well as different panels. Occasionally, the internal examiners and the external examiners could not compromise in the final marks of the candidates. Some opined that the existing assessment practices were most effective while some said that there was room for improvement.

7.2.1 Shortcomings and Benefits of Assessment Practices

Effective assessment practices can be beneficial to the educational quality assurance. However, assessment practices established on sampling, inferences and estimation appear to be problematic. Brown et al. (1997) list some shortcomings of assessment practices which could be compared with the findings of this research (Figure 37):

Figure 37: Shortcomings of Assessment Practices vis-à-vis Research Findings

Item	Shortcomings of Assessment Practices	Findings of this Research
1	Overload of the candidates and the examiners.	Registrars reported that the candidates had to attend a pre-viva in mainland China and a viva in Macau. Internal and External Examiners said that the assessments of theses were tedious.
2	Insufficient time for the candidates to complete the action research in the time available.	Some candidates reflected that they wished to do a big research project but time did not allow it.
3	Insufficient time for the examiners to mark the theses before the viva.	Some internal and external examiners reported that marking of a single thesis might take an hour before the viva. Some thesis might take longer time to digest.
4	Inadequate or superficial feedback provided to the candidates.	Some internal and external examiners responded that they wished to provide more feedback to the candidates during the half-an-hour viva. They assessed from the candidates' performance during the viva that the feedback from the thesis supervisors to a candidate during the research appeared to be inadequate and superficial.
5	Wide variations in assessment demands of different panels.	Internal and external examiners noted that different combination of the panel members varied in their assessment demands.
6	Wide variations in marking across panels.	Internal and external examiners noted that there were wide variations across different panels.
7	Wide variations in marking within a panel.	Some internal and external examiners noted that there were variations about 10 scores in their markings before the viva. However, they could promise with each other during the viva.
8	Wide variations in marking by supervisors.	Internal and external examiners commented that there were variations in marking by supervisors. Generally, their markings were higher than that by the panels.
9	Fuzzy or non-existent criteria.	Internal and external examiners believed that criteria existed but they were subject to their own interpretation.
10	Undue precision and specificity of marking	Internal and external examiners commented that precise and specific marking schemes or criteria might not be

Item	Shortcomings of Assessment Practices	Findings of this Research
	schemes or criteria.	viable for the action research in the management education.
11	Candidates do not know what is expected of them.	Internal and external examiners commented that some candidates did not know the exact requirements of the action research as well as the thesis.
12	Candidates do not know what counts as a good or bad research or thesis.	Internal and external examiners commented that some candidates had not realized the strengths and weaknesses until they were enlightened during the viva.

(Source: Brown et al., 1997; Analytic memos of the subject research)

With an array of shortcomings, it is questionable whether the assessment practices could assure educational quality. Why should have learning assessment tasks and who would be benefited from them? It is believed that assessment supports learning and assessment task should be regarded as a learning task. This belief is supported by a number of scholars over the years. Hattie and Watkins (1985) comment that using projects and open-ended assessments tend to promote independence and deeper strategies of learning.

Though using problem-based approaches and appropriate research projects tends to promote deeper styles of learning, deeper approaches to study and independent learning tend to decline in higher education (Bain & Thomas, 1984; Biggs, 1987, 1996 & 1999; Clarke & Newble, 1987; Eysenck, et al., 1987; Harper & Kember, 1989; Blake & Vernon, 1994). The declining phenomenon may be attributed to the students' perspectives on assessment tasks. Students tend to reject deeper approaches to study since the assessment involves a great deal of reproductive learning and they reckon that deeper approaches are not important or worth learning (Ramsden, 1998 & 1992; Entwistle, 1987 & 1992).

Perhaps, the students may not realize the benefit of the assessment task as learning task. When assessment tasks are correlated to learning tasks, the assessment should be at the hearts of those who: (a) learn, i.e., the student, (b) teach, i.e., the trainers, (c) hire, i.e., the

employer, (d) develop the course or training program, i.e., the institution, and (e) accredit the course or training program, that is, the authority. All these people could benefit from assessment. Brown et al. (1997) provides a list of the benefits of effective assessment practices and their beneficiaries (Figure 38):

During the focus group interviews, the examiners and registrars of AIOU reached a consensus on the benefits of the assessment practices. Particularly, the examiners opined that 30 minutes' viva were inadequate for them to maximize the benefits. The registrars observed that some examiners habitually overran the viva by providing feedback to the candidates to improve their learning. In such circumstances, the viva would then be a formative rather than summative assessment.

If the examiners regard the viva as a counseling session, motivating the candidates, diagnosing their strengths and weaknesses, helping them to develop their skills of self-assessment, and even providing a profile of what a candidate has learnt, one can imagine how long a viva would take. Though the examiners' opinions have their grounds but the assessment task in the form of a viva could hardly afford it.

Facing the time constraint of a viva voce, the examiners and registrars expect that the counseling and feedback should take place during the research process instead of the appraisal process. Hence, AIOU expects that a thesis supervisor counsels and assesses the candidate from the commencement of the action research to the candidate's attendance at the viva. The assessment of student learning outcomes task untaken by the thesis supervisor could be a criteria-referenced assessment, a norm-referenced assessment or an objective-referenced assessment.

Figure 38: Benefits of Effective Assessment Practice

Item	Benefit	To Student	To Teacher	To Employer	To Institution	To Authority
1	Providing feedback to the candidates to improve their learning.	✓	✓		✓	
2	Motivating the candidates.	✓	✓		✓	
3	Diagnosing a candidate's strengths and weaknesses.	✓	✓		✓	
4	Helping candidates to develop their skills of self-assessment.	✓	✓		✓	
5	Providing a profile of what a candidate has learnt.		✓		✓	
6	Passing or Failing a candidate.		✓		✓	
7	Grading or Ranking a candidate.		✓		✓	
8	Licensing the candidates to proceed.	✓				
9	Licensing the candidates to practice.	✓		✓		
10	Selecting candidates for future training programs.				✓	
11	Predicting the candidates' success in employment.			✓	✓	
12	Selecting the candidates for future employment.			✓		
13	Providing feedback to the trainers.		✓		✓	
14	Improving teaching.		✓		✓	
15	Evaluating the strengths and weaknesses of a training program.				✓	✓
16	Making a training program appear "respectable" and creditworthy to other institutions and employers.			✓	✓	✓

The assessment task of the program can be termed a criteria-referenced assessment rather than a norm-referenced assessment. The candidates are asked to identify a real-life problem in their workplaces and conduct an action research project and then compile a thesis which would be subject to a "pass-or-fail" rating system. When the overall marks earned by each candidate would not yield a rank order, they are not subject to a norm-referenced assessment in the viva because a candidate's performance does not tend to be relative to the performance of the other candidates.

However in the focus group interviews, the registrars reported that the thesis supervisors in mainland China had a deep belief on norm-referenced assessment. They apparently believed that those candidates going to the viva should be placed on the top echelon. When the candidates were coming from the same institution but under the guidance of different supervisors, the thesis supervisors tended to be leniently in rating the candidates under their own supervision. Understandably, the thesis supervisors guided the candidates to complete the research and the theses. They naturally rated their guided work as the most promising theses in the class.

From the perspectives of the internal and external examiners, they regarded the assessment task as an objectives-referenced assessment. During the focus group interviews, the internal and external examiners reflected that the assessment task could tell whether the objectives of the subject program had been met. On this belief, they rated the candidates regardless to the overall distribution of scores. However, a review of the assessment forms surfaced that different panels of internal and external examiners had significant difference in rating a group of the candidates from the same class and same institution. This "Criteria-Referenced Assessment at Variance with Norm-Referenced Assessment" phenomenon will be discussed later in this chapter.

Apart from the difference in rating amongst the thesis supervisors, the internal examiners and the external examiners, the assessment forms also suggested a phenomenon under objectives-referenced assessment as defined in Chapter III. Regarding the assessment task as an objectives-referenced assessment, the candidates and the examiners of the subject program appeared to have conflicting views on the objectives of the subject program. The candidates' views on "Student Learning Experiences" versus the objectives of the subject program conflicting with the examiners' views on "Student Learning Outcomes" versus the objectives of the subject program. This phenomenon coincided with the notions of "fitness for purpose" and "fitness of award" that were reviewed in the relevant literature from Chapter I.

The links between these two notions and their inherent conflict in "educational quality assurance" were highlighted in Chapter III. The former notion examined the links between particular student learning experiences and specific objectives of the subject program. In contrast, the latter notion concerned the links between the student learning outcomes against the national certification standards of Career Manager through subject program.

Analysis of "Student Learning Experiences" in Chapter V revealed that the candidates had good learning experiences and believed that the objectives of the subject program were met upon their graduation. On the other hand, analysis of "Student Learning Outcomes" in Chapter VI disclosed that the examiners had reservation about the abilities, skills and knowledge (ASK) of the candidates upon their graduation. Though the candidates successfully earned an MBA degree and gained the occupational title of Career Manager, there were still room for improvement in their abilities, skills and knowledge (ASK) of being a Career Manager. Particularly, the external examiners concluded that the candidates' ASK should be enhanced for the sake of their management career.

7.2.2 Student Learning Experiences versus Student Learning Outcomes

This "good Student Learning Experiences versus maybe better Student Learning Outcomes" phenomenon in an objectives-referenced assessment replicated the subsistence of the Aristotelian[53] and Platonic[54] philosophical interpretation of "good" (Brown & Knight, 1994; Sou, 2008). Aristotelian notion established that the subject program was "good for the candidates and for their career development". "Good" was construed on the favorable "Student Learning Experiences" of the candidates and their interpretation of "Fitness for Purpose" along with their perceived objectives of the subject program. The candidates regarded the subject program as a "good" MBA degree program because it was contributory to their career development.

On the other hand, the internal and external examiners believed that "fitness of award" was independent of the candidates. Platonic notion established "maybe better Student Learning Outcomes" should be construed on the basis of the pre-determined standards of the society (Brown & Knight, 1994; Sou, 2008). In the context of this research, the national requirements for Chinese Career Manager Certification may be regarded as the pre-determined standards of the society. The internal and external examiners believed that there were ideal standards to which Career Managers should aspire.

To this end, it is established that criteria-referenced assessment (CRA) made by the examiners may be jeopardized by the norm-referenced assessment (NRA) made by the

---

[53] Aristotelian notion is established on the premises of "good for what and for whom". "Good" is construed on the basis of the students' reflections upon experience. The reflections lead to differing interpretation of "good" according to the purposes, context and the shifting boundaries of experience. "Assessment of educational quality" in terms of "fitness for purpose" is conceptualized by this belief. (Brown & Knight, 1994; Sou, 2008)

[54] Platonic notion is established in the principles that are pre-determined by the "guardians" of the society. "Good" is construed on the basis of the pre-determined standards of the society. Plato believes that there are ideal standards to which all human beings should aspire. National standards of a degree program and rank ordering of an educational institution is conceptualized by this belief. A Platonic model of quality assessment calls for a measure against "ideal" standards that are independent of the curriculum of the educational program, processes of the teaching and learning, and abilities of the students. (Brown & Knight, 1994; Sou, 2008)

thesis supervisors. It is impossible to have both norm- and criterion-referenced assessments at the same time. Instead, objectives-referenced assessment could be another form of criterion-referenced assessment with the criteria being derived from the objectives that are coupled with "fitness for purposes" as well as "fitness of award". Furthermore, candidates' perception of "fitness for purpose" may contradict with the examiners' perception of "fitness of award" under the Aristotelian and Platonic notions (Brown & Knight, 1994; Sou, 2008).

In the preceding paragraph, a "CRA at Variance with NRA" phenomenon emerged. To understand more about this phenomenon concerning "reliability", I randomly drew 5 from 23 panels for in-depth inquiries. These 5 panels examined 72 candidates during the sampling period: 15 candidates by Panel 1, 14 candidates by Panel 2, 15 candidates by Panel 3, 14 candidates by Panel 4, and 14 candidates by Panel 5 (Figure 39).

Figure 39: Marking Variances amongst Thesis Supervisors, Internal and External Examiners of 5 Panels

Panel	Thesis Supervisor	Internal Examiner	External Examiner	Final Mark
1	75	67	69	**70**
	75	76	70	**73**
	73	66	69	**68**
	75	73	69	**72**
	75	75	71	**80**
	88	70	66	**71**
	75	67	66	**68**
	76	68	69	**68**
	78	64	68	**66**
	76	66	64	**65**
	75	73	68	**74**
	87	74	70	**75**
	75	73	70	**71**
	90	65	65	**65**
	92	70	73	**74**
2	72	72	65	**68**
	76	73	71	**72**
	74	66	70	**67**
	80	68	70	**72**

Panel	Thesis Supervisor	Internal Examiner	External Examiner	Final Mark
	76	65	68	**0**
	87	70	70	**70**
	80	74	79	**68**
	80	65	69	**70**
	75	71	68	**74**
	81	69	73	**72**
	83	60	66	**68**
	72	69	70	**74**
	68	68	64	**68**
	75	72	74	**74**
3	74	67	67	**67**
	81	64	69	**66**
	79	67	69	**66**
	76	64	57	**67**
	77	66	69	**70**
	82	70	70	**72**
	91	65	77	**71**
	78	65	67	**62**
	79	72	70	**69**
	75	63	68	**65**
	78	63	67	**68**
	73	72	69	**71**
	80	73	69	**67**
	83	68	74	**71**
	85	73	72	**60**
4	80	64	64	**67**
	78	66	70	**70**
	79	64	72	**72**
	75	65	70	**67**
	75	67	72	**63**
	78	65	72	**62**
	84	72	70	**75**
	73	68	69	**76**
	79	68	70	**73**
	85	64	68	**64**
	81	63	67	**66**
	77	62	59	**55**
	73	70	62	**74**
	87	65	69	**76**
5	77	72	77	**72**
	80	73	76	**73**
	88	64	71	**70**
	79	65	71	**65**
	78	68	70	**72**
	81	72	70	**73**
	80	75	79	**74**

Panel	Thesis Supervisor	Internal Examiner	External Examiner	Final Mark
	85	64	67	**60**
	78	64	70	**70**
	79	67	70	**50**
	80	75	68	**50**
	82	64	59	**50**
	80	68	69	**70**
	82	75	76	**70**

(Source: Assessment forms of the sample candidates)

Note: *The correlation of the above table is (a) marks given by a thesis supervisor before the viva, (b) marks given individually by the internal and external supervisors before the viva, and (c) the final marks given collectively by the academic panel comprising the internal and external examiners.*

Amongst the 72 candidates, there are five failing cases: one in Panel 2, one in Panel 4, and three in Panel 5. They are excluded from the statistical analysis for homogeneity of variances. However, those failing cases (Figure 40) are studied successively in the following paragraphs.

Figure 40: Five Failing Cases of Viva Voce in Panel 2, 4 and 5

Failing Case	Thesis Supervisor	Internal Examiner	External Examiner	Final Score
1 (Panel 2)	76	65	68	**0**
2 (Panel 4)	77	62	59	**55**
3 (Panel 5)	79	67	70	**50**
4 (Panel 5)	80	75	68	**50**
5 (Panel 5)	82	64	59	**50**

(Source: Assessment forms of 5 sampled candidates)

In the first failing case, the panel discovered that the candidate only worked in the enterprise under study for one year as a supervisor. She was unable to answer the queries of the internal and external examiners about the research on "brand management". Eventually, she admitted that one of her friends wrote the thesis for her. In the second failing case, the candidate was the director of public service organization. During the viva,

she failed to prove that she had grasped the basic knowledge about "total learning organization" under study.

In the third failing case, the candidate conducted a questionnaire survey that had a strong theoretical foundation such as schools of thought of Maslow and Hertzberg. However, he could not answer the examiners' queries about these theories in a satisfactory manner. In the fourth failing case, the candidate unreasonably used some outdated data from 2002-04 for SWOT (Strength, Weakness, Opportunity and Threat) analysis. Furthermore, she could not reason the use of BCG (Boston Consulting Group) Matrix in her study. (Aakar, 1998; Huang, 2004)

In the fifth failing case, the candidate had six years' experience in the securities market. She submitted a thesis looked like a feasibility study on a new product in the securities market. The panel ruled that the theoretical base of the research and the application of knowledge and skills gained from the subject program to the real life problem appeared to be inadequate. In five failing cases, the external examiners detected two of them before the viva while the panels collectively detected another three during the viva.

7.2.3 Fairness, Reliability and Validity of Assessment

The external examiners and academic panels appear to be essential instruments of fairness, reliability and validity. The presence of the external examiners and formation of panels were to protect the candidates and to safeguard standards. Protecting students implies checking on the fairness, reliability and validity of the assessment practice of the subject program. Safeguarding standards involves checking on the design of the assessment task, monitoring the student learning outcomes and making an estimate of worth of the subject program.

With the participation of external examiners, the collective assessment of the panel has particular advantages in upholding the reliability of the assessment practice of the subject program. McCormick (1979) relates reliability to the degree of relationship between or among the assessments of two or more independent assessors. For a sample of jobs, reliability is often measured by correlating pairs of independent assessments.

Previous studies (Scott, 1963) ascertained that the combination of the assessments of several people tended to increase the reliability of the composite assessments as long as all of them were good assessors. Reliability in this context is referred to the degree of relationship between or among the assessments of two or more independent assessors, or between separate assessments at different times by the same assessor. Reliability is measured by correlating pairs of independent assessment for a sample job, or by correlating separate (test-retest) assessments made by the same assessor at different times. Assuming that the average test-retest reliability of an assessor is 0.80, the reliability coefficients for various numbers of assessors (Figure 41) could be:

Figure 41: Reliability Coefficients for Various Numbers of Assessors

Sample Size	**1**	**2**	**4**	**6**	**8**	**16**	**20**
Reliability	.80	.89	.94	.96	.97	.98	.99

(Source: Scott, 1963)

The pooled reliability of assessments tends to increase appreciably with even three or four assessors, and then increases more gradually (McCormick, 1979). Incidentally, there are hints that the pooled reliability of assessments made independently by the several assessors tends to be a bit higher than the reliability of assessments made collectively by panels of three to five assessors (Hoggatt and Hazel, 1970). These hints suggest that it is preferable to obtain individual assessor from two or more assessors and average them, rather than obtaining group assessments by consensus.

Panels may be advantageous in one way – independent assessment other than collective assessment. It is observed that the thesis supervisors, internal and external examiners are allowed to make individual assessment before the viva. Then, a panel comprising internal and external examiners makes the collective assessment after the viva. Such arrangement can be regarded as a good assessment practice of panel system.

However, it is noteworthy that there are variances between the markings of thesis supervisor, internal and external examiners on the 67 sampled candidates. If the final marks awarded by the academic panels can be regarded as a "Control Group", what are the variances between three individual assessments and the joint assessments made by the internal and external examiners in the viva?

7.2.4 Marking Variances in Assessment

Using Statistical Product and Service Solutions (SPSS) 15.0 for Windows, the Contrast Coefficients of Assessments are worked out as per Figure 42: (1) Marks given by Thesis Supervisors, (2) Marks given by Internal Examiners, and (3) Marks given by External Examiners while (4) Final Marks awarded in the Viva by the Panels are regarded as a "Control Group". There are significant variances between the three individual assessments (1), (2) and (3) and the joint assessments (4) made by the internal and external examiners in the viva.

Figure 42: Contrast Coefficients of Assessments

	Assessment			
Contrast	1	2	3	4
1	1	0	0	-1
2	0	1	0	-1
3	0	0	1	-1

One way Analysis of Variance (ANOVA) test results are: F=94.918, P=0.000<0.05. Statistically, there are significant variances in the assessments (Figure 43) between and within Group (1), Group (2), Group (3) and the Control Group (4).

Figure 43: One Way ANOVA Test Results of the Marks given in 4 Assessments

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4952.966	3	1650.989	94.918	.000
Within Groups	4591.970	264	17.394		
Total	9544.937	267			

Test of Homogeneity of Variances (Figure 44) is conducted. The Levene statistics show that P=0.023<0.05 does not assume equal variances. Therefore, further reference should be made to the "Does not assume equal variances" section of the Dunnett's tD Test.

Figure 44: Test of Homogeneity of Variances in Assessments

Levene Statistic	df1	df2	Sig.
3.227	3	264	.023

Findings of Dunnett's tD Test (Figure 45) indicate that the t value of Contrast 1 is 12.028 whereas P=0.000<0.05. Statistically, there are significant variances between the assessments made by (1) Thesis Supervisors and (4) Academic Panels. Contrast 2 with t value of -1.892 and P=0.061>0.05 indicates that there are insignificant statistical variances between the assessments made by (2) Internal Examiners and (4) Academic Panels. Contrast 3 with t-value of -0.045 and P=0.064>0.05 also indicates that there are insignificant statistical variances between the assessments made by (3) External Examiners and (4) Academic Panels.

Figure 45: Contrast Tests of Assessments by Thesis Supervisors, Internal and External Examiners

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Marks	Assume equal variances	1	9.42	.721	13.070	264	.000
		2	-1.28	.721	-1.781	264	.076
		3	-.03	.721	-.041	264	.967
	Does not assume equal	1	9.42	.783	12.028	126.014	.000
	variances	2	-1.28	.678	-1.892	131.766	.061
		3	-.03	.667	-.045	131.199	.964

The findings of above priori comparisons suggest that there are significant variances between the assessments made by the thesis supervisors and the assessments made by the academic panels comprising internal and external examiners. Statistically, the assessments made individually or jointly by internal examiners and external examiners do not show significant variances. It will be clearer to study the Honestly Significant Difference (HSD) of these assessments by using Newman-Kuels Studenized Range Test (S-N-K) in the following paragraphs.

Findings of multiple comparisons (Figure 46) show that the mean difference of assessments made by the thesis supervisors at the 0.05 level when compared with the assessments made individually or jointly by the internal and external examiners.

Figure 46: Multiple Comparisons of Assessments

Multiple Comparisons

Dependent Variable: Marks

	(I) Assessment	(J) Assessment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	10.701*	.721	.000	8.84	12.56
		3	9.448*	.721	.000	7.58	11.31
		4	9.418*	.721	.000	7.55	11.28
	2	1	-10.701*	.721	.000	-12.56	-8.84
		3	-1.254	.721	.305	-3.12	.61
		4	-1.284	.721	.285	-3.15	.58
	3	1	-9.448*	.721	.000	-11.31	-7.58
		2	1.254	.721	.305	-.61	3.12
		4	-.030	.721	1.000	-1.89	1.83
	4	1	-9.418*	.721	.000	-11.28	-7.55
		2	1.284	.721	.285	-.58	3.15
		3	.030	.721	1.000	-1.83	1.89

*. The mean difference is significant at the .05 level.

Findings of S-N-K Test (Figure 47) show the difference between the assessments made individually or jointly by the internal and external examiners and prematurely by the thesis supervisors. All the samples (67 candidates) show significant difference. Tukey HSD Test comes up with identical findings.

Figure 47: Post Hoc Multiple Comparisons of Assessments

	Assessment	N	Subset for alpha = .05	
			1	2
Student-Newman-Kuels[a]	2	67	68.31	
	3	67	69.57	
	4	67	69.60	
	1	67		79.01
	Sig.		.178	1.000
Tukey HSD[a]	2	67	68.31	
	3	67	69.57	
	4	67	69.60	
	1	67		79.01
	Sig.		.285	1.000

Means for groups in homogeneous subsets are displayed.

[a] Uses Harmonic Mean Sample Size = 67.000.

(Source: The Subject Research)

Brown et al., (1997) suggest that the two main measures of reliability in assessment are measure of marking differences between examiners and within examiners. Historically, there has been plenty of evidence on the marking differences between examiners, even when using marking schemes. Figure 48 shows findings of pervious research on marking differences:

Figure 48: Research and Findings on Marking Differences between Examiners from 1890 to 1994

Researcher	Previous Research	Findings
Edgeworth (1890)	Twenty-eight qualified examiners were invited to mark a Latin prose as if it were by candidates for the Indian Civil Service.	Marks ranged from 45 to 100 while the modal mark was 75.
Hartog and Rhodes (1935 & 1936)	Marking of English, history and chemistry papers in a school certificate examination.	Different examiners marked the same candidates as failed, passed or passed with credits.
Diederich (1957), Bell (1980), Newstead and Dennis (1994)	Fifty-three experts were invited to mark 300 short essays of year one university students.	All essays received five or more grades of the nine possible while 34 percent of the essays obtained all the grades.

(Source: Literature review of the subject research)

In the subject program, assessments by different examiners produce considerable marking differences. Their statistical variances are summarized in Figure 49 for cross reference:

Figure 49: Marking Difference of Thesis Supervisors, Internal/External Examiners and Academic Panels

Contrast with	Thesis Supervisors	Internal Examiners	External Examiners
Academic Panels comprising Internal and External Examiners	Significant Variances	Insignificant Variances	Insignificant Variances

(Source: Analytic memos of the subject research)

Such differences influenced the fairness, reliability and validity of the assessment practice. In determining the appropriate marks, the part played by the supervisors or the examiners can be greater than that of the student's performance. However as compared with the previous research on marking differences, the variances in marking amongst the thesis supervisors, internal and external examiners appears to be nominal. Nonetheless,

"variances in marking" is one of the areas to be improved in the assessment practice of the subject program.

7.2.5 Rating Errors in Assessment

With reference to the field-notes and analytic memos, the "variances in marking" phenomenon might be attributed to the common "Rating Errors" (Fisher et al., 1991; Fleenor & Scontrino, 1982; Median, 1981; Sou, 2000; Swan, 1991). The most common "Rating Errors" are listed in Figure 50 for cross-reference. Prevailing rating errors appeared to be Item 1, 2, 3, 7, 9, and 11.

Figure 50: Common Rating Errors

Item	Common Rating Errors	Descriptions
1	Contrast Effect	Rating the candidate relative to the others instead of basing the rating on the candidate's performance expectations.
2	Central Tendency	Rating towards the middle of the scale to avoid both high and low ratings in all aspects of performance. A tendency to give average marks to avoid justification and queries.
3	Halo Effect	Rating the candidate whose performance is outstandingly high in one aspect, high in all aspects. In other words, a tendency for good performance in one area of work to 'color' the assessment in other areas of work.
4	Horns Effect	Rating the candidate whose performance is very low in one aspect, low in all aspects.
5	Incompatible Ratings	Ratings are not tally with the "Remarks"; "Overall Grading" is not tally with individual ratings and 'Remarks'.
6	Negative Leniency	Rating the candidate inappropriately low in a number of aspects.
7	Over-Rating or Leniency	A tendency to appraise more leniently than is warranted, for instance, reward candidate for effort; avoid confrontation about weaknesses; do not wish to disadvantage own students, …
8	Personal Bias	Judgment may be influenced by own need or value over

Item	Common Rating Errors	Descriptions
		appearance and attitude.
9	Positive Leniency	Rating the candidate inappropriately high in a number of aspects.
10	Recency Effect	Recent events are over-emphasized.
11	Reputation Effect	Rating the candidate based on general reputation and previous performance rather than on performance for the research project.
12	Self-Values Effect	Rating the candidate higher or lower based on the degree to which he or she has attitudes, beliefs, interests and values similar to the thesis supervisor, internal examiner or external examiner.
13	Seniority Manipulation	A tendency to give low grades to greenhorn and higher grades to candidates of seniority in recognition of experience. This can also guarantee continuous improvement of performance of a candidate.
14	Stereotyping Effect	Rating the candidate based on one's general perception of a group (for example, 'older employees') instead of as an individual.
15	Unsubstantiated / Poorly Substantiated	Ratings are not substantiated or are poorly substantiated by evidence or incidents. This makes the marking less credible.

(Source: Fisher et al., 1991; Fleenor & Scontrino, 1982; Median, 1981; Sou, 2000; Swan, 1991)

To name a few but not all, rating errors may also be Criteria or Reference System, External Impression, Flattery Effect, Hierarchy Effect, Interaction Effect, Pre-Judgment and Projection (Bartz et al., 1993). There are so many possible errors in rating one's performance. Indeed, attention should be paid to anything such as Rating Errors that may go wrong in a robust Assessment Practices. The thesis supervisors as well as the internal or external examiners must guard against common rating errors. However with a better understanding about these rating errors, the thesis supervisors, the internal examiners and the external examiners may develop their own skills to make the appraisal processes error-free within the assessment practices of the subject program.

VIII. CONCLUSION

8.1 Consolidation of Data Analysis

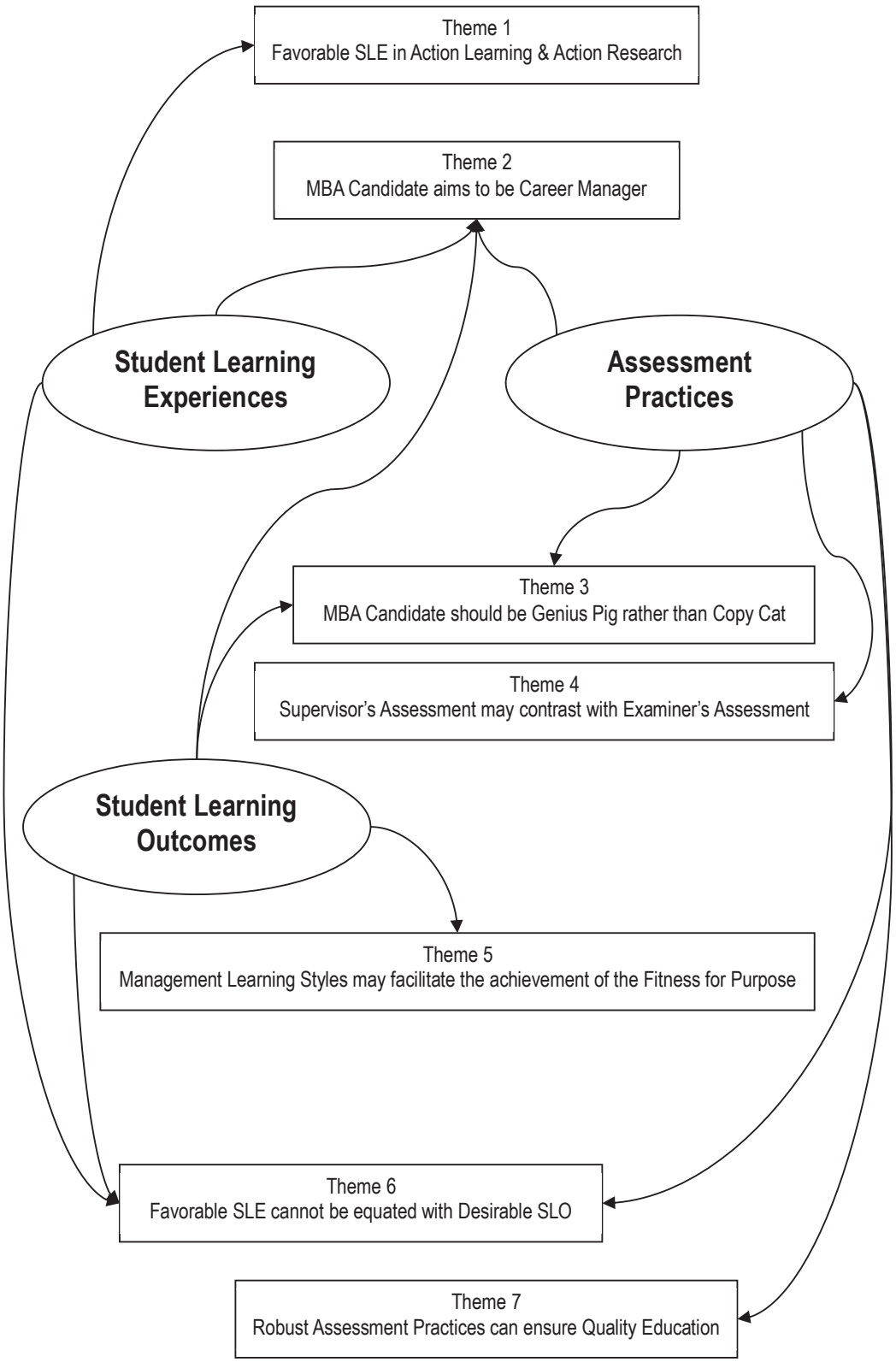
The three chapters of "Data Analysis" (Chapter V, VI and VII) are not merely descriptive findings. I aim to transcend from descriptive to analytic. Established themes are justified by qualitative data and/or supported by quantitative data. Besides, I already presented the findings of this research to 88 candidates of the subject programs in two discussion forums 2009-10. The candidates gave positive feedback which refined the themes to be reported in this Chapter. In 2011, I also presented the findings of this research in the International Conference of Learning Society and Open Education Research and got the conference proceedings (Sou, 2011) published by the conference organizer. The international scholars' constructive comments were used to refine this Chapter as well.

As planned, data saturation through coding has led to thematic emergence (Figure 10 in Chapter IV). In this chapter, readers will be guided to understand the themes emerged. The themes can answer the research problems framed in the first chapter. At the end of this chapter, answers to the two research problems will be provided. Before that, the seven thematic statements substantiating the conclusion of this research will be introduced successively. The classical and contemporary theories relevant to the themes are also deduced in due course.

8.2 Themes

Altogether, there are seven themes (Figure 51) emerged from data saturation in relation to Student Learning Experiences (SLE), Student Learning Outcomes (SLO) and Assessment Practices:

Figure 51: Seven Themes developed by Coding



(Source: Sou, 2011)

8.2.1 Favorable SLE in Action Learning and Research

Theme 1 is "Favorable SLE in Action Learning and Action Research". Statistically, data analysis in Chapter V reveals that the candidates (n = 1,074) of the subject program appeared to react favorably to the action learning and action research throughout the program. Amongst 17 core subjects, the candidates perceive (Figure 52) that:

- ✧ two core subjects (Human Resources Management and Business Strategic Management) are both relative important and relative satisfactory;
- ✧ three core subjects (Organization and Management, Financial Management and Marketing Management) are relative important; and
- ✧ two core subjects (Economic for Business and Organization Behavior) are relative satisfactory.

Figure 52: Relative Important and Relative Satisfactory Subjects perceived by the Candidates

Core Subject	Relative Important	Relative Satisfactory
(2)	Organization and Management	N/A
(3)	Financial Management	N/A
(7)	N/A	Economics for Business
(8)	N/A	Organization Behavior
(9)	Human Resources Management	
(10)	Marketing Management	N/A
(16)	Business Strategic Management	

(Source: Figure 17 in Chapter V of this Dissertation; Sou, 2011)

On the contrary, the candidates perceive that three core subjects and three elective subjects are both relative unimportant and relative unsatisfactory (Figure 53):

Figure 53: Relative Unimportant and Unsatisfactory Subjects perceived by the Candidates

Core Subject	Elective Subject	Relative Unimportant	Relative Unsatisfactory
(4)	N/A	Quantitative Methods Analysis	
(5)	N/A	Production Management	
(12)	N/A	Small Business Management	
N/A	**(2)**	Auditing	
N/A	**(4)**	Comparative Management	
N/A	**(5)**	Total Quality Management	

(Source: Figure 17 in Chapter V of this Dissertation; Sou, 2011)

As for the research process (Figure 54) of the subject program, the candidates perceive that three aspects of the research process as relative important and relative satisfactory; and another three aspects as relative unimportant and relative unsatisfactory:

Figure 54: Relative Important/Satisfactory and Unimportant/Unsatisfactory Aspects of Research Process

Aspect	Relative Important/Satisfactory	Relative Unimportant/Unsatisfactory
1	Guidance on the Choice of Research Methodology	Guidance on the Choice of Research Tools
2	Guidance on the Application of Research Methods	Guidance on the Compilation of Thesis
3	Close Supervision of Research Activities	Guidance on Literature Review

(Source: Figure 19 in Chapter V of this Dissertation; Sou, 2011)

As for the appraisal process (Figure 55) of the program, the candidates perceive that four aspects of the appraisal process as relative important and relative satisfactory:

Figure 55: Relative Important and Satisfactory Aspects of Appraisal Process perceived by the Candidates

Aspect	Relative Important	Relative Satisfactory
1	Earnestness of Internal Examiner	
2	Earnestness of External Examiner	
3	Professional Knowledge of Internal Examiner	
4	Proliferation of Personal Knowledge	

(Source: Figure 20 in Chapter V of this Dissertation; Sou, 2011)

Simultaneously, the candidates perceive (Figure 56) that:

- ✧ one aspect of the appraisal process (Appropriateness of Appraisal Mode) is relative unsatisfactory;
- ✧ three aspects of appraisal process (Appropriateness, Arrangement and Environment of Viva Voce) are both relative unimportant and relative unsatisfactory; and
- ✧ one aspect of the appraisal process (Length of Viva Voce) is relative unimportant.

Figure 56: Relative Important/Satisfactory and Unimportant/Unsatisfactory Aspects of Appraisal Process

Aspect	Relative Unimportant	Relative Unsatisfactory
1	N/A	Appropriateness of Appraisal Mode
2	Appropriateness of Viva Voce	
3	Arrangement of Viva Voce	
4	Environment of Viva Voce	
5	Length of Viva Voce	N/A

(Source: Figure 20 in Chapter V of this Dissertation; Sou, 2011)

When the strength of the subject program lies on the two core subjects (Human Resources Management and Business Strategic Management) which the candidates find them relative important and satisfactory (Figure 52), the candidates enjoy a favorable Student Learning Experiences in the learning process. This perception may be reinforced by the other core subjects (Economics for Business and Organization Behavior) which the candidates find them relative satisfactory. Though the candidates find three core subjects (Quantitative Methods Analysis, Production Management and Small Business Management) and three elective subjects (Auditing, Comparative Management and Total Quality Management) relative unsatisfactory (Figure 53), they do not have much unfavorable Student Learning Experiences in the subject program because those subjects are relative unimportant.

The candidates appear to have favorable Student Learning Experiences during the research process when there is proper guidance on the choice of research methodology and methods (Figure 54). Besides, they are happy that there is close supervision of the research activities (Figure 54). Simultaneously, the candidates perceive that guidance on the choice of research tools, compilation of thesis and literature review are both relative unimportant and relevant unsatisfactory (Figure 54). On the contrary, the examiners during the appraisal process perceive that these three aspects are relative important. Conflicting views of the candidates and the examiners may be attributed to the unfavorable Student Learning Experiences in the appraisal process.

The candidates find that both the internal and external examiners are earnest in the appraisal process (Figure 55). They appreciate the professional knowledge of the internal examiners. They also feel that the viva voce enables them to increase their personal knowledge (Figure 55). All these give the candidates favorable Student Learning Experiences. However, the candidates appear to have less favorable Student Learning Experiences in the appropriateness of the appraisal mode. It may be attributed to be

different expectation between the candidate and the examiners on the Student Learning Outcomes. When the examiners attack the candidates on their choice of research tools, literature review and compilation of the thesis (Figure 54), the candidates may have negative feeling toward the appropriateness of the appraisal mode. Such negative feeling may lead to the candidates' relative unsatisfactory perception on the appropriateness, arrangement and environment of the viva voce (Figure 55).

Although the candidates originally perceive the above three aspects of viva voce as relative unimportant, their less favorable Student Learning Experiences may be intensified when the length of viva voce (Figure 55) does not allow the candidates to have adequate interaction with the examiners during the viva. Some examiners tend to adopt "Tell and Sell" or "Tell and Listen" approach in the viva instead of "Problem-Solving" or "Composite Approach" (Hall & Goodale, 1990; Sou, 2001d). This phenomenon (Figure 20 in Chapter V) reasons out that the candidates have favorable Student Learning Experiences in learning and research processes but less favorable SLE in the appraisal process.

8.2.2 MBA Candidate becoming Career Manager

Theme 2 is "MBA Candidate aims to be Career Manager". Data analysis in Chapter VI reveals that the candidates of the subject program have a clear aim of being career managers. It is noted that some candidates work in public sector or stated-owned enterprises as high-ranking comrades or mid-career civil servants while some work in private sector or small-to-medium enterprises. They wish to pursue their career in the management profession. Completing the action research and passing the viva enables the candidates to get the national certification of Chinese Career Manager (CCMC).

In conducting the action research, some candidates may choose big topics (Figure 23 in Chapter VI) because they are working in public sector or state-owned enterprises. A research project with large scope of study does not meet the requirements of a valuable action research. Therefore, they have to revise their research projects in order to prove that they are fit for the award of an MBA degree and the vocational title of Chinese Career Manager. Despite the obstacle in the form of a viva, the candidates still find the action learning (Gregory, 1994; Revans, 1980; Sou, 2000 & 2001a-c) and action research of the subject program rewarding.

In reality, the candidates of the subject program may have to face more obstacles in their career development even if they could pass the viva and earn an MBA degree or get the vocational title of Chinese Career Manager. Most of them are not proficient in English (Figure 24 and Figure 25 in Chapter VI) when the subject program is intended for Chinese career managers. Therefore, the examiners believe that the candidates should have reasonable proficiency in English language whilst doing business in a global environment. Otherwise, the candidates may not be effective career managers even though they are proficient in problem-solving, data-analysis and decision-making through the learning and research processes of the subject program.

8.2.3 MBA Candidate should be Genius Pig rather than Copy Cat

Theme 3 is "MBA Candidate should be Genius Pig rather than Copy Cat". The learning and research processes of the subject program aim to change the Attitudes, Skills and Knowledge (ASK) of the candidates and equip themselves to be career managers. During the appraisal process, the examiners diagnose the candidates' attitudes, skills and knowledge (ASK) and assess their performance in tacking issues or problems in their action research projects. In marking the candidates' theses and assessing the candidates' reflections

(Figure 26 in Chapter VI), the examiners discover that some candidates appear to manifest "Descriptive Writing" or "Descriptive Reflection" rather than "Dialogic Reflection" or "Critical Reflection" (Hatton & Smith, 1995).

The examiners diagnosed that the candidates might simply copy what they had learnt from the subject program or what their thesis supervisors knew. The examiners long to see that the candidates of the subject program may become "understanding seekers" rather than "knowledge seekers" (Entwistle, 1987 & 1992; Gibbs, 1990). Mere copying could hardly develop genuine "understanding seekers". Data analysis in Chapter VI suggests that understanding seeking candidates may have more desirable "Student Learning Outcomes" than knowledge seeking candidates.

Candidates with desirable "Student Learning Outcomes" will be awarded an overall grading of "outstanding" in their assessment. However, only a great minority of the candidates can earn the highest overall grading in their action researches, master theses and viva performance. This type of observable Student Learning Outcomes (Figure 27 in Chapter VI), namely "Extended Abstract", is scarce in the subject program. Instead, lower levels of observable student learning outcomes like "Relational, Multistructural, Unstructural or Pre-Structural" (Biggs, 1996; Hatton & Smith, 1995) are predominant amongst the candidates of the subject program.

It is preferable for a candidate to be a genius pig than a copy cat. The worst copy cats are those candidates stealing other people's intellectual property or buying theses for submission as their own works. In a less worse case, there are some obvious misconduct in referencing protocol which may be attributed to the candidates' unfamiliarity with the academic conventions. Through the appraisal process of the subject program, examiners could assess which candidates are copy cats and which candidates are genius pigs.

Besides, the thinking, learning and writing styles of the candidates may be affected by their background, mindsets and even their organizational climate. Some candidates are nurtured in planned economy or political job settings which influence their learning orientations, mindsets and writing styles. They are used to a paradigm of political writing (Figure 28 in Chapter VI) which annoys the examiners. This phenomenon is once popular, particularly amongst those candidates working in public sector or state-owned enterprises.

Background, mindsets or organizational climate may also affect the student learning outcomes in the aspects of the candidates' Attitudes, Skills and Knowledge ((Johnson et al., 1993 & Lenning, 1988)). Data analysis in Chapter VI also suggest a correlation between the candidates' "Attitudes, Skills and Knowledge" and "Understanding" (ASKU) as reported in the preceding paragraphs. The correlation of these four learning components, namely ASKU (Brown et al., 1997) is profound. Nonetheless, it can be illustrated in a tetrahedron (Figure 29 in Chapter VI). This research surfaces that the desirable student learning outcomes should be a common output of these four learning components.

Output of the above four learning components could develop the candidates' cognitive, social and cognative skills (Brown et al., 1997; Sou, 2009). These three skills are essential to the personal growth and career development of the candidates. In the subject program, the skills of the candidates are taught through the medium of the core and elective subjects and the guidance of the thesis supervisors. Ironically, the thesis supervisors assess that the candidates have grasped these three skills before the viva. The examiners may think the other way in the viva.

8.2.4 Supervisor's Assessment contrasting with Examiner's Assessment

Theme 4 is "Supervisor's Assessment may contrast with Examiner's Assessment". Data analysis in Chapter VII reveals that the thesis supervisors had a deep belief on Norm-Referenced Assessment (Brown et al., 1997). Therefore, they tend to mark over-generously the candidates before the viva. In contrast, the examiners regard the assessment task as an objectives-referenced Assessment instead of a criterion-referenced assessment (Osterlind, 2008). They tend to assess the candidates' performance against the objectives of the subject program. Therefore, there are variations between the supervisors' assessment and the examiners' assessment. Data analysis in Chapter VI also sheds light to this phenomenon.

In Chapter V, the history of significant variations between the supervisors' assessment and examiners' assessment is traced (Bell, 1980; Diederich, 1957; Edgeworth, 1890; Hartog, 1935 & 1936; Newstead and Dennis, 1994). Such phenomenon has subsisted in the assessment practices of the subject program since 2003. In Chapter VII, statistical tests (n = 67) elaborate this phenomenon in an explicit way (Figure 41 to Figure 47). Quantitative data analysis suggests that there are considerable variances between the markings of thesis supervisor, internal and external examiners on the sampled candidates.

Further study of technical literature in Chapter VII discloses that there is a long history of marking variances between assessors from 1890 (Figure 48). In comparison, the variances in marking amongst the thesis supervisors, internal and external examiners of the subject program appear to be nominal. Nonetheless, Theme 4 substantiates that "variances in marking" is one of the areas to be improved in the assessment practices of the subject program.

8.2.5 Management Learning Styles facilitating Fitness for Purposes

Theme 5 is "Management Learning Styles may facilitate the achievement of the Fitness for Purposes". During the appraisal process, the examiners perceive that the candidates have four styles of management learning (Mumford, 2008), namely Intuitive Learners, Incidental Learners, Retrospective Learners, and Prospective Learners. The examiners think that the thesis supervisors have a role in developing the candidates' most favorable management learning styles. Otherwise, the candidates may not have the desirable student learning outcomes through the research process.

Data analysis in Chapter VI surfaces four management learning styles which are more likely to be observed over different types of candidates. Firstly, candidates working in public sector prone to be "Intuitive Learners" (Figure 31). Secondly, candidates working in state-owned enterprises prone to be "Incidental Learners" (Figure 32). Thirdly, candidates working in the private sector prone to be "Retrospective Learners" (Figure 33). Fourthly, candidates working in the transnational enterprises prone to be "Prospective Learners" (Figure 34).

Whilst developing Theme 3, it is proposed that the candidates' background, mindsets or organizational climate may affect the student learning outcomes. Theme 4 reinforces this proposition. Furthermore, Theme 4 proposes that initial preference for one of the four management learning styles should not bar the candidates from shifting to other management learning styles. Thesis supervisors can play a proactive role to develop candidates' appropriate management learning styles during the research processes. This would facilitate the candidates' learning from managerial experiences (Figure 35 in Chapter VI) and the achievement of desirable student learning outcomes.

8.2.6 Favorable SLE not equating with Desirable SLO

Theme 6 is "Favorable Student Learning Experiences cannot be equated with Desirable Student Learning Outcomes". Data analysis in Chapter V reveals that the overall Student Learning Experiences of the candidates on the subject program is favorable. As read with the data analysis in Chapter VII, it is evident that the less favorable Student Learning Experiences stem from the candidates' perception on the "fairness" and "reliability" (Brown et al., 1997) of the assessment task. Marking variances between the thesis supervisors and the examiners are contributory to this phenomenon.

Data analysis in Chapter VI also disclose that some candidates with favorable Student Learning Experiences do not achieve desirable Student Learning Outcomes. From the perspectives of the examiners, the observable student learning outcomes of most candidates are far from the ideal state such as "Understanding Seekers" (Entwistle, 1987 & 1992) or "Prospective Learners" (Mumford, 2008). Therefore in the subject program, the candidates' favorable SLE cannot equate with the desirable student learning outcomes.

Data analysis in Chapter VII further surfaces a "Good Student Learning Experiences versus maybe better Student Learning Outcomes" phenomenon. The candidates consider the subject program a good one on their SLE and their interpretation of "Fitness for Purpose" (Brown and Knight, 1994; HEFCE, 1993) along with their perceived objectives of the subject program. On the other, the examiners consider the subject program on the basis of the pre-determined standards of the society and their interpretation of "Fitness of Award" (HEFCE, 1993) along with the performance of the candidates in the assessment tasks. Conflicting views of the candidates and the examiners are also contributory to the emergence of Theme 6.

8.2.7 Robust Assessment Practices ensuring Quality Education

Theme 7 is "Robust Assessment Practices can ensure Quality Education". Data analysis in Chapter VII reveals that robust assessment practices are built on fairness, reliability and validity (Figure 36 in Chapter VII). On the other, robust assessment practices can be built on effective assessment tasks taking a sample of what the candidates do, making inferences and estimating the worth of their actions (Brown et al., 1997). This research surfaces some shortcomings of the above assessment task undertaken by the candidates of the subject program.

The shortcomings (Figure 37 in Chapter VII) include:

- ✧ Overload of the candidates and the examiners.
- ✧ Insufficient time for the candidates to complete the action research in the time available and insufficient time for the examiners to mark the theses before the viva.
- ✧ Inadequate or superficial feedback provided to the candidates during the viva.
- ✧ Wide variations in assessment demands of different panels and marking across panels or within a panel.
- ✧ Wide variations in marking by supervisors.
- ✧ Fuzzy or non-existent criteria; undue precision and specificity of marking schemes or criteria.
- ✧ Candidates do not know what is expected of them or what counts as a good or bad research or thesis.

Despite the above shortcomings, the examiners and registrars of AIOU have a strong belief on effective assessment practices which could benefit all the stakeholders of the subject program (Figure 38 in Chapter VII). To ensure the effectiveness of assessment practices in the subject program, there is a robust assessment system with thesis

supervisors, internal and external examiners in place. Although there are variations in the inferences drawn by the different examines and different panels, AIOU has placed extra efforts in ensuring quality of the subject program through assessment practices.

In estimating the worth of the candidates' action in their research projects, the internal and external examiners play an important "quality" role to ensure the "fairness", "reliability" and "validity" (Hogatt and Hazel, 1970; McCormick, 1979; Scott, 1963; Sou, 2001d & 2008) of the assessment task. Apparently, the four-tier markings of the candidates by (1) thesis supervisors, (2) internal examiners, (3) external examiners and (4) academic panels (Figure 39) are robust assessment practices for quality assurance of the subject program. The appraisal process in the form of a viva also plays a critical role of gatekeeper (Figure 40) for quality assurance.

8.3 Conclusion

In Chapter I and Chapter II of this dissertation, it comes to light that MBA degree programs are growing in mainland China (Beijing Youth Daily, 2003; A-Performers.com, 2009 & 2010). The society needs management education to groom career managers for the societal development. Under the accountability system (Madaus & Stufflebeam, 1984), quality assurance of management education program is essential for grooming career management to suit societal needs. In the global context of management education, accreditation has a place in the quality assurance of educational programs (The Association of MBAs, 2007).

The Association of MBAs accredits MBA degree programs worldwide. Its accreditation criteria lie on the purpose, outcomes, curriculum and assessment scheme of an MBA degree program seeking global accreditation. Accreditation of The Association of MBAs does not find a place in mainland China because MBA degree has been introduced to the

management education market of mainland China since 1982. Nonetheless, this global benchmark mirrors that **purpose (objectives)**, **outcomes**, **curriculum** and **assessment scheme** are essential elements in quality assurance of MBA degrees.

In Chapter III of this dissertation, literature review surfaces three classical approaches[55] towards educational quality from 1800s to 1960s (Sou & Zhou, 2007). One of those three approaches viz. Tylerian Approach concentrates on Student Learning Outcomes that is identical to the benchmark of The Association of MBAs. Then, literature review also surfaces four types of conventional educational quality assurance approaches[56] for higher education developed between 1958 and 1999 (Chun, 2002; Sou, 2006). Amongst those four approaches, "Direct Measurement of Student Learning Outcomes" is the approach for educational quality assurance in terms of student learning outcomes.

Furthermore, literature review reveals that scholars have developed 22 foundational program evaluation models (Figure 3 in Chapter III) from 1960 to 1999 (Sou and Zhou, 2007; Stufflebeam, 2000). Along with 17 alternative modes (Footnote 7 in Chapter III), there is an array of multifarious program evaluation models for educational quality assurance. Mixed application of these models further produces more "hybrid models" for quality assurance of different educational programs. However, literature review also reveals that none of those models are tailor-made for quality assurance of an MBA degree program.

In Chapter V of this dissertation, it is known that the use of quality management tools like ISO 9001:2000 has emerged in the education industry (British Standards Institute, 2000). Quality management practitioners and experts make it possible with migration of best

[55] (1) Accreditation or Professional Approach, (2) Testing Approach, and (3) Tylerian Approach.

[56] (1) Conduct of Student Surveys, (2) Comparison of Actuarial Data, (3) Ratings of Institutional Quality, and (4) Direct Measurement of Student Learning Outcomes.

practice from one sector to the other sector. Furthermore, quality management professionals facilitate the integration of quality assurance philosophy from the business world into education industry with the schools of thought in other disciplines. Nonetheless, this research reveals that one size does not fit all!

Lending the quality assurance tools of the business world, a prototype of program evaluation model (Figure 15 in Chapter V) with emphasis on the essential processes is implemented for the quality assurance and international certification (ISO) of the subject program (Sou & Zhou, 2005a-d; Zhou & Sou, 2005a-b). However, this research surfaces that customer-centered quality management system of the 21st century cannot be transplanted from the business world to an educational institution across organizational, social and cultural boundaries.

Customer-centered process approach lays much emphasis on the Student Learning Experiences on the curricular processes[57] of the subject program (Sou, 2006a-c; Sou & Zhou, 2005a-e & 2007). Other than Student Learning Experiences, Student Learning Outcomes should be duly considered for the quality assurance of the subject program. In Chapter VI of this dissertation, data analysis discloses that student learning outcomes in terms of the sampled candidates' performance in the master research and viva voce can be observed through their reflection[58] (Schön, 1983, 1988 & 1991) upon the learning and research processes of the subject program. In addition, desirable Student Learning Outcomes[59] (Biggs, 1996) of the sampled candidates are influenced by their learning

[57] Teaching Process, Learning Process, Research Process, and Appraisal Process.

[58] (1) Descriptive Writing, (2) Descriptive Reflection, (3) Dialogic Reflection, and (4) Critical Reflection.

[59] (1) Prestructural, (2) Unstructural, (3) Multistructural, (4) Relational, and (5) Extended Abstract.

orientations[60] (Entwistel, 1987 & 1992) and management learning styles[61] (Mumford, 2008).

Desirable Student Learning Outcomes appears to be the end product of four learning components, namely Attitudes, Skills, Knowledge and Understanding (ASKU). These four learning components with an acronym of ASKU have interaction with each other (Brown et al., 1997). With proper interaction, the cognitive, social and cognative skills (Brown et al., 1997) of the students of the subject program could be enhanced to the desirable level. Ultimately, they can be cultivated to become effective career managers.

This research establishes that it would be of value to learn more about Student Learning Outcomes than Student Learning Experiences. Apart from the Student Learning Outcomes, assessment scheme of an MBA degree program is also a global benchmark of The Association of MBAs for quality assurance of MBA degree program. In Chapter VII of this dissertation, data analysis highlights the assessment practices of the subject program wherein the issues of quality assurance are being tackled. Shortcomings (Figure 37) and benefits (Figure 38) of the assessment practices are analyzed.

In Chapter VII, variances of assessments amongst thesis supervisors, internal and external examiners, and academic panels are statistically tested. Eventually, the fairness, reliability and validity (Hogatt and Hazel, 1970; McCormick, 1979; Scott, 1963; Sou, 2001d & 2008)) of the assessment practices in the subject program are measured. Finally in this chapter, thematic statements weave facts and findings of this research into conclusion. As a result, conclusion (Sou, 2011) provides the answers to the two research problems framed in Chapter I:

---

[60] (1) Knowledge Seekers and (2) Understanding Seekers.

[61] (1) Intuitive Learners, (2) Incidental Learners, (3) Retrospective Learners, and (4) Prospective Learners.

1. **Quality of Educational Provision**

- ✧ Is this (the research leading to a master thesis) a good learning experience for the students in the educational program?
  - ⌑ **The action research leading to a master thesis appears to be a favorable Student Learning Experiences (SLE) in the subject program. However, it appears that favorable SLE cannot be equated with desirable Student Learning Outcomes (SLO).**
- ✧ Does this (the research leading to a viva) contribute to a good learning experience for the students in the educational program?
  - ⌑ **The research leading to a viva contributes to a less favorable Student Learning Experiences (SLE) in the subject program. It appears that the assessment tasks undertaken by the thesis supervisors have not aligned with the assessment tasks undertaken by the internal and external examiners. Therefore, supervisors' assessment may contrast with examiners' assessment.**

2. **Fitness for Purpose**

- ✧ How well does the educational experience (performance in the master research and the viva) of the students match the claims of the educational program provider?
  - ⌑ **It appears that the students consider the subject program a good one on the favorable Student Learning Experiences (SLE) and their interpretation of "Fitness for Purpose" along with their perceived objectives of the subject program.**
  - ⌑ **It appears that the examiners consider the subject program a**

can-be-better one on the desirable Student Learning Outcomes (SLO) and their interpretation of "Fitness of Award" along with their perceived objectives of the subject program.

- **It appears that the examiners of the subject program have reservation about the Attitudes, Skills, Knowledge or Understanding (ASKU) of the candidates in relation to their performance in the master research and the viva.**
- **It appears that there are still room for the candidates to develop their Attitudes, Skills, Knowledge or Understanding (ASKU) in order to become effective career managers upon their graduation from the subject program.**

IX. RECOMMENDATIONS

9.1 Recommendations

In Chapter VIII, it is concluded that there are some areas for improvement in the subject program. In the last chapter of this dissertation, specific recommendations on the basis of the seven themes (Figure 51 in Chapter VIII) will be made to address nine areas for improvements. At the end of each recommendation, the relevant themes are bracketed into groups for cross-reference.

1. **Manageable Action Research is a Must.** The candidates of the subject program reflect that there is insufficient time for them to complete the action research while the examiners say that the assessment task is tedious. It is partially attributed to the wide scope of study. Therefore, the thesis supervisor should ensure that the candidate is conducting a manageable action research which would not overload either the candidate or the examiners. *(Theme 1 + 6)*
2. **Standard Format facilitates the Marking of a Thesis.** The supervisor should encourage the student to use a 5-chapter format in compiling the thesis. Standard format helps the examiners digesting and marking the thesis in a timely and effective manner. Above all, the 5-chapter format is more compatible with the marking criteria for the thesis. *(Theme 1 + 6 + 7)*
3. **Continuous Assessment by the Thesis Supervisors helps Students' Reflection.** When the examiners judge that the students should have deep learning during the action research, the thesis supervisors should continuously assess the students during the research process. A thesis supervisor should pay frequent attention to the student's reflection upon learning orientations (knowledge seekers versus understanding seekers) and development of management learning styles (Intuitive

Learner, Incidental Learner, Retrospective Learner and Prospective Learner). *(Theme 2 + 3 + 7)*

4. **Pre-Notification of the Examiners' Assessments helps Students' Reflection.** The examiners wish to give more feedback to the students in the viva but the time does not allow it. The examiners should provide their assessment forms to the students before the viva. Pre-Notification would facilitate a composite approach of interaction between the examiners and the student during the viva which may foster more favorable SLE on the appraisal process of the subject program. *(Theme 1 + 7)*
5. **Variances in Assessment should be avoided.** Variances in the assessments made by the thesis supervisors and the panels cause the students to have less favorable SLE. Regular statistical tests on the variances of the markings by (a) the thesis supervisor, (b) the internal examiner, (c) the external examiner, and (d) the academic panel should be conducted. Findings of the statistical tests should be made known to the thesis supervisors and examiners. Intensive study of the assessment forms completed by the supervisors and examiners would lead to alignment in individual and collective markings. *(Theme 4 + 7)*
6. **Change on Combination of Academic Panels eliminates Marking Variances.** Examiners and registrars sometimes note that different combination of the panel members varies in their assessment demand which jeopardize the fairness, reliability and validity of the assessment practices of the subject program. Additionally, there are wide variations in marking across and within a panel. Change on the combination of the panel members would minimize the phenomenon of "variable assessment demands of different panels". Regular statistical tests on their marking variances and intensive study of the findings would give more insight into this area. Then, combination of academic panels can be optimized. *(Theme 4 + 7)*

7. **Copy Cat should be transformed into Genius Pig.** Thesis supervisor should play an active role in transforming the students' learning orientations. Transforming the students of the subject programs from "knowledge seekers" into "understanding seekers" meet the objectives of the subject program. Personal and career development of the students can be achieved by enhancement of their Attitudes, Skills, Knowledge and Understanding (ASKU). *(Theme 2 + 3)*

8. **Observable Student Learning Outcomes (SLO) can be Internal Benchmark.** Examiners believe that precise and specific marking scheme and criteria exist in the assessment practices of the subject program. However, they may not be viable for the action research in the management education and are subject to the thesis supervisors' and examiners' own interpretation. In such circumstances, five levels of Bigg's Schema of Structure of the Observed Learning Outcome should be promoted amongst the thesis supervisors, internal and external examiners. *(Theme 4 + 7)*

9. **Management Learning Styles of the students should be identified.** Management learning styles of the students have a correlation to their background, organization climate and culture. Management learning styles influence the students' achievement of the desirable Student Learning Outcomes. Management learning styles can be shifted from one to the other during the research processes. Thesis supervisors can play an active role of identifying the students' predominant management learning styles and help them change[62] as appropriate. *(Theme 1 + 5 + 6 + 7)*

[62] This theme assumes that people can change their learning style whereas the literature suggests that this is not always possible.

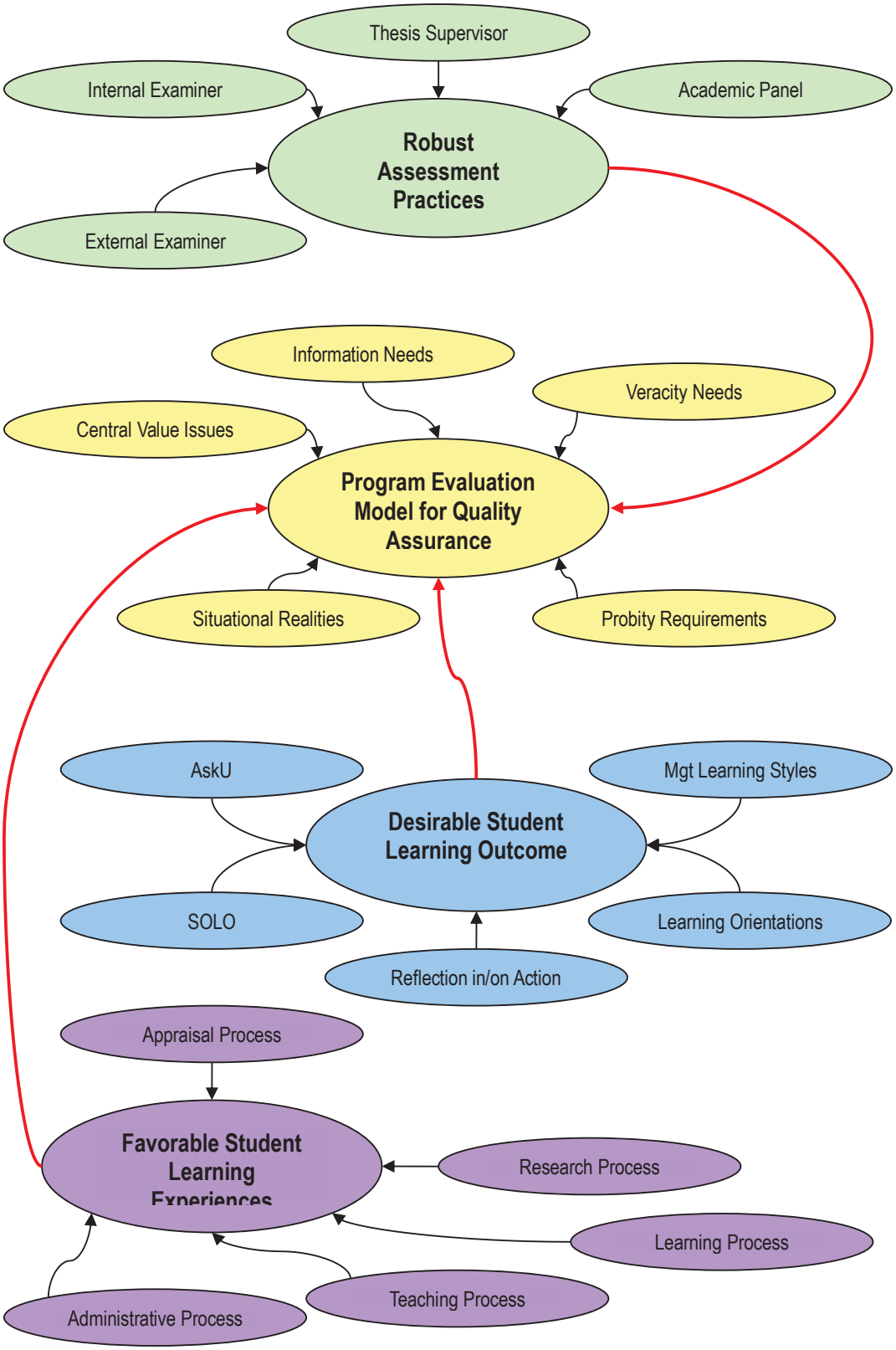
9.2 Contribution to Body of Knowledge and Professional Practices

Though no assessment practices can be transplanted from one institution to another across organization, social and cultural boundaries, a program evaluation model based on assessment practices with due consideration of Student Learning Experiences and Student Learning Outcomes is feasible for quality assurance of an educational program. Definitely, one size does not fit all but a feasible program evaluative model is always of referential value to kindred programs. Anyhow, a holistic program evaluation model should be more powerful than the Tylerian approaches for evaluation.

Acquiring and retaining a competitive edge in business requires modern organizations to constantly seeking improvement opportunities for advancement. Economic changes and structural shifts in the many business environments have created formidable challenges for modern organizations, stimulating them to re-evaluate and improve their management practices. Educational institutions, like other modern organizations, have been inspired to adopt a more well-rounded and holistic approach in driving institutional improvement initiatives, placing greater focus on quality management and social expectation.

Many organizations are working very hard to offer quality services and products, which provide a route towards business excellence. Only modern organizations demonstrate their ability and commitment in delivering quality services will excel in today's highly competitive market environment. Establishing an effective and efficient quality assurance system has become an indispensable component of modern competitiveness. Nowadays, it is encouraging that world class quality management tool has been introduced into the education institutions which might normally be beyond their reach previously.

Figure 57: Program Evaluation Model for Quality Assurance of AIOU's MBA Degree Program



(Source: Sou, 2011)

This research contributes to the wide body of knowledge in the establishment of a refined program evaluation model with emphasis upon Student Learning Experiences, Student Learning Outcomes and Assessment Practices of the subject program (Figure 57). This model may be beneficial to the future development of the subject program in the education market of the Mainland. Furthermore, this model can be regarded as a reference of continually offering MBA degree programs to the prospective Chinese Career Managers. Educators may regard it as food for thought in the provision of quality management education. This is the contribution to the professional practices of management education in mainland China.

9.3 Implications for Further Research

When development of program evaluation model must be home-grown, recognizing its context of the program. It is expected that program evaluation for quality assurance may move beyond external processes of accreditation to internal mechanisms of assessment practices. Program evaluation will promote the growth and sustenance of a culture of quality within an institution.

To name a few but not all, further research on the following aspects would lead to a more holistic view on the quality assurance of management education program:

- ✧ Utilization of Action Research as an Appropriate Assessment Task for the MBA Degree Program in Mainland China
- ✧ Implementation of Effective Assessment Practices for the MBA Degree Program in Mainland China
- ✧ Development of a Program Evaluation Model for Quality Assurance of MBA Education in Mainland China (Sou, 2011)

Appendix A: English Translation of MBA Thesis Assessment Form

Name of Student : _____ Student Number : _____

Name of Supervisor : _____

Title of Thesis : _____

Score: A = Outstanding (80 or above); B = Credit (70-79); C= Pass (60-69); D= Failed (59 or below).

Assessment Guidelines		Examiner's Comments	Score
Choice of Title of Thesis and Abstract (10 marks)	Title of Thesis should be direct and clearly reflects the Research Questions. Thesis should be submitted with Abstract in Chinese and English. Abstracts with Keywords should clearly show the core of Research Problems and concisely describe the Subject of Study, Research Methodology, Main Contents of the Thesis and the Findings, Conclusions and Recommendations of the Research.		
Introduction and Literature Review (20 marks)	In the Chapter of INTRODUCTION, there should be Background of Study, Definition of Research Problems, Relationship between the Research and the Student's Job, and Objectives of Research or Research Questions to be Addressed. In the Chapter of LITERATURE REVIEW, there should be differentiation, integration, processing, and analysis of the relevant literature. The Student should concisely generalize and collate the relevant theories. The Student should be able to bring up his or her own perception and comments thus establishing a theoretical foundation and framework of the Research.		
Research Methods & Design (10 marks)	Size of Sampling Frame and Selection of Respondents / Participants in the Research should be justified. Choice of appropriate Research Tool(s), Implementation of the Research, and Collection of Data (including time, venue, methods, and processes) should be elaborated.		

Data Processing & Analysis, and Discussion (30 marks)	The Student should duly, exactly, and reasonably process and analyze the data. The Student should be able to differentiate the findings of the present and previous related researches. The Student should also differentiate deduction and adduction as well as integrate theories with the practices. Charts should be compiled and edited for clear illustration. They should be duly inserted and labeled while their sources should be properly cited.		
Conclusions and Recommendations (20 marks)	In the last Chapter of CONCLUSIONS, the Student should sum up Findings of the Research that include, perspectives, options, strategies, measures, etc. There should be pragmatic and specific RECOMMENDATIONS. In addition, the Student may speak out the limitation of the Research and suggest ways of improvement. Further study may be proposed as well.		
Compilation, Format and Style of Thesis (10 marks)	The overall structure of the Thesis should be logical and rational. Contents of the Thesis should be associated with the Title of Thesis. Perspectives and arguments should be thoroughly and well-reasoned. Use of words and paragraphs should ensure a smooth text-flow and clarity of presentation. Format and Style should be in compliance with the Thesis requirements. For instance, List of Contents, List of Figures, Appendices, Charts, Tables, Text and Headings should be properly numbered. In addition, Statistical Diagrams should be presented and used in an orderly manner.		
		Total Score	

Overall Comments and Suggestions

_____

_____

Signature：_____

Appendix B: Outline of Questions for Focus Group Interview

Primary Open Question:

- How well does the educational experience of the students match the claims of the educational program provider (HEFCE, 1993)?

Secondary Open Questions

- How do you think about the Student Learning Outcomes in the following aspects (Brown, Bull & Pendlebury, 1997) from the Action Researches leading to the MBA Theses and/or the performance of the MBA Candidates in the Viva Voce?
    1. Attitude
    2. Skills
    3. Knowledge
    4. Achievements
    5. Potential
    6. Intelligence
    7. Aptitudes
    8. Motivations
    9. Personality
- Are there any "Positive Case(s)" or "Negative Case(s)" to support your perception? If affirmative, please elaborate.

P.S. Strauss and Corbin (1998) assert that a researcher should not begin a project with a preconceived theory in mind unless his or her purpose is to elaborate and extend existing theory. Therefore, focus group interviews are not structured in this research. They are general and open to what the informants want to say. However, it is important to get the informants to think collectively in response to some stimuli material. Therefore, the quality framework of Higher Education Funding Council for Education (1993) and Student Learning Outcomes (Brown et al., 1997) are sensitively used in Focus Group Interviews.

Appendix C: Serial Research on AIOU's MBA Degree Program

Item	Survey	Data & Subjects	Data for Analysis
1	2002 Student Survey on Perceived Importance of 17 Core Subjects in the Curriculum of AIOU's MBA Degree Program (Zhou, 2002)	Available (n=580 MBA Students)	5-out-of-17 Most Important Core Subjects as perceived by the MBA Students in 2002.
2	2003-04 Student Survey on Teaching and Learning Processes in AIOU's MBA Degree Program in mainland China (Zhou, 2004)	Available (n=544 MBA Students)	Perceived Degree of Importance and Satisfaction of Pedagogy adopted by the Mainland/AIOU Faculty in the subject Program
3	2004 Student Survey on the Teaching and Learning Processes of AIOU's MBA Degree Program in Tianjin, PRC (Li, 2004)	Available (n=80 MBA Students)	Apparent Social Needs of Higher Education for Working Adults in Tianjin. Perceived Degree of Importance & Satisfaction of the Pedagogy adopted by the Mainland Faculty in the subject Program.
4	2004-05 Graduate Survey on Perceived Importance and Satisfaction of AIOU's MBA Degree Program in the Mainland (Sou & Zhou, 2004)	Available (n=924/1074 Candidates)	Degree of Importance and Satisfaction as perceived by AIOU's MBA Graduates: ✧ 17 Core Subjects ✧ 3-out-of-6 Electives ✧ Training Materials ✧ Action Research ✧ Viva Voce
5	2004 Educator Survey on AIOU's MBA Degree Program in mainland China (Sou & Zhou, 2004)	Available (n=10+ Educators)	Opinions from the Educators about the Curriculum of the subject Program.

BIBLIOGRAPHY

Aakar, D. A., Day, G. S., & Kumar, V. (1998). *Marketing Research*. New York: Wiley.

Abrahamson, M. (1983). *Social Research Methods*. New York, Prentice Hall.

Ainscow, M., Balshaw, M., Black-Hawkins, K., Hargreaves, D., & Hopkins, D. (1994). *Mapping Change in Schools. The Cambridge Manual of Research Techniques (1st Edition)*. Cambridge: Institute of Education.

Ainscow, M., Beresford, J., Harris, A., Hopkins, D., & West, M. (1997). *Creating the Classroom Conditions for School Improvement*. London: David Fulton.

Anaya, G. (1999). College Impact on Student Learning: Comparing the Use of Self-Reported Gains, Standardized Test Scores, and College Grades. *Research in Higher Education, 40*(5), 499-526.

Andrews, F. M. (1984). Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach. *Public Opinion Quarterly, 48*, 409-442.

Anonymous. (2007, 30.01.07). Australian points out the Poor English Efficiency of Asia Students *am730,* p. 3.

Anonymous. (2007). Professor of Yale University condemns Plagiarism of Chinese Students. *Oriental Daily,* p. A42.

A-Performers Editorial Team A (Ed.). (2009). *China Executive Education Guide 2009*. HKSAR: Job Market Publishing Limited.

---.. (2010). *China Executive Education Guide 2010*. HKSAR: Job Market Publishing Limited.

Argyris, C. & Schön, D. (1974). *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.

---. (1978). *Organizational Learning: A Theory of Action Perspective*. Massachusetts: Addison Wesley.

---. (1996). *Organizational Learning II: Theory, Method and Practice*. Massachusetts: Addison Wesley.

Asia International Open University. (2006a). *Prospectus - Master of Business Administration (MBA) / Diploma in Management (DipM)*. Macau SAR: Asia International Open University.

---. (2006b). *Prospectus - Doctor of Business Administration (DBA)*. Macau SAR: Asia International Open University.

Astin, A., Ayala, F., & Jacobi, M. (1987). *College student outcomes assessment: A talent development perspective.* (No. ASHE-ERIC Higher Education Report No. 7). Washington, DC:

Association for the Study of Higher Education.

Astin, A. W. (1968). Undergraduate achievement and institutional 'excellence'. *Science, 161*(August 1968), 661-668.

---. (1977). *Four critical years: Effects of college on beliefs, values, and knowledge.* San Francisco: Jossey-Bass.

---. (1991). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education.* New York: American Council on Education / Macmillan.

---. (1992). *Cognitive Development among College Undergraduates.* Unpublished Doctoral Dissertation, University of California, Los Angeles.

---. (1993). *What matters in college? Four critical years revisited.* San Francisco: Jossey-Bass.

Atkins, M. J., Beattie, J., & Dockrell, B. (1993). *Assessment Issues in Higher Education*. Sheffield: Employment Department.

Augustine, C. H., Benjamin, R., Bikson, T. K., Derghazarian, E., Gates, S. M., Kaganoff, T., et al. (2001). *Ensuring the quality and productivity of education and professional development activities: A review of approaches and lessons for DoD*. Santa Monica, CA: National Defense Research Institute, RAND.

Augustine, C. H., & Peterson, M. W. (2000). *The influences of regional accreditation associations on institutions' approaches to, support for and use of student assessment.* Unpublished manuscript, Chicago.

Babbie, E. (2004). *The Practice of Social Research* (10th ed.). Victoria: Wadsworth/Thomson Learning.

Bain, G., & Thomas, P. (1984). Contextual Differences of Learning Approaches: The Effects of Assessment. *Human Learning* (3), 227-240.

Baird, L. L. (1976). *Using self-reports to predict student performance.* New York: The College Board.

Baker, M. (2006). Measuring Corporate Social Impact - Art or Science? *Hong Kong Engineer, 3-4*(6), 16-17.

Banta, T. W., Lambert, E. W., Pike, G. R., Schmidhammer, J. L., & Schneider, J. A. (1987). Estimated student score gain on the ACT Comp Exam: Valid tool for institutional assessment. *Research in Higher Education, 273*, 195-217.

Banta, T. W., Lund, J. P., & Oblander, F. W. (Eds.). (1996). *Assessment in practice: Putting principles to work on college campuses.* San Francisco: Jossey-Bass.

Banta, T. W., & Palomba, C. A. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education.* San Francisco: Jossey-Bass.

Barnett, R. (1997). *Realizing the University*. London: University of London, Institute of

Education.

Bartz, D., Anderson, D., McCormick, W., & William, L. (1993). Improving the Performance Appraisal Conference. *Modern Management, 7 (4)*, 15-17.

Beaty, L., Bourner, T., Lawson, J., & O'Hara, S. (1997). Action Learning Comes of Age - Part 3: Action Learning for What? *Education + Training, 39*(5), 184-188.

Beere, J. (2001). Towards the Learning School. *Education today, 51*(4), 47-48.

Beijing Youth Daily. (2003). Demand of MBA Degrees in China Management Field. China Youth Daily.

Bell, R. C. (1980). Problems in improving the Reliability of Essay Marks. *Assessment and Evaluation in Higher Education, 5*, 254-263.

Berdie, R. F. (1971). Self-claimed and tested knowledge. *Educational and Psychological Measurements, 31*, 629-636.

Beresford, J. (2003). Developing Students as Effective Learners: The Student Conditions for School Improvement. *School Effectiveness and School Improvement, 14*(2), 121-158.

Beresford, J., Hopkins, D., & West, M. (1995). *Creating the Conditions for Classroom Improvement.* (An Interim Report). Bath: ECER.

---. (1998). Creating the Conditions for Classroom and Teacher Development. *Teachers and Teaching: Theory and Practice, 4*(1), 115-141.

Bettis, R., Friga, P., & Sullivan, R. (2003). Changes in Graduate Management Education and New Business Strategies for the 21st Century. *Academy of Management Learning and Education, 2*(3).

Bickerstaffe, G. (2000, 01.07.04). which MBA? Retrieved 01.07.04, 2004, from http://store.eiu.com/index.asp?layout=show_sample&product_id=970000297&country

Biggs, J. B. (1987). *Student Approaches in Learning and Studying*. Melbourne: Australian Council for Educational Research.

---. (1996). Enhancing Teaching through Constructive Alignment. *Higher Education, 32*, 347-364.

---. (1999). *Teaching for Quality Learning in University*. Buckingham: Society for Researcher in Higher Education and Open University Press.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the Quality of Learning*. New York: Academic Press.

Black, S. (1993). Portfolio Assessment. *The Executive Educator, 15*, 28-31.

Blake, R. L., & Vernon, D. T. A. (1994). Does Problem-Based Learning Work? A

Meta-Analysis of Evaluative Research. *Academic Medicine, 69*, 550-563.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. New York: Prentice-Hall.

Bogdan, R. & Biklen, S. (1992). *Qualitative Research for Education: An Introduction to Theory and Methods*. Boston, Allyn & Bacan.

Bohr, L., Nora, A., Pascarella, E. T., & Terenzini, P. T. (1994). *Is differential exposure to college linked to the development of critical thinking?* Chicago: Illinois University National Center on Postsecondary Teaching, Learning and Assessment.

Bohr, L. E., Bulakowski, C., Desler, M., Jacobs, M., Nora, A., Pascarella, E., et al. (1994). Cognitive effects of two-year and four-year institution: A preliminary study. *Community College Review, 22*(1), 4-11.

Boote, D. N., & Beile, P. (2005). Scholars before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational Researcher, 34*(6), 3-15.

Boulter, N., Dalziel, M., & Hill, J. (Eds.). (1996). *People and Competencies*. London: Kogan Page.

Bourner, T., Beaty, L., Lawson, J., & O'Hara, S. (1996). Action Learning Comes of Age: Questioning Action Learning. *Education + Training, 38*(8), 32-35.

Bourner, T., Bowden, R., & Laing, S. (2001). Professional Doctorates in England. *Studies in Higher Education, 26*(No. 1), 65-83.

Boykin, D. (2003). PEs find Growing Online Education Opportunities. *Engineering Times, 25*(9), 11.

Bradburn, N. M., & Sudman, S. (1988). *Polls and surveys: Understanding what they tell us*. San Francisco: Jossey-Bass.

Bradley, J. L., Draper, G. F., & Smith, M. K. (1993). *A national survey on assessment practices*. Knoxville, TN: University of Tennessee, Knoxville, Clearinghouse for Higher Education Assessment Instruments.

Brandt, R. M. (1958). The accuracy of self-estimates. *Genetic Psychology Monographs, 58*, 55-99.

Brennan, M. (1997). Struggles over the Definition and Practice of the Educational Doctorates in Australia. *Australian Educational Researcher, 25*(No. 1), 77-89.

Brennan, M., Green, B., & Lee, A. (1999). *The Rise of the Professional Doctorate: Changing Doctoral Education in Australia*. Rockhampton: Central Queensland University Press.

British Standard Institute. (2000). *ISO 9001:2000 Manual*. London: BSI.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*. London & New York: Routledge.

Brown, S., & Knight, P. (1994). *Assessing Learners in Higher Education*. London: Kogan Page.

Bunda, M. A., & Muffo, J. A. (1993). *Attitude and Opinion Data. In Banta, Trudy and Associates, eds. Making a difference: Outcomes of a decade of assessment in higher education, 139-50.* San Francisco: Jossey-Bass.

Bunnage, J. C., Carini, R. M., Kuh, G. D., & Ouiment, J. A. (2001). *Using focus groups to establish validity and reliability of a college student survey.* Paper presented at the 2001 AIR Forum, Long Beach, CA.

Butcher, B., & Hope, A. (2006). Embracing Change: Quality Assurance at the Open University of Hong Kong. In B. N. Koul & A. Kanwar (Eds.), *Perspectives on Distance Education: Towards a Culture of Quality* (pp. 113-124). Vancouver: Commonwealth of Learning.

Candoli, C. I., Cullen, K., & Stufflebeam, D. L. (1997). *Superintendent Performance Evaluation: Current Practice and Directions for Improvement (Evaluation in Education and Human Services).* MA: Kluwer Academic Publishing.

Carless, D., Joughin, G., & Liu, N. F. (2006). *How Assessment supports Learning - Learning-Oriented Assessment in Action*. Hongkong: Hong Kong University Press.

Carpenter, B., & Tait, G. (2001). The Rhetoric and Reality of Good Teaching: A Case Study across Three Faculties at the Queensland University of Technology. *Higher Education, 42*, 191-203.

Chaffee, E., & Sherr, L. (1992). *Quality: Transforming Postsecondary Education, ASHE/ERIC Higher Education Report 3*. Washington DC: George Washington University.

Chan, K. C. (1994). Learning for Total Quality an Action Learning Approach. *The Learning Organization, 1*(1), 17-22.

Chan, K. C., & Anderson, G. C. (1994). Academia - Industry Fusion Action Learning for Teaching Enterprise. *Industrial and Commercial Training, 26*(4), 28-32.

Chiang Jiang News. (2010, 05.01.10). Dissertation Black Market worth Ten Billion Dollars in Mainland China. *Ming Pao Daily News,* p. 1 of 1.

China Education Center. (Eds.) (2011). *MBA in China.* Nanjing: China Education Center.

China Social Sciences Academy. (2008, 04.01.08). CSSA Blue Book: Metropolitan's Degree of Satisfaction drops in Blooming Economy. *Oriental Daily,* p. A34.

Chun, M. (2002). Looking Where the Light is Better: A Review of the Literature on Assessing Higher Education Quality. *Peer Review* (Winter/Spring 2002), 16-25.

Clark, P. M. (2003). Quality in Digital Age. In A. Szucs, E. Wagner & C. Tsolakidis (Eds.), *The Quality Dialogue - Integrating Quality Cultures in Flexible, Distance and e-Learning. Proceedings of the 2003 European Distance Education Network Annual Conference, Rhodes,*

Greece. (pp. 1-6). Rhodes: EDEN.

Clarke, J. A. (1996, 16-18 October 1996). *Why EdD and not PhD? Some Students Perceptions*. Paper presented at the 'Which Way for Professional Doctorates' Conference, Coffs Harbor.

Clarke, R., & Newble, D. I. (1987). *Approaches to Learning in a Traditional and an Innovative Medical School*. Milton Keynes: SRHE & Open University.

Cohen, L., Manion, L., & Morrison, K. (2003). *Research Methods in Education* (5th ed.). London & New York: Routledge Falmer.

Cole, J. J. K., Nettles, M. T., & Sharp, S. (1997). *Assessment of teaching and learning for improvement and accountability: State governing, coordinating board and regional accreditation policies and practices*. Ann Arbor: University of Michigan, National Center for Postsecondary Improvement.

Converse, J. M., & Presser, S. (1989). *Survey questions: Handcrafting the standardized questionnaire*. Newbury Park, CA: Sage.

Cornesky, R. A., Clark, J. M., & Emerson, R. W. (1992). *Using Deming to Improve Quality in Colleges and Universities*. Madison, WI: Magma Publications.

Cornesky, R. A., McCool, S., Byrnes, L., & Weber, R. (1991). *Implementing Total Quality Management in Higher Education*. Madison, WI: Magma Publications.

Creighton, E., Mark, W., Robinson, G., Trafford, V., & Wisker, G. (2003). Recognizing and Overcoming Dissonance in Postgraduate Student Research. *Studies in Higher Education, 28*(1).

Daniel, J. (2006). Preface. In B. N. Koul & A. Kanwar (Eds.), *Perspectives on Distance Education: Towards a Culture of Quality* (pp. vii-viii). Vancouver: Commonwealth of Learning.

Davis, B., & Sumara, D. J. (2000). Curriculum Forms: on the Assumed Shapes of Knowing and Knowledge. *Journal of Curriculum Studies, 32*(6), 681-645.

Day, C., & Leithwood, K. (Eds.). (2007). *Successful School Principals in Changing Times: International Perspectives*. The Netherlands: Springer.

Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., et al. (2009). *The Impact of School Leadership on Pupil Outcomes* (No. DCSFRR018). Nottingham: National College for School Leadership.

Deming, W. E. (1982). *Out of Crisis*. Cambridge, MA: Massachusetts Institute of Technology.

DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology, 62*, 641-644.

Department of Employment, Education and Youth Affairs. (1998). *Learning for Life - Final*

Report. Review of Higher Education Financing and Policy (April 1998, chaired by Roderick West). Canberra: Australian Government Publishing Service.

Dey, E., Guan, F., Hurtado, S., Inkelas, K. K., Rhee, B., & Wimsatt, L. A. (1997). *Improving research on postsecondary outcomes: A review of the strengths and limitations of national data Sources*. Stanford, CA: National Center for Postsecondary Improvement.

Dhanarajan, G. (1998, 4 - 6 November 1998). *Access to Learning and Asian Open Universities: In Context*. Paper presented at the 12th Annual Conference of the Asian Association of Open Universities, HKSAR.

Dhanarajan, G., & Hope, A. (1992). Quality Assurance at the Open Learning Institute. In A. Craft (Ed.), *Quality Assurance in Higher Education* (pp. 207-219). London: Falmer.

Diederichm, P. (1957). *The Improvement of Essay Examinations*. Princeton: Educational Testing Service.

Dimmock, C. (2000). *Designing the Learning-Centered School: A Cross-Cultural Perspective*. London: Palmer Press.

Drake International. (2004). Analyze Your Career to identify Your Best Attributes. *South China Morning Post,* p. 44.

Duckart, T. (2005, 23.02.05). Autoethnography. Retrieved 23.02.05, 2005, from http://www.humboldt.edu/~tdd2/Autoethnography.htm

Easterby-Smith, M., & Araujo, L. (1999). *Organizational Learning: Current Debates and Opportunities*. London: Sage.

Easterby-Smith, M., Araujo, L., & Burgoyne, J. (Eds.). (1999). *Organizational Learning and the Learning Organization: Developments in Theory and Practice*. London: Sage.

Easterby-Smith, R., Thorpe, R., & Lowe, A. (1991). *Management Research: An Introduction*. London: Sage.

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of Educational Measurement*. New York: Prentice Hall.

Edgeworth, F. Y. (1890). The Elements of chance in Competitive Examinations. *Journal of the Royal Statistical Society*, 400-475, 644-463.

Ehrenberg, R. G., & Monk, J. (1999). U. S. News and World Report rankings: Why do they matter? *Change* (November/December 1999), 43-51.

eLearning Europa. (2004, 01.03.04). What is e-Learning? Retrieved 01.03.04, 2004, from www.elearningeuropa.info

Gale Group. (Eds.) (1999). *Future of the MBA in China (Master of Business Administration)*. California: Thomson

Entwistle, N. J. (1987). *Styles of Learning and Teaching* (2nd ed.). Chichester: Wiley.

---. (1992). *The Impact of Teaching on Learning Outcomes* (Rev. ed.). Sheffield: CVCP UCoSDA.

European Foundation for Management Development. (2004a, 01.08.04). The EQUIS Quality Standards. *What is EQUIS?* Retrieved 01.08.04, 2004, from http://www.efmd.be/

---. (2004b, 01.07.04). European Quality Improvement System EQUIS. *What is EQUIS?* Retrieved 01.07.04, 2004, from http://www.efmd.be/

---. (2004c, 01.08.04). European Quality Link (EQUAL) European MBA Guidelines. Retrieved 01.08.04, 2004, from http://www.efmd.be/

Ewell, P. T. (1984). *The self-regarding institution: Information for excellence*. Boulder, CO: National Center for Higher Education Management Systems.

---. (1987). Establishing a campus-based assessment program. In D. F. Halpern, ed. Student outcomes assessment: What institutions stand to gain? *New Directions for Higher Education, 59*, 9-24.

---. (1988). *Outcomes, assessment, and academic improvement: In search of usable knowledge. In J. C. Smart, ed. Higher Education: handbook of theory and research 4, 53-108*. New York: Agathon Press.

Ewell, P. T., & Jones, D. P. (1993). Actions matter: The case for indirect measures in assessing higher education's progress on the national educational goals. *Journal of General Education, 42*, 213-248.

Eysenck, M. W., Piper, D. W., & Richardson, J. T. E. (1987). *Student Learning, Research in Education and Cognitive Psychology*. Milton Keynes: SRHE & Open University.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. California: Stanford University Press.

Fisher, C. D., Schoenfeldt, L. F., & Shaw, J. B. (1991). *Human Resource Management*. Boston: Houghton Mifflin Co.

Fleenor, C. P., & Scontrino, M. F. (1982). *Performance Appraisal: A Manager's Guide*. Iowa: Kendal/Hunter.

Fletcher, J. J., Gentemann, K. M., & Potter, D. L. (1994). *Refocusing the academic program review on student learning. In M. K. Kinnick, ed. Providing useful information for deans and department chairs. New Directions for Institutional Research 84, 31-46*. San Francisco: Jossey-Bass.

Flynn, J. R. (2007). *What is Intelligence?: Beyond the Flynn Effect*. London: Cambridge University Press.

Foddy, W. (1997). *Constructing Questions for Interviews and Questionnaires*. London: Cambridge University Press.

Fong, B. (1988). *Assessing the departmental major. In J. H. McMillan, ed. Assessing students' learning. New Directions for Teaching and Learning 34, 71-83*. San Francisco: Jossey-Bass.

Fong, C., & Fong, J. (2004). The Business School - 'Business': Some Lessons from the US Experience. *Journal of Management Studies*.

Forestier, K. (2004). When only the best will do; Quality assurance is always important but becomes crucial in a matter as vital as the choice of a school. *South China Morning Post Good Schools Guide, Special Supplement Saturday* (September 11, 2004), 15.

Forrest, A., & A Study Group on Portfolio Assessment. (1990). *Time will tell: Portfolio assisted assessment of general education*. The AAHE Assessment Forum, American Association for Higher Education.

Gamson, Z. F., & Poulsen, S. J. (1989). Inventories of good practice: The next step for the seven principles for good practice in undergraduate education. *AAHE Bulletin, 42*, 7-8.

Gibbs, G. (1990). *Learning by Doing: A Guide to Teaching and Learning Methods*. Oxford: Oxford Brookes University.

Gill, W. E. (1993). *Conversations about accreditation: Middle States Association of Colleges and Schools: Focusing on outcomes assessment in the accreditation process*. Paper presented at the Double Feature Conference on Assessment and Continuous Quality Improvement, Chicago, IL.

Glaser, B. (1978). *Theoretical sensitivity*. Mill Valley, CA: The Sociology Press

Glaser, R. O. (1992). *Classic Readings in Self-Managing Teamwork: 20 of the Most Important Articles, King of Prussia*. Prussia, Organization Design & Development Inc.

Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, Aldine Publishing Company.

Glassner, B., Ksander, M., Johnson, B., & Berg, B. L. (1983). The Deterrence Effect of Juvenile versus Adult Jurisdiction. *Social Problems, 31*(2), 219-221.

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and Qualitative Design in Educational Research*. Orlando: Academic Press.

Gonzci, A. (2004). E-Mail Interview: UTS.

Graham, A., & Thompson, N. (2001). Broken ranks: US News college rankings measure everything but what matters. And most universities do not seem to mind. *The Washing Monthly* (September).

Green, D. (2007). Perspectives on Teaching and Learning English as a Global Language. *Education Today, 57*(2, Summer 2007), 2 - 7.

Gregory, M. (1994). Accrediting Work-Based Learning: Action Learning - A Model for Empowerment. *Journal of Management Development, 13*(4), 41-52.

Gronlund, N. E. (1988). *How to construct Achievement Tests*. New York: Prentice Hall.

Gumport, P. J. (1997). *Restructuring the Academic Environment*. San Francisco: Jossey-Bass.

Gundry, J. (2003, 01.01.04). How Flexible is E-Learning? Retrieved 01.01.04, 2004, from http://www.knowab.co.uk/elflexible

Guthrie, D. S., Jones, E. A., Oehler, D., & Ratcliff, J. L. (1991). *The effect of coursework patterns, advisement, and course selection on the development of general learned abilities of college graduates.* University Park: The Pennsylvania State University, National Center on Postsecondary Teaching, Learning, and Assessment.

Hall, D. T., & Goodale, J. G. (1990). *Human Resource Management: Strategy, Design, and Implementation*. Glenview, IL: Scott Foresman.

Halpern, D. F. (1987). Recommendations and caveats. In D. F. Halpern, ed. Student outcomes assessment: What institutions stand to gain. *New Directions for Higher Education, 59*, 109-111.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research, 52*, 123-142.

Haper, C., & Kember, D. (1989). Interpretation of Factors Analyses from the Approaches to Studying Inventory. *British Journal of Educational Psychology*(59), 66-74.

Hartrog, P., & Rhodes, E. C. (1935). *An Examination of Examinations*. London: Macmillan.

---. (1936). *The Marks of Examiners*. London: Macmillan.

Hattie, J., & Watkins, D. (1983). A Longitudinal Study of the Approaches to learning of Australian Tertiary Students. *Human Learning*(4), 127-141.

Hatton, N., & Smith, D. (1995). Reflection in Teacher Education: Towards Definition and Implementation. 2010, from http://www2.edfac.usyd.edu.au/LocalResource/Study1/hattonart.html

HEFCE. (1993). *Assessment of the Quality Circular 3/93*. Bristol: Higher Education Funding Council.

Hoggatt, R. S., & Hazel, J. T. (1970). *Reliability of Individual versus Group Job Pay Rating*. Texas: USAF.

Holland, T. (2004). Research Literacies [Video Aid]. Sydney: UTS.

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Massachusetts: Addison-Wesley.

Huang, Y. (2004). *New Management Dictionary*. Beijing: Beijing Democracy & Construction Publisher.

Hughes, M., & Vass, A. (2001). *Strategies for Closing the Learning Gap*. Stafford: Network Educational Press Ltd.

Hutchings, P. (1989). *Behind outcomes: Contexts and questions*. Paper presented at the AAHE Assessment Forum.

Jaeger, K., & Newble, D. I. (1983). The Effects of Assessment and Examinations on the Leaning of Medical Students. *Medical Education*(17), 25-31.

Jennings, L. (1993). *An Introduction to Qualitative Research*. Lismore: University of New England.

Johnson, R., McCormick, R. D., Prus, J. S., & Rogers, J. S. (1993). *Assessment Options for the College Major. In T. W. Banta and Associates, eds. Making a Difference: Outcomes of a Decade of Assessment in Higher Education, 151-67*. San Francisco: Jossey-Bass.

Jones, E. A., Ratcliff, J. L., & et al. (1997). *Turning results into improvement strategies*. University Park: The Pennsylvania State University, National Center on Postsecondary Teaching, Learning and Assessment.

Jones, L. (2004). Lies, Damned Lies and University Ranking Tables. *Business Executive, Autumn 2004*(Number 93), 16-17.

Kellaghan, T., & Stufflebeam, D. L. (Eds.). (2002). *International Handbook of Educational Evaluation (Springer International Handbooks of Education)*. MA: Kluwer Academic Publishing.

Kennedy, A. (2007). A Culture of Learning: Where East meets West. *Education Today, 57*(2, Summer 2007), 8 - 13.

Keogh, B., & Naylor, S. (1999). Concept Cartoons, Teaching and Learning in Science: An Evaluation. *International Journal of Science Education, 21*(4), 431-446.

Keys, L. (1994). Action Learning: Executive Development of Choice for the 1990s. *Journal of Management Development, 13*(8), 50-56.

Kinchin, I. M. (2002). Why Professional Development should challenge Teachers' Core Beliefs. *School Science Review, 84*(306), 119-121.

---. (2003). Effective Teacher-Student Dialogue: A Model of Practice from Biological Education. *Journal of Biological Education, 37*(3), 110-113.

---. (2004). Achieving a Shared Understanding of the Learners' Role through Teacher-Student Dialogue. *Education Today, 54*(Number 1, March 2004), 28-30.

Klein, S., Koretz, D., McCaffrey, D., & Stecher, B. (1994). The Vermont portfolio assessment program: Findings and implications. *Educational Measurement: Issues and Practices, 13*(3), 5-16.

Koo, L. C. (1999a). Learning Action Learning. *Journal of Workplace Learning, 11*(Number 3),

89-94.

---. (1999b). Learning Action Learning [Training Note]. HKSAR: International Management Centers.

Koul, B. N. (2006). Towards a Culture of Quality in Open Distance Learning: Present Possibilities. In B. N. Koul & A. Kanwar (Eds.), *Perspectives on Distance Education: Towards a Culture of Quality* (pp. 177-187). Vancouver: Commonwealth of Learning.

Koul, B. N., & Kanwar, A. (Eds.). (2006). *Perspectives on Distance Education: Towards a Culture of Quality*. Vancouver: Commonwealth of Learning.

Kuh, G. D. (2001). Assessing what really matters to student learning. *Change, 33*(3), 10-17.

Lalitpat, P. (2000, June 9, 2000). Former PM wants changes made to teaching methods. *Bangkok Post,* p. 3.

Lam, W. Y. (2010, 15.01.10). EMBA blending Classical-and-Contemporary and East-and-West Wisdom. *Job Market, 448*.

Lee, N. (1980). *Tertiary Education and Planning in Hong Kong*. Hongkong: Amazon.

Lee, N., & Lam, A. (1994). *Professional and Continuing Education in Hong Kong*. Hongkong: Hong Kong University Press.

Leedy, P. D. (1994). *Practical Research Planning and Design* (5th ed.). New York: MacMillan Publisher.

Leithwood, K., & Day, C. (2008). The Impact of School Leadership on Student Outcomes (Special Issue). *School Leadership and Management, 28*(1).

Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). *Review of Research: How Leadership influences Student Learning*. New York: The Wallace Foundation.

Lenn, M. P. (1992). Global Trends in Quality Assurance in Higher Education. *World Education News & Reviews* (5)2, 1-21.

Lenning, O. T. (1988). *Use of noncognitive measures in assessment. In T. W. Banta, ed. Implementing outcomes assessment: Promise and perils. New Directions for Institutional Research 59, 41-51*. San Francisco: Jossey-Bass.

Leung, K., & White, S. (Eds.). (2007). *Handbook of Asian Management*. New York: Kluwer Academic.

Levine, D. U., & Ornstein, A. C. (2003). *Foundations of Education* (8th ed.). New York: Houghton Mifflin Company.

Liang, J., Noble, J., & Swayer, R. (1989). Accuracy of self-reported activities and accomplishments college-bound seniors. *Journal of College Student Development, 29*, 362-368.

Lincoln, Y. & Guba, E. (1985). *Naturalistic Inquiry*. Newbury Park, CA, Sage Publications.

Lodge, M., & Tursky, B. (1981). The Workshop on the Magnitude Scaling of Political Opinion in Survey Research. *American Journal of Political Science*(25), 376-419.

Lohninger, H. (1999). *Teach me Data Analysis*. Berlin–New York–Tokyo: Springer-Verlag.

Lowman, R. L., & Williams, R. E. (1987). Validity of self-ratings of abilities and competencies. *Journal of Vocational Behavior, 31*, 1-13.

Machung, A. (1995). *Changes in college rankings: How real are they?* Paper presented at the 35th Annual AIR Forum, Boston, MA.

Madaus, G. F., & Stufflebeam, D. L. (Eds.). (1988). *Education Evaluation: Classic Works of Ralph W. Tyler (Evaluation in Education and Human Services)*. MA: Kluwer Academic Publishing.

Manchester Business School. (2004a). The Difference Learning MBA. *Human Resources*(May 2004), 3.

---. (2004b). The Difference Learning MBA. *Banking Today*(March/April 2004), 9.

---. (2004c). The Difference Learning MBA for Accountants, Company Secretaries, Fund Managers, Bankers and Management Accountants. *Hong Kong Securities*(July/August 2004), 9.

Margolis, J., Walsh, J., & Weber, K. (2003). Social Issues and Management: Our Lost Cause Found. *Journal of Management*.

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School Leadership that Works: From Research to Results*. VA: Association for Supervision and Curriculum Development.

Maxwell, T. W., & Shanahan, P. J. (1996, 16-18 October 1996). *Reconceptualizing the Doctorate: Issues arising from Comparative Data on the Doctor of Education in Australia.* Paper presented at the Which Way for Professional Doctorates?, Coffs Harbor.

McCormick, E. J. (1979). *Job Analysis: Methods and Applications* (1st ed.). New York: AMACOM.

Mccormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A Study of Job Characteristics and Job Dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*(56), 347-368.

Mcguire, M. D. (1995). *Validity issues for reputational studies. In Walleri, R. D. and M. K. Moss, eds. Evaluating and responding to college guidebooks and rankings. New Directions for Institutional Research 88, Winter.* San Francisco: Jossey-Bass.

McInnis, C. (1995). *Less Control and More Vocationalism: The Australian and New Zealand Experience*. Buckingham: The Society for Research into Higher Education and Open University Press.

Median, A. (1981). *The Appraisal of Managerial Performance.* New York: AMACOM.

Merriam, S. B. (1988). *Case Study Research in Education: A Qualitative Approach.* San Francisco: Jossey-Bass.

Michael, R. K., Sower, V. E., & Motwani, J. (1997). A Comprehensive Model for Implementing Total Quality Management in Higher Education. *Benchmarking: An International Journal, 4*(2), 104-120.

Miles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis: An Expanded Sourcebook (2nd Edition). Thousand Oaks: Sage Publications.

Mills, R. (2006). Quality Assurance Distance Education towards a Culture of Quality: A Case Study of the Open University, United Kingdom (OUUK). In B. N. Koul & A. Kanwar (Eds.), *Perspectives on Distance Education: Towards a Culture of Quality* (pp. 135-148). Vancouver: Commonwealth of Learning.

Morgan, C., & O'neilly, M. (2002). *Assessing Open and Distance Learners*. London, England: Kogan Page.

Mulford, B. (2006). Leading Change for Student Achievement. *Journal of Educational Change, 7*, 47-58.

Mulford, B., & Silins, H. (2003). Leadership for Organizational Learning and Improved Student Outcomes: What do we know? *Cambridge Journal of Education, 33*, 175-179.

Mumford, A. (1995). Learning in Action. *Industrial and Commercial Training, 27*(8), 36-40.

National Center for Higher Education. (1994). *A preliminary study of the feasibility and utility for national policy of instructional and good practice indicators in undergraduate education.* (No. Contractor Report for the National Center for Education Statistics.). Boulder, CO: National Center for Higher Education Management Systems.

---. (1996). *The national assessment of college student learning: An inventory of state-level assessment activities.* Boulder, CO: National Center for Higher Education Management Systems.

National Opinion Research Center. (1987). *A review of the methodology for the U. S. News & World Report's rankings of undergraduate colleges and universities.*

National Postsecondary Education Cooperative. (2000a). *The NPEC sourcebook on assessment, volume 1: Definitions and assessment methods for critical thinking, problem solving, and writing.* Harrisonbury, VA: Center for Assessment and Research Studies, James Madison University under the sponsorship of the National Center for Education Statistics, U. S. Department of Education.

---. (2000b). *The NPEC sourcebook on assessment, volume 2: Selected institutions utilizing assessment results.* Harrisonbury, VA: Center for Assessment and Research Studies, James Madison University under the sponsorship of the National Center for Education Statistics, U. S. Department of Education.

Nettles, M. T., Cole, J. J. K., & Sharp, S. (1997). *Assessment of teaching and learning in higher*

education and public accountability. Stanford, CA: National Center for Postsecondary Improvement.

Newstead, S., & Dennis, I. (1994). Examiners examined: The Reliability of Exam Marking in Psychology. *The Psychologist, 7*, 216-219.

Newton, R. R., & Rudestam, K. E. (2001). *Surviving Your Dissertation A Comprehensive Guide to Content and Process* (2nd ed.). London: Sage Publications.

Nguyen, P., Elliott, J. G., Terlouw, C., & Pilot, A. (2009). Neocolonialism in Education: Cooperative Learning in Asian Context. *Comparative Education, 45*, 109-130.

Noble, K. (1994). *Changing Doctoral Degrees: An International Perspective*. Buckingham: Society for Research into Higher Education and Open University Press.

Nosich, G. M. (1982). *Reasons and Arguments*. California: Wadsworth.

Obler, S. S., Slark, S. J., & Umbdenstock, L. (1993). *Classroom assessment. In T. W. Banta and Associates, eds. Making a difference: Outcomes of a decade of assessment in higher education, 211-26*. San Francisco: Jossey-Bass.

O'Neil, J. (1996). A Study of the Role of Learning Advisers in Action Learning. *Employee Counseling Today, 8*(6), 39-44.

Ouiment, J. A., Carini, R. M., Kuh, G. D., & Bunnage, J. C. (2001). *Using Focus Group to establish the Validity and Reliability of a College Student Survey*. Paper presented at the 2001 AIR Forum, Long Beach, CA.

Pace, C. R. (1985). *The credibility of students self-reports*. Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles.

---. (1990). *The undergraduates: A report of their activities and progress in college in the 1980s*. Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles.

Parkes, J., & Suen, H. K. (1996). *Challenges and opportunities for student assessment in distance education*. Paper presented at the Distance Education Online Symposium, On-line Serial.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.

Pearson, M. (1999). The Changing Environment for Doctoral Education in Australia: Implications for Quality Management, Improvement and Innovation'. *Higher Education & Development, 18*(No. 3), 269-287.

Perry, C. (1999, 01.01.99). How to Write a Doctoral Thesis - PhD/DPhil. Retrieved 01.02.99, 1999, from http://www.imc.org.uk/imc/news/occpaper/cpindex.htm

Peters, J. (1996). A Learning Organization's Syllabus. *The Learning Organization, 3*(1), 4-10.

Peterson, M., Augustine, C. H., Einarson, M. K., & Vaughan, D. S. (1999a). *Designing student*

assessment to strengthen institutional performance in associate of arts institutions. Stanford, CA: National Center for Postsecondary Improvement.

---. (1999b). *Designing student assessment to strengthen institutional performance in comprehensive institutions*. Stanford, CA: National Center for Postsecondary Improvement.

---. (1999c). *Designing student assessment to strengthen institutional performance in doctoral and research institutions*. Stanford, CA: National Center for Postsecondary Improvement.

Peterson, M. W. (1987). *Academic management practices survey for the research program on the organizational context for teaching and learning.* Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning, The University of Michigan.

Philipson, R. (1992). *Linguistic Imperialism*. Oxford: OUP.

Phillips, J. L. J. (1992). *How to Think About Statistics*. New York: W. H. Freeman and Company.

Phillips, M. E. (1993). *How to get a PhD*. London: Open University Press.

Pike, G. R. (1995). The relationship between self reports of college experiences and achievements and achievement test scores. *Research in Higher Education, 36*, 1-22.

Pike, G. R. (1999). The constant error of the halo in educational outcomes research. *Research in Higher Education, 40*, 61-86.

Pohlman, J. T., & Beggs, D. L. (1974). A Study of the Validity of Self-Reported Measures of Academic Growth. *Journal of Educational Measurement, 11*, 115-119.

Poole, M., & Spear, R. H. (1997). *Policy Issues in Postgraduate Education: An Australian Perspective*. Buckingham: Society for Research into Higher Education and Open University Press.

Preston, N., & Symes, C. (1992). *Schools and Classrooms: A Cultural Studies Analysis of Education*. Melbourne: Longman Cheshire.

Queensland University of Technology. (1998). *Teaching and Learning Plan: 1998-2000*. Queensland: QUT.

Quinn, F. M. (2000). *The Principles and Practice of Nurse Education.* Cheltenham: Stanley Thornes.

Raelin, J. A. (1997). Action Learning and Action Science: Are They Different? *Organizational Dynamics, 26*(1), 21-34.

Ramsden, P. (Ed.). (1988). *Improving Student Learning*. London: Kogan Page.

---. (1992). *Learning to Teach in Higher Education*. London: Kogan Page.

RAND Corporation. (2004, 08.06.04). About RAND. Retrieved 01.07.04, 2004, from http://www.rand.org/about/

Ratcliff, J. L., Jones, E. A., Guthrie, D. S., & Oehler, D. (1991). *The Effect of Coursework Patterns, Advisement, and Course Selection on the Development of General Learned Abilities of College Graduates*. University Park: The Pennsylvania State University, National Center on Postsecondary Teaching, Learning, and Assessment.

Rea-Dickins, P. & Germaine, K. (Eds). (2003). *Evaluation*. London: Oxford University Press.

Revans, R. W. (1980). *Action Learning*. London: Blond & Briggs Limited.

Riding, R. (2002). *School Learning and Cognitive Style*. London: David Fulton.

Ries, P., & Thurgood, D. (1993). *Summary Report 1991: Doctoral Recipients from United States Universities*. Washington, DC: National Academy Press.

Riggs, M. L., & Worthley, J. S. (nd). Baseline Characteristics of Successful Program of Student Outcomes Assessment.

Rowntree, D. (1987). *Assessing Students - How shall We know Them?* London: Harper & Row.

Sandelands, E. (1998). Creating an Online Library to Support a Virtual Learning Community. *Internet Research: Electronic Networking Applications and Policy, 8*(1), 77-80.

Schön, D. A. (1983). *The Reflective Practitioner*. New York: Basic Books.

---. (1988). *Educating the Reflective Practitioner*. San Francisco: Jossey-Bass.

---. (1991). *The Reflective Turn: Case Studies In and On Educational Practice*. New York: Teachers Press.

Scott, W. E. J. (1963). *The Reliability and Validity of a Six-Factor Job Evaluation System*. New York: Purdue University.

Serpell, R. (2007). Bridging between Orthodox Western Higher Educational Practices and An African Sociocultural Context. *Comparative Education, 43*, 23-51.

Seymour, D. (1992). *Q-Causing Quality in Higher Education*. New York: American Council on Education / MacMillan Publishing.

Shanahan, P. (1997). Towards a Reconceptualization of the Doctorate: Issues arising from Comparative Data relating to the EdD Degree in Australia. *Studies in Higher Education, 22*(2), 15.

Shinkfield, A. J., & Stufflebeam, D. L. (1996). *Teacher Evaluation: Guide to Effective Practice (Evaluation in Education and Human Services)*. MA: Kluwer Academic Publishing.

Simpson, R. (1983). *How the PhD came to Britain: A Century of Struggle for Postgraduate Education*. Guildford: Society for Research into Higher Education.

Sing Tao Education. (2010). Joint Admission Reform leads to a Surge of Enrolment in

Business Administration Degree Courses. *Sing Tao News,* p. F3.

Slade, C. (2003). *Form and Style* (12th ed.). New York: Houghton Mifflin Company.

Smith, K. W., & Wengland, E. J. (1993). *Survey responses: An evaluation of their validity*. New York: Academic Press.

Sotto, E. (1994). *When Teaching becomes Learning. A Theory and Practice of Teaching*. London: Continuum.

Sou, G. (2000). Action Learning for Modern Managers. *International Journal of Action Learning, 1*(5), 141-145.

---. (2001a). Action Learning for Life. *Multi-Skills, 2*(1), 3-4.

---. (2001b). Action Learning for Life. *Professional Business & Technical Management, 1*(3), 9-10.

---. (2001c). Action Learning for Life. *The Management Specialists, 4*(2), 6-7.

---. (2001d). *An Examination of Performance Management System in Customs & Excise Service of Hongkong SAR*. Unpublished Doctoral, Southern Cross University, New South Wales.

---. (2003, 11.03.03). Action Learning for Modern Managers. Retrieved 11.03.03, 2003, from http://www.iiuedu.ie/download/Articles/article-69.doc

---. (2004a). An Examination of Performance Management System in Customs & Excise Service of Hong Kong Special Administrative Region. *Action Learning Harvest, 4*(1).

---. (2004b, 20.04.05). Professional Doctorates for Practitioners. Retrieved 20.04.05, 2004, from http://www.iiuedu.ie/download/Articles/article-80.doc

---. (2004c). Professional Doctorates in Australia. *Continuing Professional Development, 4*, 1-12.

---. (2004d). Real Value of e-Education. *Virtual University Journal, 4*(2), 1-7.

---. (2004e, 20.05.04). Real Value of e-Education. Retrieved 20.05.04, 2004, from http://www.iiuedu.ie/download.php?v=Articles/article-78.doc

---. (2004f, June 24 - 25, 2004). *Real Value of e-Education*. Paper presented at the International Conference on e-Education 2004: Review and New Perspectives, Macau SAR.

---. (2004g). Real Value of e-Education [Powerpoint Presentation]. Macau SAR: Sinocan Group.

---. (2004h). Real Value of e-Education - Abstract. *e-Education Conference 2004 - Abstracts*.

---. (2005a, June 20 - 22, 2005). *Quality Assurance Aid in Graduate Management Education of the Mainland China (Abstract)*. Paper presented at the 3rd Conference on Research in Distance and Adult Learning and The 4th Conference on Distance Education in

China (CRIDALA/DEC 2005), Hong Kong Special Administrative Region.

---. (2005b, June 20 - 22, 2005). *Quality Assurance Aid in Graduate Management Education of the Mainland China (Full Paper).* Paper presented at the 3rd Conference on Research in Distance and Adult Learning and The 4th Conference on Distance Education in China (CRIDALA/DEC 2005), Hong Kong Special Administrative Region.

---. (2005c). Quality Assurance Aid in Graduate Management Education of the Mainland China (Presentation). On *The 3rd Conference on Research in Distance and Adult Learning and The 4th Conference on Distance Education in China (CRIDALA/DEC 2005)* [Powerpoint Presentation]. Hong Kong Special Administrative Region: The Open University of Hong Kong.

---. (2005d, June 21,2005). *Quality Management System for a University under ISO 9001:2000.* Paper presented at the 11th Hong Kong Quality Management & 1st Six Sigma Convention: World Class Excellence through Quality, Hongkong, China.

---. (2005e, 04.08.05). 'Western Business Suit' of 'MBA' Brand for Occupational Managers of the Modern China. Retrieved from http://www.paper.edu.cn/index.php/default/scholar/downpaper/chouzhicheng-1

---. (2005f). Orientation of Management Education for Working Adults in the Mainland China. *China Higher Education Journal, 1*(49).

---. (2005g, 25.05.05). Modern Enterprises in Consideration of Cs. Retrieved from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200505-134

---. (2005h). Footprints of Western Management Education onto the Mainland China. *China Modern Education Teaching Magazine, 13*(63), 12-13.

---. (2006a, December 6–8, 2006). *Evaluation of an MBA Degree Program of an External University for the Career Managers in Mainland China.* Paper presented at the 4th Asia Pacific Conference on Continuing Education & Lifelong Learning, The University of Hong Kong.

---. (2006b). Evaluation of an MBA Degree Program of an External University for the Career Managers in Mainland China. In *Transnational Education in a Globalizing World: Abstracts of Papers* (Vol. 2-2, pp. 12-13). HKSAR: The University of Hong Kong.

---. (2006c, 17.02.06). Quality Assurance Aid in Graduate Management Education of Mainland China. Retrieved 02.03.06, 2006, from http://asiapacific-odl.oum.edu.my/C17/F317.pdf

---. (2008). Historical Review of Program Evaluation for Educational Quality Assurance. Retrieved 13.08.08, from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200808-173

---. (2011). A Management Education Model for the Chinese Career Managers.

International Conference of Learning Society and Open Education Research, Macau, City University of Macau & Universidale Alberta Internacional da Asia (Macau).

Sou, G., & Chow, S. K. L. (2007, 11.01.07). Use of GTA in Qualitative Research. Retrieved 11.01.07, 2007, from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200701-121

Sou, G., & Zhou, P. Q. (2005a, 01.02.05). Quality Management Model for an External MBA Degree Program. Retrieved from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200502-7

---. (2005b, 17.02.05). Curricular Quality Appraisal Model for MBA Degree Cohort Program in the Mainland China. Retrieved 17.02.05, 2005, from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200502-40

---. (2005c, 18.01.05). Curriculum Design of an External MBA Degree Program for the Working Adults in the Mainland China. Retrieved 18.01.05, 2005, from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200501-61

---. (2005d). Curriculum Design of an External MBA Degree Program for the Working Adults in the Mainland China. *New Horizons in Education, 51*(May 2005), 40-55.

---. (2005e, 04.04.05). Curricular Quality Appraisal Tool. Retrieved 04.04.05, 2005, from http://www.paper.edu.cn/index.php/default/releasepaper/downPaper/200504-3

---. (2007). Management Education Program Evaluation: An Empirical Study in Mainland China. *New Horizons in Education, 55*(No. 2, Oct 2007), 12-27.

Stake, R. (1995). *The Art of Case Study Research*. California: Sage Publications.

Steele, J. M., & Lutz, D. A. (1995). *Report of ACT's research on postsecondary assessment needs*. Iowa City, IA: American College Testing Program.

Stevens, S. (2007). Guanxi and the Art of Doing Business in China. *IET Member News, August 07*(6), 7.

Strauss, A. (1987). *Qualitative Analysis for Social Scientists*. New York: Cambridge University Press.

Strauss, A. and J. Corbin (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage Publications.

---. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Newbury Park, CA: Sage Publications.

Stufflebeam, D. L. (2000). *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*. Hingham: Kluwer Academic Publishers.

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (Eds.). (2000). *Evaluation Models - Viewpoints on Educational and Human Services Evaluation* (2nd ed. Vol. 49). MA: Kluwer Academic Publishing.

Stufflebeam, D. L., McCormick, C. H., Brinkerhoff, R. O., & Nelson, C. O. (1985). *Conducting Educational Needs Assessment (Evaluation in Education and Human Services)*. MA: Kluwer Academic Publishing.

Stufflebeam, D. L., & Shinkfield, A. J. (1984). *Systematic Evaluation: A Self-Instructional Guide to Theory and Practice (Evaluation in Education and Human Services)*. MA: Kluwer Academic Publishing.

Swan, W. (1991). *How to Do a Superior Performance Appraisal.* New York: John Wilney & Sons.

Tan, O. S., McInerney, D. M., Liem, A. D., & Tan, A. G. (Eds.). (2008). *What the West can learn from the East: Asian Perspectives on the Psychology of Learning and Motivation*. Charlotte: Information Age.

Tang, A. Y. (2007, 20.09.07). Cross-Border Experience vital for Managers. *China Daily - Hong Kong Edition,* p. 11.

Taylor, S. & Bogdan, R. (1984). *Introduction to Qualitative Research Methods: The Search for Meaning*. New York: John Wiley & Sons.

Taylor, M., & McGimpsey, I. (2008). Learning to Love Ideas. *Education Today, 58*(1, Spring 2008), 2-4.

Terenzini, P. T. (1989). Assessment with open eyes: Pitfalls in studying student outcomes. *Journal of Higher Education, 60*, 644-664.

Terenzini, P. T., & Wright, T. (1987). Influences on students' academic growth during four years of college. *Research in Higher Education, 26*, 161-179.

Tesch, R. (1990). *Qualitative Research: Analysis Types and Software Tools*. New York: Falmer.

The Association of MBAs. (2004, 02.09.04). Accreditation / Why accredit MBA Programmes. Retrieved 02.09.04, 2004, from http://www.mbaworld.com

The Association to Advance Collegiate Schools of Business. (2004, 02.09.04). Accreditation. Retrieved 02.09.04, 2004, from http://www.aacsb.edu/accreditation/default.asp

---. (2004a, 02.09.04). Accreditation Standards. Retrieved 02.09.04, 2004, from http://www.aacsb.edu/accreditation/standards.asp

The Economist. (2004, 20.05.04). But can you teach it? Retrieved 20.05.04, 2004, from http://www.economist.com

Theodorakou, K., & Zernvas, Y. (2003). The Effects of the Creative Movement Teaching Methods and the Traditional Teaching Method on Elementary School Children's Self-esteem. *Sport, Education and Society, 8*(1), 91-104.

Theodorson, G. A. & Theodorson, A. G. (1979). *A Modern Dictionary of Sociology*. New York: Barnes & Noble Books.

Thompson, N. (2000). Playing with numbers: How U. S. News mismeasures higher education and what we can do about it. *Washington Monthly*(September).

Tjosvoid, D., Wong, A., & Hui, C. (2007). *Leadership Research in Asia: Developing Relationships*. New York: Kluwer Academic.

Toporek, R. L., Gerstein, L. H., Fouad, N. A., Roysircar, G. and Israle, T. (Eds). (2006). *Handbook for Social Justice in Counseling Psychology – Leadership, Vision, and Action*. New Delhi: Sage Publications.

Tsang, P. (2005, January 14 - 15, 2005). *ISO 9001:2000 and Quality Assurance of Distance Learning Institutes*. Paper presented at the First International Conference on Quality Management and Six Sigma, Hong Kong Special Administrative Region, People's Republic of China.

Tsui, A. S. (2010). In Pursuit of High Quality Research. *The Journal of The International Association for Chinese Management Research, 6*(1), 149-150.

Turner, C. F., & Martin, E. (Eds.). (1984). *Surveying subjective phenomena* (Vol. 1). New York: Russell Sage Foundation.

University of Technology Sydney. (1999, 30.04.04). The Professional Doctorate at UTS - Questions. Retrieved 30.04.04, 2004, from http://www.gradschool.uts.edu.au/about/aboutprocess/profdoctorquestions.html

---. (2001, 30.04.04). Research Education at UTS: A Distinctive Approach. Retrieved 30.04.04, 2004, from http://www.gradschool.uts.edu.au/about/aboutprocess/futureresearch.html

---. (2004a, 17.09.04). Guide to Writing a Research Proposal. Retrieved 17.09.04, 2004, from http://www.education.uts.edu.au/research/degrees/guide.html

---. (2004b, 30.04.04). The Professional Doctorate at UTS: A Discussion Paper. Retrieved 30.04.04, 2004, from http://www.gradschool.uts.edu.au/about/aboutprocess/professionaldr.html

Vandament, W. E. (1987). A state university perspective on student outcomes assessment. In D. F. Halpern, ed. Student outcomes assessment: What institutions stand to gain. *New Directions for Higher Education,, 59*, 25-28.

Vaus, D. A. d. (1996). *Surveys in Social Research* (4th ed.). London: UCL Press.

Vogt, P. W. (1993). *Dictionary of Statistics and Methodology: A Non-Technical Guide for the Social Sciences*. London: Sage Publications Ltd.

Wallis, E. (1999). *QUT Staff Teaching and Learning Development Programme: Student Centered Learning*. Queensland: QUT Department of Teaching and Learning Support Services.

Waluconis, C. J. (1993). *Student self-evaluation. In Trudy Banta, ed. Making a difference: Outcomes of a decade of assessment in higher education, 244-55*. San Francisco: Jossey-Bass.

Wang, H. (2006). Teaching Asian Students Online: What matters and why? *PAACE Journal of Lifelong Learning, 15*, 69-84.

Wax, R. (1971). *Doing Fieldwork: Warnings and Advice*. Chicago: University of Chicago Press.

Wentland, E. J. & Smith, K. W. (1993). *Survey responses: An evaluation of their validity*. New York: Academic Press.

Wilkinson, C. (2003). The Effects of a School Physical Education and Health WebSite on Parental Knowledge of the Program. *Physical Educator, 60*(3, Fall 2003), 162-168.

Winfield, G. (1987). *The Social Science PhD; the ESRC Inquiry on Submission Rates*. London: Economics and Social Research Council.

Winser, W. N. (1996). Literacy Development and Teaching East and West: Culture and Context-Text Relations. *Hong Kong Journal of Applied Linguistics, 1*(1), 19-37.

Winter, D. G., McClelland, D. C., & Steward, A. J. (1981). *A new case for the liberal arts*. San Francisco: Jossey-Bass.

Wittrock, M. C. (Ed.). (1986). *Handbook of Research on Teaching* (3rd ed.). New York: Macmillan Publishing Co.

Wolverton, M. (2003). *A New Alliance: Continuous Quality and Classroom Effectiveness*. Washington, DC: George Washington University.

Wong, T. M. (Ed.). (2010). *2010 Postgraduate Plus Guidebook*. HKSAR: Ming Pao News.

Xia, L. B. (2009, 04.12.09). University Ranking 2009. *Jiu Jik,* A14.

Zhang, F. F. (Eds.) (2005). Qualitative Data Analysis: An Expanded Sourcebook. Taipei: Double Leaf.

Zhou, P. Q. (2006). *Quality Assessment of the Master of Business Administration Degree Program of Asia International Open University (Macau) in the Mainland China.* Unpublished Doctoral, Asia International Open University (Macau), Macau Special Administrative Region.

Zhou, P. Q., & Sou, G. (2005a, March 28 - 30, 2005). *Quality Management Model for Higher Education.* Paper presented at the 10th International Conference on ISO9000 and TQM (10-ICIT): TQM & Six Sigma for Competitiveness, Shanghai, People's Republic of China.

---. (2005b, March 28 - 30, 2005). *Quality Management Model for Higher Education.* Paper presented at the 10th International Conference on ISO9000 and TQM and The 5th Shanghai International Symposium on Quality: TQM & Six Sigma for Competitiveness, Shanghai, People's Republic of China.